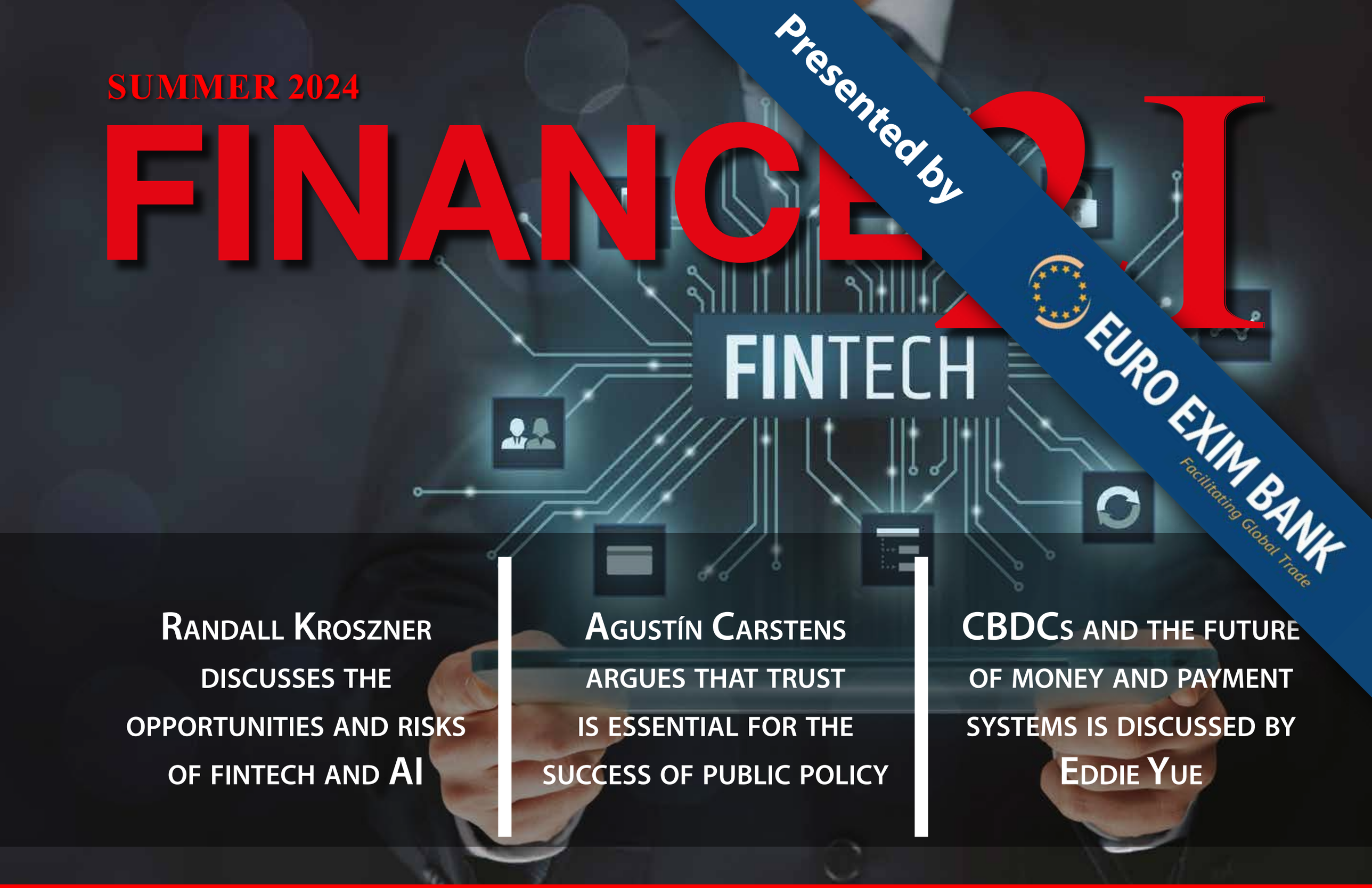
SUMMER 2024

FINANCE 21

Presented by

EURO EXIM BANK

Facilitating Global Trade



RANDALL KROSZNER DISCUSSES THE OPPORTUNITIES AND RISKS OF FINTECH AND AI

AGUSTÍN CARSTENS ARGUES THAT TRUST IS ESSENTIAL FOR THE SUCCESS OF PUBLIC POLICY

CBDCS AND THE FUTURE OF MONEY AND PAYMENT SYSTEMS IS DISCUSSED BY EDDIE YUE

21ST CENTURY FINANCE

Foreword

Welcome to the Summer edition of **FINANCE21**. This publication has been prepared in response to readership demand for an overview of the financial sector in these turbulent and unique times.

All aspects of the sector are examined, with the most respected authors providing the reader with the most comprehensive information available. Our brief is to provide all the data necessary for the readership to make their own informed decisions. All editorials are independent, and content is unaffected by advertising or other commercial considerations. Authors are not endorsing any commercial or other content within the publication. ■

CONTENTS

CONTENTS

CONTENTS

CONTENTS

CONTENTS

CONTENTS

CONTENTS

CONTENTS

WHY BVI?

- Compliance with international regulatory standards
- Competitive start-up costs
- Innovative legislation
- Internationally renowned commercial court
- No currency controls
- Qualified professional pool of practitioners
- Strong partnership between public and private sectors

JURISDICTION OF CHOICE

BRITISH VIRGIN ISLANDS

Pioneering, innovative and leading the way in global business solutions, the **British Virgin Islands (BVI)** is an internationally respected business and finance centre with a proven committment to connect markets, empower clients and facilitate investment, trade and capital flow.

BVI Finance
your international business partner

3rd Floor, Cutlass Tower, Road Town, Tortola, BVI VG1110
T: +1 (284) 852-1957 E: info@bvifinance.vg
W: www.bvifinance.vg | www.bviglobalimpact.com

A virtuous circle

Trust is essential for the success of public policies. Agustín Carstens argues that policymakers' success in dealing with recent crises was due to the trust they had built over the years, which allowed them to take decisive action when required

I will use my time to share some lessons learned during my career as a policymaker, which has now spanned more than four decades. At various times, I have had a front row seat during episodes of economic and financial crisis. Early in my career at the Bank of Mexico, I witnessed the debt crisis that hit many Latin American countries in the early 1980s and the 1994 'Tequila crisis' in Mexico.

The Great Financial Crisis (GFC), which exploded in 2008, coincided with my tenure first as Mexican Finance Minister, and then as Governor of the central bank. The COVID-19 pandemic and its aftermath took place while I was in my current role at the BIS.

The crises of the 1980s and 1990s were felt most keenly in emerging market economies. Some considered advanced economies immune from such events. But the more recent crises have hit advanced economies as hard as emerging markets, if not harder. Indeed, in many respects the GFC had its epicentre in the most mature and largest advanced economies.

Crises often prompted important changes in policy frameworks. In many emerging markets, the crises of the 1980s and 1990s led to the adoption of more flexible exchange rate regimes, greater central bank independence from governments, a heightened focus on low inflation as the main monetary policy objective, fiscal discipline and a remarkable improvement in banking sector supervision and regulation.

Some jurisdictions also saw major structural reforms, including the liberalisation of product and labour markets and privatisation of public enterprises. The GFC and the subsequent period of low inflation led to a broadening of central bank policy instruments, particularly in advanced economies, including the introduction of quantitative easing and forward guidance. Needless to say, the recent inflation surge has prompted further reassessments in many policy dimensions.

A key development that has taken place over the course of my career is that economies have become significantly more integrated. They trade more with each other. And, due to the emergence of global value chains, they trade more intensively. We felt this intensely in Mexico, particularly after the introduction of the North American Free Trade Agreement (NAFTA) in 1994.

Today, I see this international trade integration from the vantage point of a resident of Basel, which lies at the intersection of Switzerland, France and Germany. And it is very evident here in Frankfurt, at the heart of the manufacturing powerhouse of Europe.

Building resilient and robust economies and financial systems is the best way to ensure that policies remain effective, so that they can be deployed when they are needed the most

More generally, the global economy has become extremely dynamic. Information flows are unfettered, and firms and consumers very sophisticated. Even more importantly, as compared with when I began my career, financial markets have become much larger, more integrated and more fast-paced.

A side effect of all this is that the scope to sustain fundamentally flawed policy frameworks has diminished. In particular, global financial markets' immense size and speed discipline policymakers and, at times, force policies to be realigned.

To be sure, financial markets do not get everything right. They can miss important signals and remain calm in the face of rising vulnerabilities. This can be a serious problem. The years leading up to the GFC were a case in point. It is precisely in tranquil times that the seeds of future distress are often sowed.

But when financial markets smell weakness, they can move very quickly. To quote another highly influential German economist, the late Rudiger Dornbusch: *"Financial crises take much, much longer to come than you think and then they happen much faster than you would have thought."*[1] When that happens, policymakers need to react very quickly, often amidst great uncertainty and with their credibility dented. Unsustainable policies disappear swiftly.

What are the key lessons that I take from my experience?

The first is that crises are costly and best avoided, and while adopting sensible policy frameworks can make them less likely, some crises, sooner or later, will occur.

The second is that as economies and financial markets continue to evolve rapidly over time, policy settings and even frameworks that seem appropriate today will need to change, at times very quickly.

Thus, it is critical that policymakers have the nimbleness and flexibility to adapt both to crises and to evolving economic and financial circumstances – it is a perennial challenge to adjust and recalibrate policy.

This takes me to the core issue that I would like to focus on today: the value of trust in policymaking. It is much easier to make the necessary changes if the public trusts policymakers and their policies. But the importance of trust goes beyond enabling change. In my view, only when trust is present can public policies succeed.

Trust in public policies

What does 'trust' refer to? Essentially, it refers to society's expectation that public authorities will act predictably in the pursuit of predefined objectives and that they will succeed in their task.

Why is trust so important? If the public trusts authorities' actions, they will incorporate those actions in their own behaviour. This makes it more likely that the authorities will achieve their objectives. In addition, if the public has trust in public policies, they will be more willing to accept measures that impose short-term costs but deliver long-term benefits. In sum, trust underpins the effectiveness and legitimacy of policies.

Policymakers acquire trust by achieving their objectives over time. Hence the importance of setting clear policy goals, which provide a benchmark against which to evaluate policy actions and assess their success or failure. But setting targets alone is not enough. Policymakers must also pursue them decisively, particularly when the environment changes.

There is a positive feedback loop in the dynamics of trust. Effective and legitimate policies make it easier for the authorities to achieve their objectives. This, in turn, feeds back onto trust, producing a virtuous circle. However, this dynamic can also work in the other direction and, at times, very quickly.

In the extreme, if trust evaporates, the capacity to make effective public policies disappears. Preserving credibility is therefore a constant challenge, and it requires consistency in public policies over time. Institutional arrangements, like independent central banks, can be very valuable for this purpose.

To establish, enhance and preserve macro-financial stability, it is essential that the public retain trust in all of the key macro policy dimensions – monetary, financial and fiscal policies – individually and as a group. This requires coherence between them.

Let me elaborate. I will begin with the most fundamental aspect of central banking: the nature of money. The social convention of money, as we know it today, is based on the trust placed in it by the public. And as money is the basis for the entire financial system, the system's stability depends also on trust.

Fiat money is an asset that has no intrinsic value. It's worth derives from the social convention that underpins it, and from the institution that enables it to function: the central bank. Money only has value if the public believes that others will honour that value, today and in the future. This ensures that when a person wants to use it, they know that there will be finality in the payment.

Thus, the value of money clearly comes from trust. That is why the issuer of money is so powerful. This power carries with it great responsibility. Those who abuse their ability to issue currency deprive money of its value and forfeit the trust of the public. Germany knows this all too well.

The consequences of the state abusing the privilege of issuing money can be disastrous. These can range from high inflation and sharp exchange rate depreciations to the substitution of the national currency in favour of a foreign one. In the extreme, for example in hyperinflationary episodes, there could even be a return to barter. Such events

typically go hand in hand with financial instability, sharply lower economic growth, widespread job losses and soaring inequality.

The consequences of losing trust in money were a key reason for the emergence of central bank autonomy. After all, autonomous central banks are nothing more than institutions within the state with a mandate to preserve the purchasing power of the national currency. Their autonomy is the social engineering that solidifies society's trust in money.

Germany's experience in the 20th century illustrates vividly why trust in money matters. The contrast between the hyperinflation of the 1920s and the monetary stability that followed the foundation of the Deutsche Bundesbank could not be starker. And it is fair to say that the success of the Bundesbank inspired the emergence of central bank independence worldwide.

In recent decades, many central banks have followed the Bundesbank's lead and adopted monetary arrangements that allow them to anchor expectations and preserve money's purchasing power. Inflation targeting regimes are the most common framework to ensure this. The Bank of Mexico and the European Central Bank apply their own versions. But what does inflation targeting consist of?

Central banks do not control inflation directly. But their policy tools can influence it. When a central bank adopts an inflation target, it commits to use its tools to achieve that target. If the public trusts the central bank, then the inflation target, rather than current inflation, becomes a key reference point for price and wage decisions.

This contributes to low and stable inflation. Inflationary episodes are usually short-lived, reflecting changes in relative prices. Inflation becomes self-equilibrating and ceases to have a material influence on the behaviour of households or businesses.

That is why the inflation outbreak that followed the COVID pandemic and the onset of the Ukraine war was so concerning. The trust central banks had gained over many years could have been lost if society had started to doubt their commitment to price stability. Some generations experienced for the first time the risk of the economy transitioning to a high-inflation regime.

Once that transition starts, it can become increasingly difficult to stop. Therefore, it was necessary and appropriate for central banks to tighten policy forcefully and decisively through higher interest rates to restore price stability. The tighter stance may need to be maintained for a long time, for only through resolve, perseverance and success can trust in money be preserved.

Commercial bank money also needs to command trust. It is well known that the money issued by the central bank, known as primary money, is not the only money that circulates in a modern economy. Commercial bank money, in the form of bank deposits and credits, is what most households and businesses use for the bulk of their day-to-day transactions. It is thus fundamental to the monetary system. At the same time, for most people primary and commercial bank money are indistinguishable. That is by design.

Over time, institutional arrangements have evolved to extend society's trust in primary money to commercial bank money in a two-tiered monetary system. The central bank lays the foundation, and on the first floor are commercial banks.

The key is that interbank payments ultimately settle on the central bank's balance sheet, through the exchange of primary money between commercial banks. This guarantees the finality of payments and the singleness of commercial bank money.

The ultimate settlement of the banking system at the central bank is made possible by the central bank's ability to create liquidity by lending to the banking system. At times of great instability, the central bank can also provide additional liquidity through its well-known function of lender of last resort. In doing so, it safeguards the public's trust in the entire monetary system.

To put into perspective the enormous value of the framework I have just described which supports trust in primary and bank money, it is useful to consider recent failed attempts to issue private money through technologies that allow transactions based on decentralised ledgers.

These alleged forms of money function without central bank intervention, a lender of last resort or a reliable regulatory and supervisory framework. They have led to the proliferation of so-called cryptocurrencies, which cannot guarantee finality of payments nor a stable value, and so clearly lack the fundamental attributes of money.

These developments reinforce the point that what sustains fiat money over alternatives based on novel technologies is the institutional framework and the social convention that support it, which are precisely what makes it reliable for the public.

However, the mere existence of a two-tier monetary system is not enough to guarantee trust. The banking system must also remain solvent. Because banking crises have large social costs, the system should be extensively regulated and supervised.

A complement is deposit insurance, which exists to forestall potential bank runs. These layers of protection aim to safeguard the public's savings and are manifested in trust in both primary and commercial bank money.

Banks' resilience has increased markedly since the GFC. We reaped the benefits of the comprehensive regulatory and supervisory response to that crisis in the COVID pandemic, as banks were able to play a vital role in keeping economies afloat. Even when banking stress emerged in 2023, the post-GFC reforms and authorities' rapid deployment of crisis management tools limited the fallout to only a handful of institutions.

Nonetheless, there is still work to do to bolster the banking sector's resilience. Make no mistake, the core responsibility lies with banks themselves. There is no substitute for sound business models, adequate risk management and effective governance. But banking supervision needs to up its game to identify and remedy problems at banks proactively[2]. And we need timely, full and consistent implementation of banking reforms and regulations, including Basel III.

Recent decades have also seen rapid growth in the non-bank financial system. This sector comprises mainly activities involving securities, including debt instruments and broader forms of intermediation performed by insurance companies, private credit, investment service companies and hedge and pension funds, among others. In many countries, non-bank financial intermediation has for some time now accounted for over half of the financial system.

The need for greater supervision and regulation of the non-bank sector has become more pressing in the light of recent episodes of extreme instability. One reason is to prevent nefarious arbitrage between regulated and unregulated financial activities.

In addition, the sector's interconnectedness with the traditional banking system and the tendency of non-bank intermediation to generate opaque and excessive leverage and substantial liquidity mismatches create systemic risks. Unforeseen events in this sector can trigger systemic financial crises.

In recent years, some central banks have had to act as 'market-makers of last resort' to defuse crises and preserve trust in the broader financial system. Because such actions may conflict with central banks' measures to preserve price stability, greater regulation and supervision of the non-bank financial sector are indispensable.

Within the universe of debt instruments, public debt is of particular importance. If used appropriately, public debt allows governments to successfully function. But, from a macro-financial point of view, it is important that any public debt is, and is seen to be, sustainable. Investors must trust the government to meet its financial obligations, without resort to central bank financing.

Public debt plays a strategic role. It is considered the instrument with the lowest credit risk, making it essential for grounding the risk of asset portfolios, particularly those of banks. In addition, public debt serves as the main reference for valuing other forms of debt, for example corporate debt.

Hence, defaults on public debt compromise the stability of the whole financial system. They also threaten monetary stability since the central bank, even if it is formally autonomous, could find it necessary to finance debt service with primary issuance, leading to fiscal dominance of monetary policy. Under these circumstances, economies would cease to have a nominal anchor and would be cast adrift.

The result would be rising inflation and sharp exchange rate depreciations. We can thus appreciate the vulnerabilities that can be triggered if trust in public finances is lost.

In the light of these considerations, it is imperative for fiscal authorities to curb the relentless rise in public debt. The post-GFC low interest rate environment flattered fiscal accounts. Large deficits and high debt seemed sustainable, allowing fiscal authorities to avoid hard choices. But the days of ultra-low rates are over.

Fiscal authorities have a narrow window in which to get their house in order before the public's trust in their commitments starts to fray. As I pointed out earlier, financial markets can remain calm in the face of large imbalances until suddenly, one day, they no longer are.

That is why fiscal consolidation in many economies needs to start now. Muddling through is not enough. In many countries, current policies imply steadily rising public debt in the coming decades. Demands for more public spending will only increase, not least due to population ageing, climate change and, in many jurisdictions, higher defence spending.

Fiscal authorities must provide a transparent and credible path to safeguard fiscal solvency, ideally supported by stronger fiscal frameworks. And they must follow through on their commitments.

Fiscal health is not only about avoiding crises. It also brings material benefits. The lower long-term interest rates and debt service burdens enjoyed by Germany, compared with some of its advanced economy counterparts, are a prime illustration. Greater trust in public finances also increases fiscal space. This allows fiscal authorities to maintain trust even in the face of adverse events that require expansionary policy responses as, once again, Germany's recent experience has illustrated[3].

It is clear from what I have discussed that trust in the various aspects of macroeconomic policy – monetary, financial stability and fiscal – is closely interrelated. Monetary instability imperils financial stability, erodes the willingness of investors to hold public debt and hammers public confidence. Financial crises have large fiscal costs. And loss of confidence in public finances compromises the stability of the whole financial system and can undermine price stability.

This could happen because of political pressure to keep interest rates low to maintain fiscal space. But it can also occur if central banks perceive that raising interest rates risks triggering a sovereign debt crisis. Either way, monetary and financial stability are seriously undermined.

Thus it is essential to preserve trust in all pillars of a country's macro-financial frameworks, and for there to be consistency between them. In practice, this represents a great challenge due to the multiple authorities involved and the existence of unavoidable political motivations, especially with regard to fiscal policy. This is not an insurmountable problem, but it highlights the need for consistency and coordination of public policies.

In this context, I think it is unavoidable to mention the need for consistent policy frameworks in the euro area. While the institutional environment features a single monetary policy, there is no fiscal authority, and movement towards a fuller banking and capital union has been slow.

This hinders coherence and can make the euro area more vulnerable, as we witnessed during the sovereign debt crisis. The best institutional framework to deliver policy coherence is open for debate, but the value of coherence itself seems self-evident.

Let me add a final reflection on the credibility of fiscal and monetary policy today. Recent experiences should prompt a reassessment of the appropriate role of monetary and fiscal policy and greater realism about what they can deliver. Fostering unattainable expectations about policymakers' ability to smooth out every economic pothole will ultimately lower trust in public policies.

For monetary policy, a prudent approach would be to avoid excessive 'fine-tuning'. Central banks should not be called upon to stabilise inflation at very short horizons and within narrow ranges.

This is particularly important because, as recent events have shown, inflation will partly depend on factors that are not under central banks' control.

For fiscal policy, prudence requires allocating scarce fiscal resources to measures that can raise future growth. In addition to the green transition, this includes improving healthcare systems, spending on education and improving infrastructure.

Above all, we must remember that structural reforms are the best tool to sustainably increase a country's growth potential.

Towards a soft landing?

Let me start wrapping up by highlighting the recent positive developments in the fight against inflation. It seems that we are on route to a soft landing, thanks to the forceful, opportune and decisive monetary policy response.

Lower inflation, combined with surprisingly resilient activity and labour markets, suggests that we are on the right course. Financial markets seem to agree – the prices of shares and other risky asset classes have reached new highs in recent months.

In the light of the enormous strains placed on the global economy in recent years, a soft landing would be an impressive outcome. It would surely bolster trust in macroeconomic policymaking. So, where do we stand?

On the inflation front, the news is good. The monetary medicine is working. A year ago, inflation averaged 7% in advanced economies. Today it is 3%. In emerging market economies, excluding a few outliers, inflation averages 4%.

Admittedly, central banks cannot claim all of the credit. As the blue bars in Graph 1 show, lower commodity prices and the easing of pandemic-related supply disruptions also played a role. I am showing here estimates for the United States, although the story for the euro area would be similar.

Central bank actions were felt in other ways, however. Tighter monetary policy restrained demand. Just as importantly, central banks' accumulated trust allowed them to bring inflation down without the need for a large recession.

This was a stark contrast to the end of the most recent global inflationary outbreak in the late 1970s, which occurred at a time when many central banks lacked credibility as inflation fighters. Central banks' public commitment to restore price stability, and decisive actions in pursuit of this objective, prevented changes in 'inflation psychology' from taking hold and kept second-round effects at bay.

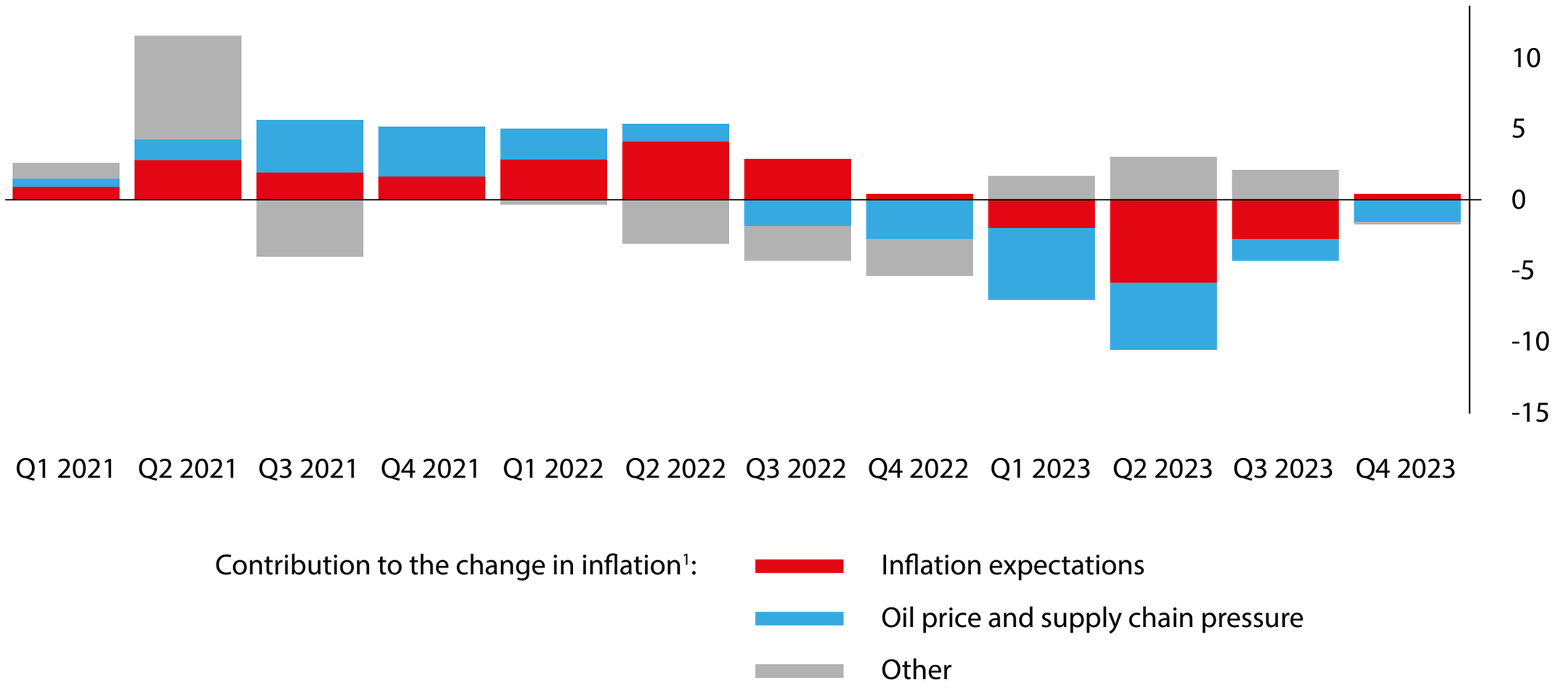
As the red bars in Graph 1 show, measures of inflation expectations, which rose concerningly at the start of the inflation outbreak, began to fall shortly after central banks started to raise rates.

Simulations by some of my BIS colleagues, shown in Graph 2, illustrate that, if central banks had not tightened policy, inflation would have stayed high, even as the pandemic-related supply shocks faded[4]. Through their actions, central banks showed their firm commitment to achieving their mandates.

Lower inflation has come at a remarkably small cost to the real economy. To be sure, global growth has slowed. Here in Europe, we narrowly avoided a recession last year. But labour market conditions remain firm and, at a global level, the growth slowdown is shallow. Against a backdrop of the largest and most synchronised monetary policy tightening in decades, this is an impressive achievement.

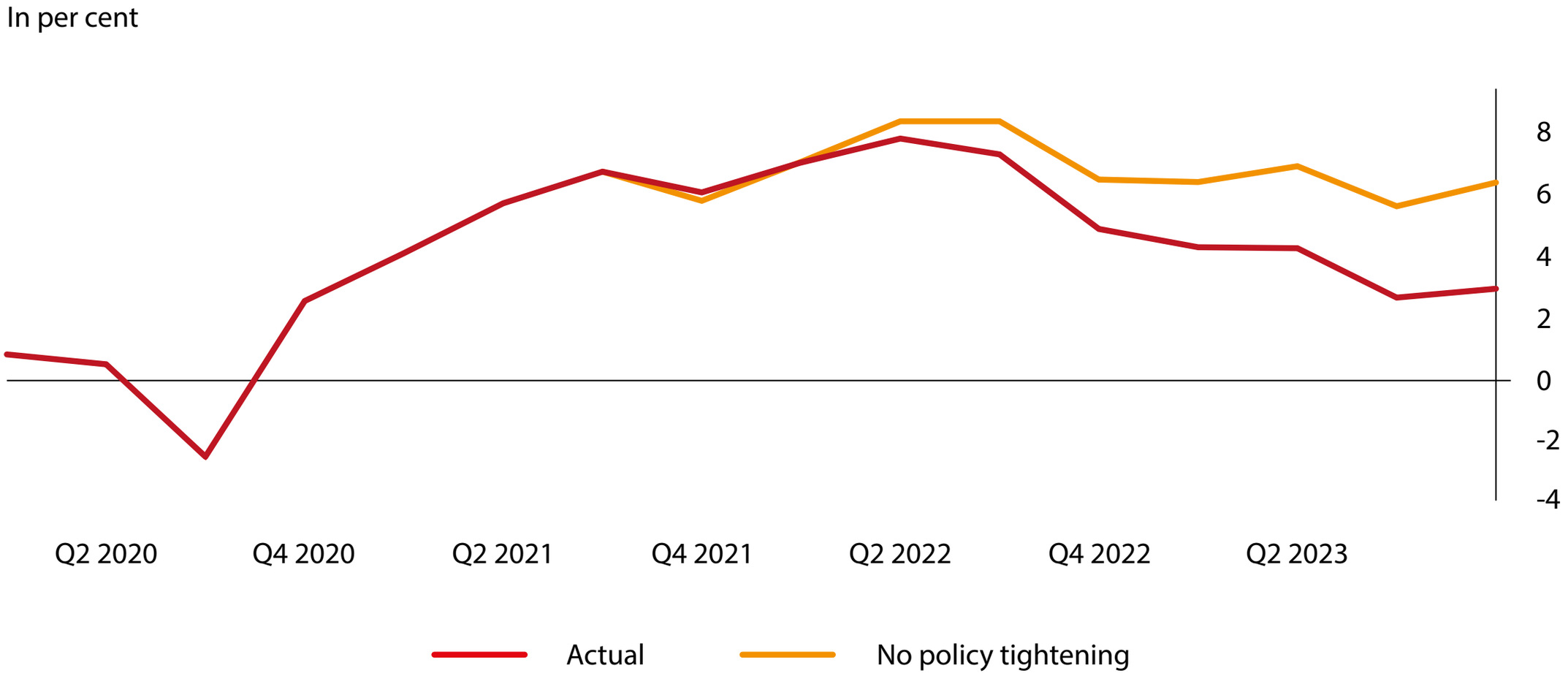
Graph 1. Contributors to disinflation: decomposition

In percentage points



1 Contributions to the change in quarter-on-quarter inflation over past year based on a linear regression model using data for the United States.
Sources: Federal Reserve Bank of New York; Federal Reserve Bank of St Louis, FRED; Bloomberg; BIS.

Graph 2. Inflation would have remained much higher without central bank actions



Based on simulations from a medium-scale macroeconomic model of the United States.
Sources: Federal Reserve Bank of St Louis, FRED; BIS.

A strong recovery in aggregate supply was an important contributor. It helped support output while lowering inflation. Without the supply recovery, disinflation would have been harder. But we should not forget that, without tighter monetary policy, there would have been no disinflation.

A soft landing is not guaranteed, however. Central banks' job is not done. While inflation is lower, it is still above central banks' targets. And there will surely be more bumps in the road. The medium-run risks to inflation – such as deglobalisation, economic fragmentation, adverse demographic trends and the need to fight climate change – reinforce the need for central banks to stay the course. It is only in this way that the public's trust in money can be preserved.

Conclusion

The events of the past decades presented policymakers with frequent and intense challenges. Facing extraordinary strains, policymakers strived to preserve the value of money, and keep their economies and the financial system functioning. These challenges showed the importance of the public's trust in policymakers, which allowed for decisive action. And policies worked better when they were each part of a coherent whole.

As Goethe wrote in Faust: *"Es irrt der Mensch, solang er strebt"* – man errs as long as he strives. Policymakers did not get everything right. Nonetheless, their attempts to make good choices were noticed by society.

If the events of the 21st century so far are a guide, we should not expect plain sailing in the coming decades. Building resilience will require policymakers to apply an appropriate policy mix and communicate it effectively. Monetary policy will need to prioritise the inflation fight, until it is decisively won and price stability is restored.

Financial stability policy needs to ensure a resilient banking sector and address remaining regulatory gaps. Fiscal policy will need to rationalise expenditure while making room for vital investments in our future. But these policies are unlikely to be sufficient even if applied jointly.

Ultimately, to improve economic resilience and enhance sustainable growth, governments must rediscover the appetite for structural reforms that has been absent for far too long.

I would like to leave you with this last thought. Part of preserving trust is to know the limits of what policies can deliver. Expecting policymakers to deploy extraordinary macroeconomic policies to respond to every challenge is a sure way to erode the public trust.

Building resilient and robust economies and financial systems is the best way to ensure that policies remain effective, so that they can be deployed when they are needed the most. ■

Agustín Carstens is General Manager of the Bank for International Settlements

Endnotes

1. R Dornbusch and S Fischer, "International financial crises", CESifo Working Paper, no 926, 2003.

2. See A Carstens, "Investing in banking supervision", speech at the European Banking Federation's International Banking Summit, Brussels, 1 June 2023.

3. See J Nagel, "Fiscal policy challenges in a high inflation environment", speech at the Forum Finanzpolitik und Steuerrecht, Baden-Baden, 10 November 2023.

4. See P Amatyakul, F De Fiore, M Lombardi, B Mojon and D Rees, "The contribution of monetary policy to disinflation", BIS Bulletin, no 82, December 2023.

This article is based on a lecture by delivered at the Center for Financial Studies and the House of Finance at Goethe University, Frankfurt, 18 March 2024.

Why do financial crises happen so often?

Jon Danielsson argues that failures in regulation are a key reason for financial crises, and proposes that the authorities adopt diversification to build a more resilient financial system

One of my favourite exam questions is: 'Given our extensive knowledge about the causes of financial crises and the measures needed to prevent them, why do they happen so frequently?' We have had a deep understanding of financial crises for over 200 years. A 19th-century central banker dealing with the severe crisis of 1866 would find few surprises in the more recent ones. All crises share the same fundamental causes. Excessive leverage renders financial institutions vulnerable to even small shocks. Self-preservation in times of stress drives market participants to prefer the most liquid assets.

System opacity, complexity, and asymmetric information make market participants mistrust one another. These three fundamental vulnerabilities have been behind almost every financial crisis in the past 260 years, ever since the first modern one in 1763 (Danielsson 2022).

If we know why crises happen, preventing them should be straightforward. But given their alarming frequency, it does not appear to be so. When looking at the various alternative explanations and ignoring the political ones, we find two different narratives: consensus and diversification. Let's start with the consensus one.

The consensus narrative

We find the consensus narrative in financial stability and 'lessons learned' reports published by the financial authorities. These reports have become very common ever since the authorities re-started taking the financial system seriously after the crisis in 2008.

At the risk of oversimplification, the consensus narrative is as follows: some financial institutions bypass the spirit of regulations – deliberately or inadvertently – amassing large, illiquid, and risky positions that are increasingly vulnerable to stress. One of the best manifestations of this view is the Financial Stability Board's 2020 holistic review of the COVID March 2020 market turmoil.

The consensus narrative drives the recommended responses. Most parts of the system are safe because of regulations implemented after 2008. However, some undesirable activities have slipped through the cracks, necessitating tighter supervision, expanded regulatory coverage, and stronger capital and liquidity buffers.

The problem with the consensus narrative

I fear that the consensus narrative and its ever-increasing regulatory intensity will not protect us. Finance is crucial and requires risk to deliver on its promises to society. Unfortunately, the private sector's incentives are not fully aligned with society, giving rise to crises that have cost Europe and the US trillions of dollars.

Use the authorities' powers to push for a more diversified financial system – one that absorbs shocks and increases efficiency – instead of the current set-up, which drives homogeneity, procyclicality and deadweight loss

To mitigate that very high cost, after-the-fact public bailouts of private risk are unavoidable, further misaligning incentives for private risk-taking. This problem is the rationale for before-the-event regulation.

This then begs the question: why not regulate finance heavily? Well, we already do, and it has become very hard to do even more, as it appears that we are getting close to the upper limits of regulatory intensity.

The reason it is so difficult to regulate finance is that the financial system is one of the most complicated of all human constructs. In effect, it is infinitely complex. And when a system is infinitely complex, there are infinite areas where excessive risk and misbehaviour can emerge.

When market participants optimise, they are actively searching for overlooked areas in which to take risk, so it is almost axiomatic that crises happen where nobody is looking. How can we regulate something we have yet to see?

If crises are to be prevented, the architects of regulations have to foresee all the areas where vulnerability can emerge, and the supervisors must patrol all of them. That is not enough. They also need to identify all the latent links between the disparate areas of the system, channels that only emerge in times of stress.

Meanwhile, the authorities have to contend with political forces that benefit from the pre-crisis bubble and do not want regulations that threaten the perceived benefits to society. Add to this the dismissive attitude of the monetary policy and supervision authorities to the macroprudential agenda.

The objectives of the consensus-founded macroprudential narrative are impossible to achieve, in part because effectiveness demands much higher resources – human capital, politics, data, compute – than those available to the authorities. Even worse, it requires far more resources than the private sector needs.

The consequence is a cat-and-mouse game where the mice have the advantage.

Of course, this is well understood, and there is a consensus solution: build tall buffers to protect financial institutions against bad outcomes. Unfortunately, that will not prevent crises. There are several reasons why.

The first is how consensus narrative regulations harmonise beliefs and action. The practical implementation of the regulations not only favours market concentration because of high fixed costs but also compels financial institutions to measure risk and respond to it in the same way. This makes them behave like a herd.

While that is fine if we are regulating visible conduct, such as traffic police measuring speeding and issuing tickets, it is different with finance since the risk is latent (Danielsson 2024). The consequence is procyclical amplification of the financial cycle, increasing booms and deepening busts.

Regulations based on the consensus narrative also amplify the complacency channel of financial instability. If we believe that the authorities understand the system, have everything under control, and are confident that they will step in with bailouts if needed, it leads to overconfidence and excessive risk-taking, particularly in the parts of the system that the supervisors are not patrolling, which is most of it.

The resulting complacency and short-termism are key factors in most crises, such as the one in 2008. Such Minsky-type responses make crises more rather than less likely, as empirically shown in Danielsson *et al* (2018).

The consensus narrative further amplifies the political channel for instability. When the government takes increasing responsibility for financial activities, controlling risk and protecting us from the adverse consequences of that risk, it makes the state, rather than the private sector, responsible for finance.

That, in turn, has two consequences. The first, as argued by Chwieroth and Walter (2019), is that bailouts become a middle-class good that cannot be politically forsworn. The more the state gets involved with finance, the higher the chance of bailouts, which makes crises more likely.

Furthermore, when the state regulates risk-taking and underwrites losses, how can one respond to political extremists who question why we have a private-sector financial system in the first place?

Meanwhile, the fiscal and monetary resources to fight crises have mostly been exhausted by fighting minor stress. We could marshal very significant fiscal and monetary resources in 2008. If the same event happened today, that would not be possible. Knowing this undermines the credibility of financial policy, making crises more frequent and severe.

In addition, the consensus approach to regulations neglects efficiency while stressing stability. The aim of regulations is not financial stability. It is to support prosperous and stable economic activity. Not recognising that makes regulations subject to increasingly vicious political attacks. If the cost of regulating increases faster than the economy grows, the authorities will be forced to change direction.

Finally, the consensus narrative has led to rapid concentration and ever larger too-big-to-fail financial institutions, exacerbating systemic risk and inducing the further raising of protective buffers.

A different way forward

There are three related reasons why a financial system composed of a large number of relatively small and diverse financial institutions is more stable and prosperous than one with few large and similar institutions.

First, when financial institutions differ from one another, excessive systemwide risk-taking is less likely because, at any given time, the actions of some institutions inflate bubbles while others do the opposite.

Second, when faced with shocks, relatively homogeneous institutions will respond similarly, buying and selling the same assets at the same time. This leads to disastrous selling spirals. In contrast, when they are diverse, some institutions will buy and others will sell, dissipating shocks. In other words, a system with relatively homogeneous institutions acts as a shock amplifier, whereas a more diversified system absorbs and dissipates shocks and, hence, is more stable.

Finally, a system with many small and diverse institutions will be more prosperous. It allows better tailoring of financial services to the needs of the economy while also requiring lower buffers against systemic risk. This means it offers the cheaper provision of financial services.

The benefit of diversification

The consensus approach to regulating reduces institutional diversity since it is a partial equilibrium approach subject to a fallacy of composition: if every individual part of the system is made safe, the system is safe.

The problem is that the consensus approach makes financial institutions increasingly homogeneous, leading to herd behaviour that causes booms and busts. These financial institutions amplify shocks when stress occurs because they are compelled to seek safety in the same way.

It is better to borrow a basic principle from finance: diversification. Just as we should not put all our savings into one investment, a system composed of diverse institutions is more stable and prosperous.

This suggests learning from how competition authorities use their powers to increase competition. In practice, we can actively use the licensing regime to facilitate start-ups with diverse business models and tailor regulatory regimes to suit different types of institutions, including reducing the fixed cost of compliance.

Why does diversification not happen?

A competitive financial system simultaneously promotes and opposes diversification. Start-ups compete against incumbents by having better business models, diversifying the system. The profit motive drives incumbents towards the short-term selection of successful business models and, hence, homogeneity.

Since financial institutions are highly regulated, the authorities wield powerful tools that can either help or hinder the forces of positive diversification.

The incumbents' incentives are clear: keep entrants out. That means lobbying for regulations with high fixed costs and uniform licensing regimes and regulations, using arguments of 'fair play' and 'level playing fields'.

The consensus narrative for financial regulations also pushes for homogeneity. There are three reasons for this:

- first, regulatory capture, which is driven by the two-way traffic between banks and the authorities and political pressure;
- second, the mistaken belief that there is no trade-off between stability and efficiency;
- finally, the erroneous impression that a system composed of a few large players is easier to understand and to control than one that is more complex and diverse.

Conclusion

There are two acceptable answers to my favourite exam question: 'Given our extensive knowledge about the causes of financial crises and the measures needed to prevent them, why do they happen so frequently?' A coherent answer based on the consensus approach to regulations will earn the student an A.

But there is a better answer. Use the authorities' powers to push for a more diversified financial system – one that absorbs shocks and increases efficiency – instead of the current set-up, which drives homogeneity, procyclicality and deadweight loss.

We can easily achieve this by leveraging the licensing regime to actively permit start-ups that use innovative business models, and by tailoring regulations to the type of institution.

Such a diversified financial system will be more efficient, robust and stable – a win-win-win. ■

Jon Danielsson is Director of the Systemic Risk Centre at the London School of Economics and Political Science

References
Chwieroth, J and A Walter (2019), "The wealth effect: The middle class and the changing politics of banking crises", VoxEU.org, 3 June.
Danielsson, J (2024), "When risk models hallucinate", VoxEU.org, 3 February.
Danielsson, J (2022), The Illusion of Control, Yale University Press.
Danielsson, J, M Valenzuela and I Zer (2018), "Low risk as a predictor of financial crises", VoxEU.org, 26 March.
FSB – Financial Stability Board (2000), "Holistic Review of the March Market Turmoil", 17 November.

This article was originally published on VoxEU.org.

Why Europe must safeguard its global currency status

Amid geopolitical shifts Piero Cipollone argues Europe needs to further develop the infrastructure for making crossborder payments in euro with key partners

For the last quarter of a century, the euro has been a key global currency, second only to the dollar. It has demonstrated its resilience despite the coronavirus pandemic, Russia's war in Ukraine and the tragic conflict in the Middle East. The euro's estimated share of international currency use stands at over 19 per cent, a level that has remained broadly stable over the past five years.

Nevertheless, the currency's place on the global stage cannot be taken for granted, as a recent report by the European Central Bank on the international role of the euro shows. More reforms are needed.

China's increasingly large role in global trade is encouraging use of its currency. By 2023, the renminbi's share of China's trade invoicing had risen to around one-quarter for goods and one-third for services. It is racing with the euro to become the second most used currency for trade finance[1].

History shows that the evolution of global currencies is deeply intertwined with that of the global geopolitical order. In an increasingly multipolar world, there are signs that the fragmentation of the global monetary system is no longer a remote possibility.

To diversify and protect against geopolitical risks, central banks — led by China's — are accumulating gold at the fastest pace seen since the second world war. And anecdotal evidence suggests that some countries are exploring ways of using their own currencies more in international trade transactions instead of those of countries sanctioning Russia.

Yet nowhere else are the risks of global monetary system fragmentation more visible than in international payments. At a time when we should be integrating payment systems to reduce their complexity and cost to users, some nations are deliberately creating separate platforms as alternatives to existing global infrastructures.

For example, China, Iran and Russia have created their own crossborder payment messaging systems, while BRICS members have started to discuss a 'bridge' platform for linking digital payments and settlement. These developments could potentially disrupt the smooth flow of capital and reduce the efficiency of the global financial system.

Given these shifts, there are compelling economic and political reasons for seeking to preserve the euro's global currency status. This status brings tangible benefits to European citizens, such as low borrowing costs in international capital markets and protection from exchange rate volatility.

By bolstering safety, liquidity and connectivity, we can ensure that the euro continues to strengthen as a cornerstone of the global monetary system

Moreover, in a fragmented geopolitical landscape, the euro's international currency status provides strategic autonomy by shielding Europeans from external financial pressures.

Internally, the euro's appeal to foreign investors hinges on maintaining confidence in its stability, supported by well-anchored expectations of price stability and sound economic policies. And its appeal depends on the size and liquidity of the market for safe euro-denominated debt securities and the resilience of the underlying market infrastructures, particularly as a haven in times of stress.

A majority of official reserve managers have expressed an interest in increasing their euro holdings but note that the currency's attractiveness is hampered by a lack of highly-rated assets and centrally-issued debt[2].

So building a stable, technically resilient, and deeper market for internationally accepted euro debt securities is essential. To be a reliable haven in times of stress, this market could be supported by a robust and flexible supply of common instruments[3].

Providing a broader pool of euro-denominated safe assets, which would act as a European risk-free benchmark, would also be crucial to deepening euro-denominated capital markets. That is why building a genuine European capital markets union must go hand in hand with efforts to further strengthen the fiscal dimension of the EU economic and monetary union.

Externally, Europe needs to further develop the infrastructure for making crossborder payments in euro with key partners. This could, for example, involve interlinking the euro area's Target Instant Payment Settlement system with fast payment systems in other jurisdictions, either through bilateral links or by connecting to a common, multilateral platform.

Such steps could strengthen the trade and financial relations with key partners, including emerging economies, especially where legislation on combatting money laundering and terrorist financing is fully aligned with the international standards established by the Financial Action Task Force. They could also pave the way for central bank digital currencies to be used to make crossborder payments in the future.

Robert Mundell — the late international economist whose Nobel Prize-winning work was so influential for the creation of Europe's single currency — once said of the euro: *"In all the aspects in which it was expected economically to make an improvement, it has performed spectacularly."*[4]

By bolstering safety, liquidity and connectivity, we can ensure that the euro continues to strengthen as a cornerstone of the global monetary system. ■

Piero Cipollone is a Member of the Executive Board of the European Central Bank

Endnotes

1. "China's renminbi pips Japanese yen to rank fourth in global payments", Financial Times, December 21, 2023.

2. "HSBC Reserve Management Trends: How central bank reserve managers are adapting their strategies amid a rapidly changing environment", 15 April 2024.

3. Panetta, F (2020), "Covid-19 crisis highlights the euro's untapped potential", Financial Times, June 12.

4. Wallace, L (2006), "Ahead of His Time – Interview with Robert Mundell", Finance and Development, Volume 43, No 3.

This article is based on an ECB blog post.

Bank liquidity, regulation, and the Fed's role as lender of last resort

Bank regulators are planning to introduce a package of liquidity regulations in response to last year's bank failures. Michelle Bowman discusses potential changes to improve liquidity for banks in times of stress

We have recently passed the one-year anniversary of the failures of Silicon Valley Bank (SVB) and Signature Bank. The long shadow of these bank failures, and the subsequent failure of First Republic, have prompted a great deal of discussion about the bank regulatory framework, including capital regulation, the approach to supervision, and the role of tailoring, among other topics.

It is my hope that we consider the appropriate role of the Federal Reserve in providing liquidity to the US banking system and, of course, its role as the 'lender of last resort' through the discount window and authority under section 13(3) of the Federal Reserve Act.

I look forward to a deeper examination of important policy questions, including the lessons that should be learned from the banking system stress experienced last spring, the broader stress in financial markets during the COVID-19 crisis, potential approaches to operationally enhance and optimize tools like the discount window to more effectively meet industry liquidity needs, and the importance of effective resolution mechanisms in the banking system.

I would like to briefly touch on three main themes: (1) the broader framework in which the Federal Reserve supports liquidity in the banking system, particularly how this function complements other regulatory requirements and sources of liquidity; (2) how this function can be optimized to work within the evolving liquidity framework; and (3) the challenges we face in making the Federal Reserve's liquidity tools, particularly the discount window, effective.

The Federal Reserve's role in banking system liquidity

The complexity of the US financial system makes it difficult to predict where the next stress (or in the worst case, the next crisis) will arise. While we focus on recent episodes that required the Federal Reserve to employ its liquidity tools—the COVID crisis and the early 2023 banking stress—it is helpful to consider how the Federal Reserve's authority has evolved in the aftermath of the 2008 financial crisis.

Let's review the historical context, which could be helpful for framing the discussion. In 1913, Congress established the Federal Reserve at least in part to help address the pattern of cyclical financial panics and the ensuing economic turmoil that followed by allowing the Fed to create a more elastic money supply to meet demand for liquidity during times of stress.

We will see a growing momentum to 'do something' that would help address the banking stress from 2023. While some reforms may be necessary, we should think about the response to banking stress more broadly

This authority included tools like open market operations, later used as a tool for monetary policy[1]. Since its establishment, the Federal Reserve was granted the authority to engage in discount window lending[2].

In addition, during the Great Depression, the Fed was given a broader set of tools to engage in emergency lending under section 13(3) of the Federal Reserve Act[3].

More recently, in 2003, the Federal Reserve restructured its previous discount window lending programs and established the Primary Credit Facility (PCF) and Secondary Credit Facility[4]. Primary credit enabled financially strong banks to obtain secured loans from the discount window at a penalty rate.

The secondary credit provided discount window loans at a higher rate, and with higher collateral haircuts and other more stringent terms than apply for primary credit, to solvent institutions that did not qualify to borrow from the PCF[5].

This evolution of the discount window function more closely aligned operations with a theory, often attributed to Walter Bagehot, that central banks should lend freely to solvent institutions against good collateral, at a penalty rate of interest[6].

The Fed used its lending tools extensively during the 2008 financial crisis. Relying heavily on discount window lending authority and emergency lending facilities under section 13(3) of the Federal Reserve Act, the Fed provided emergency liquidity to support individual firms that were under severe stress, and to facilitate the flow of credit more broadly.

Of course, the financial crisis left a lasting imprint on many Americans who suffered significant economic harm, many of whom have not yet fully recovered. It also prompted Congress to review and amend the Fed's authorities through the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act).

The banking system today is stronger and more resilient than it was before the 2008 financial crisis with significantly more capital and substantially more liquidity. US banks are also subject to a host of supervisory tools that did not exist prior to the Dodd-Frank Act, like new stress testing requirements[7].

Many of the regulatory changes implemented at that time were designed to reduce the probability of large bank failures, but the statute also mandated other changes designed to improve the likelihood that failing large banks could be resolved without broad systemic disruptions[8].

Of course, these changes were additive to existing authorities that are meant to promote banking system resilience, particularly the other core element of the federal safety net, deposit insurance[9].

Congress also made significant changes to the Fed's emergency lending authority. For example, section 13(3) facilities must now be broad-based, rather than designed only for individual firms, and must be approved by the US Treasury Secretary.

In addition, loans can only be made to solvent institutions, and there are new collateral and disclosure requirements[10]. Further, while the Dodd-Frank Act preserved the Fed's ability to make discount window loans to eligible borrowers, including depository institutions and US branches of foreign banks, it made some modifications.

Notably, one change that I will return to later is the new requirement that discount window lending is no longer confidential. These loans, including the names of borrowing institutions, are now required to be disclosed with a two-year lag[11].

Changes made by the new law and other subsequent changes have attempted to strike a balance between making firms more resilient to stress and adding additional parameters to the Fed's liquidity tools. The complementary tools we have—the prudential bank regulatory framework, tools to promote banking system liquidity and stability, discount window lending and 'lender of last resort' authority, and resolution tools—all contribute to the safety and soundness of individual banks, and more broadly, to financial stability.

Broadly defined, the challenge we face is that banking crises and banking stress can arise from unpredictable events. They can be the product of external events (like a global pandemic) or can arise from cascading failures of bank management and regulators to identify and effectively address and mitigate the buildup of risk. This risk can occur at a single institution, like we saw in the lead-up to the failure of SVB, or more broadly throughout the financial system, as we saw during the last financial crisis.

When we consider banking system stress and potential crises in the broader context, our primary goal should always be prevention, particularly so that we can avoid contagion risks that lead to financial instability and more significant government intervention. We should be reluctant to intervene in private markets, including using emergency government lending facilities to support private enterprises.

The federal safety net that covers the banking system—including discount window lending and deposit insurance—is meant to make the US banking system and broader economy more resilient. Where market disruptions affect liquidity, it is important that these tools—particularly discount window lending—function effectively.

So, we must ask whether there are steps we can take to optimize the functioning of these tools and identify some of the key challenges we face in making these tools effective, including preserving industry standard access to liquidity outside of the Fed's tools for day-to-day liquidity management, like advances from the Federal Home Loan Banks.

Optimizing the lender of last resort function

When we think about the Fed's lender of last resort function, we must think about the broader framework that supports bank liquidity, including liquidity regulation, bank supervision, deposit insurance, and day-to-day liquidity resources. While my discussion focuses primarily on discount window lending, I will also briefly address design issues that we experienced with the recently expired Bank Term Funding Program.

I think we can all agree that the discount window remains a critical tool, but it does not operate in isolation. It operates to support bank liquidity, but it is an additional resource in the federal safety net that allows eligible institutions to weather disruptions in liquidity markets and access other resources.

First, there are questions about the utility of the discount window in light of its scope and the evolution of the banking system. There are a limited set of entities that have access to discount window loans, including depository institutions and, in unusual or exigent circumstances, designated financial market utilities1[2].

As activities continue to migrate out of the regulated banking system, what are the implications of more activity occurring outside the banking system as it relates to the effectiveness of the discount window as a tool?

Second, are there ways in which the Fed can enhance the technology, the operational readiness, and the services underpinning discount window loans to make sure that they are available when needed? Here, the events in the

lead-up to the failure of SVB are illuminating—SVB experienced difficulty in accessing the discount window before its failure.

We must understand and evaluate these difficulties and determine whether there are improvements the Federal Reserve System can make to ensure the discount window is an effective tool to provide liquidity support. Are there operational issues that can be improved, whether by improving the technology or extending business hours for the discount window and other Reserve Bank payment services like FedWire® and ACH (automated clearinghouse), particularly during times of stress?

The Federal Reserve System must also take a close look at our operational readiness and capacity. Banking stress can manifest quickly and outside of regular business hours in different time zones, and we must make sure that the tools we have are available and prepared with trained and experienced staff ready to deal with the evolving risks of liquidity stress and pressure.

Finally, are there changes that need to be made to support contingency liquidity on the borrower side? One prominent issue that has come to light recently is whether there should be some form of pre-positioning requirement—whether banking institutions should be required to hold collateral at the discount window, in anticipation of the need for accessing discount window loans in the future[13].

Arguably, requiring pre-positioning at the discount window may serve a variety of purposes. One use case is ensuring the system is efficient enough to allow borrowers to access discount window loans in a timely manner, including by getting collateral to the discount window to support loans. We have much work to do on this front.

To fulfil its function, the discount window must be able to provide liquidity quickly. The failure of SVB demonstrated how rapidly a run can occur and revealed that the discount window must be able to operate in a world in which new technologies, rapid communications, and the growth of real-time payments may exacerbate the speed of a bank run.

Identifying and mitigating the technology and operational issues that affect the discount window should go a long way to addressing this concern. Understanding that these problems exist and requiring pre-positioning of collateral at the discount window may not fully address any technological and operational shortcomings of the discount window.

But as a secondary matter, the notion of required collateral pre-positioning has also been proposed as a complementary liquidity requirement for banks, in part to ensure greater liquidity certainty to balance perceived 'runnable' funding sources, as with SVB's significant proportion of uninsured deposits.

While this could be an effective approach, we do not fully understand the consequences of a new pre-positioning requirement or whether, given the unique nature of SVB's business model and lax supervision, other institutions would have similarly runnable uninsured deposits or if this was an idiosyncratic event.

Further, would required pre-positioning of collateral impede a bank's ability to manage its day-to-day liquidity needs (including from private sources at lower cost)? Would pre-positioning collateral increase operational risk, or otherwise change bank activities? Would there be any unintended consequences from requiring banks to encumber more assets on their balance sheets?

More fundamentally—is a change of this magnitude, requiring a new daily management of discount window lending capacity, necessary and appropriate for all institutions, or are there particular bank characteristics that may

warrant this additional layer of liquidity support? These are all important but as yet unanswered questions that need to be explored and understood before imposing such a radical shift.

Currently, banks are not mandated to use the discount window to access liquidity. In fact, one of the core functions of bank management is to make the day-to-day decisions about how the institution will manage liquidity and other responsibilities.

While it may be appropriate for supervisors to encourage banks to test contingency funding plans and to evaluate whether those plans are adequate in the context of examination, we must be cautious to not cross the line from supervisor to member of the management team and to avoid interfering with the decision making of bank management by mandating across-the-board changes in response to the failure of a single unique institution.

We need to ask whether having one standardized set of rules for institutions with different activities, risk profiles, and funding structures is the most efficient and effective way to support bank liquidity, particularly as we think about not only stressed conditions and liquidity disruptions in the market, but also day-to-day management and activities.

Challenges

I will briefly touch on a number of the challenges and issues about liquidity support of the banking system and the special role of the Federal Reserve as lender of last resort.

Stigma

A long-standing challenge to the utility of discount window borrowing is the perception of stigma. During times of stress, signs of banking sector weakness are often magnified through small and independent actions of institutions, which may add to the reluctance to borrow from the central bank when other sources may be available.

The perception of stigma existed long before the new Dodd-Frank disclosure requirements, and it is possible that public disclosure of the borrowing—even with a two-year delay—may create a greater deterrent.

Regardless of the timing of the disclosure, the reality is that market participants have a strong interest in identifying any public signals of bank financial health, including discount window lending.

Even where the market is just making educated guesses about discount window lending (for example, by looking at public-facing liquidity management activities of banking institutions), the stigma risk can be an important consideration for banks trying to manage public perceptions of their financial condition.

The Federal Reserve cannot entirely eliminate discount window borrowing stigma through regulatory fiat. One of the key sources of stigma seems to be the spectrum of reasons that a bank may choose to borrow from the discount window: the need for borrowing could be due to market disruptions in the provision of liquidity or a scarcity in the total amount of reserves in the banking system but could also indicate a specific borrower's growing financial stress.

Of course, it is possible that a combination of factors may lead a bank to access the discount window—as stress on banking institutions builds, there may be a 'pullback' on the ordinary liquidity tools banks use, accompanied by increased demands for liquidity.

In this context, discount window lending becomes one additional data point for the market to interpret—while the signal it may send is unclear, one can easily imagine that the market may be skittish and fixate on any sign of financial weakness.

The broader issue, however, is the health of the banking system and particular financial institutions, which can be affected by a number of other factors. For example, as we saw with SVB, the public messaging around the sale of

securities and the prospective capital raised were both public announcements that altered the perception of the institution's financial health and risk profile.

In short, while discount window 'stigma' is an important issue, it is a subset of a broader concern—the perception of the institution's financial health—that each bank must confront as it manages its funding resources, risk profile, and liquidity.

At the same time, we should explore ways that the Federal Reserve can work to mitigate stigma concerns. In some ways, the design of primary discount window credit, where a borrower must meet financial standards for borrowing, suggest that the 'market signal' of discount window borrowing should perhaps speak more toward market liquidity disruption than an individual institution's financial condition.

We should explore ways to validate the use of discount window lending in our regulatory framework. While the federal banking agencies have encouraged institutions to be prepared to access discount window loans, we should also seriously consider whether we should finally recognize discount window borrowing capacity in our assessment of a firm's liquidity resources, including in meeting a firm's obligations under the Liquidity Coverage Ratio[14].

One of the emerging arguments about how the Federal Reserve can mitigate stigma concerns is simply by mandating that banks pre-position collateral and periodically borrow from the discount window.

The notion is that the 'signalling' effect of discount window borrowing becomes more muted when more participants are essentially forced to use it to meet a regulatory requirement or a supervisory expectation. I question whether this approach will truly address the underlying stigma concern.

The discount window has not historically functioned as a source of ordinary day-to-day liquidity for the banking system, but rather as a backup liquidity resource and it is priced as such.

Our expectation should not be that the Federal Reserve replaces existing sources of market liquidity for banks in normal times. As a source of backup liquidity, the question becomes whether requiring pre-positioned collateral would mitigate the stigma of drawing on the discount window.

To be effective, banks must be willing to obtain discount window loans when needed, and it is not clear that required pre-positioning or even testing requirements will address the perceived stigma associated with a bank's need to access the discount window for emergency liquidity purposes. The market will continue to take signal from a bank's external activities in liquidity markets—and try to extrapolate whether a bank is using the discount window—and draw a negative inference from this borrowing.

Broad-based approach to bank liquidity

The discount window is a small but important element of bank liquidity, but banks manage liquidity in many ways for day-to-day business needs and during times of market stress.

Considering discount window reform narrowly ignores the interrelationships among various liquidity resources, liquidity requirements and regulations, and liquidity planning. Building resiliency in the financial system requires policymakers to think about these variables together, ensuring that reforms are rational and contribute to a complementary liquidity framework.

The complexity of liquidity issues warrants a broad-based review before we embark on piecemeal changes. That review should endeavour to understand not only the need for reform, but also the tradeoffs of different

approaches, including the economic cost. However, the proposed change raises many questions about not only cost and effectiveness, but also unintended consequences.

Another example is the use of Federal Home Loan Bank (FHLB) advances by some banks as a supplemental source of liquidity, and how this resource functions along the continuum of day-to-day liquidity management to instances of widespread stress in the banking system and among individual firms.

The FHLBs are an important source of liquidity for many banks. At the same time, the operational design of FHLB advances make these advances poorly suited to function as emergency liquidity support for the banking system.

By contrast, the Fed's discount window lending authority, and the flexible authority to lend under section 13(3) of the Federal Reserve Act, place the Fed well to function as the lender of last resort in support of banking system liquidity during times of stress.

A note on the design of emergency lending facilities: the Bank Term Funding Program

Before closing, I would like to briefly reflect on events we saw this winter, when design flaws with the Bank Term Funding Program (BTFP) were first identified. On March 12, 2023, the Federal Reserve, with the approval of the Treasury Secretary, announced the creation of the BTFP, which was designed to make additional funding available to institutions to *"help assure banks have the ability to meet the needs of all their depositors."*[15]

This program was initially authorized to make new loans for a full year, even though at the time, it was not clear that 'unusual and exigent' circumstances would continue to exist for a full year that would warrant the ongoing availability of loans under the program[16].

Under the BTFP, eligible depository institutions were able to pledge Treasury securities, agency debt, and agency mortgage-backed securities—valued at par—to obtain one-year loans. This program allowed institutions to avoid selling those assets to generate additional liquidity.

By valuing the collateral at par—when the market value had declined due to the rising interest rate environment—the program allowed eligible borrowers to obtain a greater amount of liquidity than they would have been able to by simply pledging collateral to the discount window.

These generous collateral terms were accompanied by generous rate terms and prepayment flexibility. As originally designed, the interest rate for loans under the program was set at the one-year overnight index swap rate, plus 10 basis points[17]. Borrowers were also entitled to prepay loans at any time without penalty[18].

As has been well documented, the combination of these terms over time created a significant arbitrage opportunity, which the Fed sensibly cut off as the program was approaching the end of its term for originating new loans[19].

We must learn from this experience. When we identify flaws in program design or ways to improve our tools in the future, we should avail ourselves of the knowledge we have learned through experience, including by shutting down an authorized section 13(3) facility when it is no longer needed, and lending at a true penalty rate so the usage of the facility naturally declines as market conditions normalize.

Conclusion

As regulatory attention turns toward the liquidity framework and liquidity regulation, I expect we will see a growing momentum to 'do something' that would help address the banking stress from 2023. While some reforms may be necessary, we should think about the response to banking stress more broadly.

We should continue to focus on improving the targeted approach of supervision, to enhance the 'prevention' of banking system stress. We should think about the liquidity framework in a broad-based manner to ensure that the available tools, resources, and requirements are working in a complementary way. And we should understand what changes we need to make discount window lending and other emergency lending programs more efficient and effective. ■

Michelle W Bowman is a Member of the Board of Governors of the Federal Reserve System

Endnotes

1. Federal Reserve Act, ch. 6, 38 Stat. 251 (1913).

2. Federal Reserve Discount Window, "General Information: The Primary & Secondary Lending Programs."

3. Section 13(3) of the Federal Reserve Act was added by section 210 of the Emergency Relief and Construction Act of 1932.

4. Mark Carlson and Jonathan D Rose, "Stigma and the Discount Window," FEDS Notes (Washington: Board of Governors of the Federal Reserve System, December 19, 2017).

5. Carlson and Rose, "Stigma and the Discount Window."

6. Walter Bagehot, Lombard Street: A Description of the Money Market (New York: Charles Scribner's Sons, [1873] 1897).

7. 12 U.S.C. § 5365(i) (2010).

8. 12 U.S.C. § 5365(d) (2010).

9. The Dodd-Frank Act also increased the deposit insurance limit from $100,000 to $250,000. See Pub. L. No. 111–203, 124 Stat. 1540 (2010), § 335; 12 U.S.C. § 1821(a)(1)(E).

10. Dodd-Frank Act, Pub. L. No. 111–203, 124 Stat. 2113, 2118 (2010), §§ 1101 and 1103.

11. Dodd-Frank Act, Pub. L. No. 111–203, 124 Stat. 2118 (2010), § 1103(b).

12. Federal Reserve Act, 12 U.S.C. 347b; and Dodd-Frank Act, 12 U.S.C. 5465(b). For designated financial market utilities, this would require an affirmative vote by a majority of the Board after consultation with the Secretary of the Treasury.

13. See Acting Comptroller of the Currency Michael J Hsu, "Building Better Brakes for a Faster Financial World (PDF)" (speech at the Columbia Law School, January 18, 2024); and Working Group on the 2023 Banking Crisis, Bank Failures and Contagion: Lender of Last Resort, Liquidity, and Risk Management (PDF) (Washington: Group of Thirty, January 2024).

14. See Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, National Credit Union Administration, Office of the Comptroller of the Currency, "Agencies Update Guidance on Liquidity Risks and Contingency Planning," news release, July 28, 2023. "The updated guidance encourages depository institutions to incorporate the

discount window as part of their contingency funding plans. Consistent with other contingency funding sources, the guidance reinforces the supervisory expectation that if the discount window is part of a depository institution's contingency funding plans, the depository institution should establish and maintain operational readiness to use the discount window, which includes conducting periodic transactions."

15. Board of Governors of the Federal Reserve System, "Federal Reserve Board Announces It Will Make Available Additional Funding to Eligible Depository Institutions to Help Assure Banks Have the Ability to Meet the Needs of All Their Depositors," news release, March 12, 2023.

16. See section 13(3)(A) of the Federal Reserve Act, 12 U.S.C. § 343(3)(A).

17. "Bank Term Funding Program (PDF)," Board of Governors of the Federal Reserve System, March 12, 2023.

18. "Bank Term Funding Program."

19. Board of Governors of the Federal Reserve System, "Federal Reserve Board Announces the Bank Term Funding Program (BTFP) Will Cease Making New Loans as Scheduled on March 11," news release, January 24, 2024. "As the program ends, the interest rate applicable to new BTFP loans has been adjusted such that the rate on new loans extended from now through program expiration will be no lower than the interest rate on reserve balances in effect on the day the loan is made. This rate adjustment ensures that the BTFP continues to support the goals of the program in the current interest rate environment. This change is effective immediately." See "How America Accidentally Made a Free-Money Machine for Banks," The Economist, January 18, 2024.

The views expressed here are my own and are not necessarily those of my colleagues on the Federal Reserve Board or the Federal Open Market Committee. This article is based on a speech delivered at The Roundtable on the Lender of Last Resort: The 2023 Banking Crisis and COVID, sponsored by the Committee on Capital Markets Regulation, Washington, D.C., April 03, 2024.

Building operational resilience at the heart of the financial system

The operational resilience of Financial Market Infrastructures is crucial to UK financial stability. Sasha Mills outlines some key expectations for FMIs to meet the Bank of England's policy in this area

I'm going to take a risk and talk about something recently described in the press as a dull and tedious topic – likely to be of interest only to serious financial market anoraks! I am talking about financial market plumbing. To help bring the topic to life, lets run through a hypothetical scenario together. It's 7am in the morning and my phone's just rung. It's the CEO of a payment system operator letting me know their critical technology systems are down. As a result, they are unable to authorise or settle any new payments.

They don't yet know what has caused the issue – it could be a cyber-attack, extreme weather damaging a datacentre, a critical systems failure while implementing an IT change programme, anything. What they do know is that all around the country customers are standing at tills unable to pay for their morning coffee, many businesses can't buy the materials they need to work today, and it's all over social media. And it is only 7am.

My priority, as supervisor of the payments firm, is that the crucial services the firm provides can be recovered as soon as possible. Firstly, that means diagnosing the problem – so I am looking for the CEO to tell me, *"I know which services are affected, and what the critical components of providing those services are."*

Secondly, that means having contingency plans in place. Say they find out it's a cyber issue, I want the CEO to tell me *"We (the FMI) have prepared for this scenario, we can contain the cyber threat and know how to respond and recover. The services will be back online before there is a major threat to the payments ecosystem."*

The scenario I have outlined provides a clear demonstration of the outcome we are seeking to achieve by March 2025 with the Bank's Operational Resilience Policy – that crucial bits of financial market infrastructure are able to respond to and recover from an extreme but plausible disruption scenario before the market or payments ecosystem it serves is destabilised.

And I would like to talk in a bit more detail about where the focus of such market infrastructure providers should be in terms of building this resilience.

The importance of financial market infrastructure

The scenario above referred to a payment system – which is a type of 'Financial Market Infrastructure' or 'FMI'. FMIs provide the 'pipes' and infrastructure which interconnect and underpin modern financial markets and the real economy.

FMIs are important to financial stability. Firms providing infrastructure for the UK's financial markets and payments are critical to the resilience and safe functioning of those financial markets and the real economy

We all benefit from these services daily – when was the last time you tapped your phone, made a card payment in a shop, or paid a bill with a direct debit? You will likely be more familiar with these sorts of services provided by payments systems, but post-trade clearing and settlement services provided by central counterparties and securities depositories are also critical to the smooth functioning of the economy – they ensure trades are settled and seek to mitigate counterparty credit risk in financial markets.

Confidence in FMI services is critical to having a vibrant and prosperous economy. Households and businesses want to be confident that payments are going through, transactions are being settled, and (in the financial markets) that post-trade activities are completed. And they should be able to have this confidence.

But when the underlying infrastructure provided by an FMI fails, this confidence can be damaged, and this puts financial stability and growth at risk – and that's why the Bank of England (Bank) supervises key FMIs in the UK.

A major focus of our supervisory activity is in maintaining confidence in FMIs through ensuring they are operationally resilient. In other words, we want to be sure that FMIs can continue to provide the vital payments, clearing and settlement services they're meant to deliver even when they are beset by operational disruption.

When we talk about firms being 'operationally resilient', we mean firms can prevent, respond to, recover from, and learn from these disruptions. Disruptions could come from a variety of places. Cyber-attacks are one of the most frequently cited risks to UK financial stability we see in our industry engagement, but we are also concerned about events like natural disasters or operational errors.

Over recent years, the Bank has put in place policies on operational resilience and outsourcing and third-party risk management. We are about to finalise a third plank of these policies later this year with the publication of rules for firms that provide critical services to the financial sector.

But first, I'm going to go through the key principles of the operational resilience policy for FMIs, the work that the Bank and FMIs have carried out so far, and the key outstanding areas FMIs need to focus and improve on ahead of the policy deadline.

How does the Bank define operational resilience?

Coming up with a standard for operational resilience is more complex than simply asking firms to always run flawlessly, across all business areas. Firstly, it is impossible to prevent every disruption or disruptions of every conceivable kind. And secondly, some operations are more important than others.

The first component of our operational resilience policy asks FMIs to identify which business services are important to financial stability – or put another way, services which, if disrupted, could threaten financial stability. Then, we ask firms to say what level of disruption those important business services could experience before risking financial stability, and we call this an 'impact tolerance'.

While expressing impact tolerances in terms of time is necessary to plan for continuity of an important business service, FMIs should consider if there are other metrics that could play a useful role.

FMI's also need to consider how data integrity (or lack of) may impact time to recover – any recovered data that will be used in critical processes, once restored, needs to be checked to be accurate, complete, valid, and reliable. Obviously as supervisors we will probe how FMIs are thinking about these questions – this is not 'one size fits all'.

Having identified the important business services and impact tolerances, we expect FMIs to show they can meet those impact tolerances – that is to recover their services within tolerance – under a variety of extreme but plausible disruption scenarios. Now, having processes and operations which meet this bar doesn't happen overnight, so we have given FMIs several years and a deadline of March 2025 to meet this required standard of resilience.

Before talking about what FMIs have left to do, I wanted to emphasise two points about our expectations: firstly, we assume that some operational disruptions will happen (even though we expect FMIs to have excellent incident prevention mechanisms, there will always be some incidents that are very difficult (or even impossible) to avoid and FMIs need to prepare for that).

Second, we focus on financial stability outcomes, and so don't prescribe which technological solutions or operating models FMIs should use. Resilience is about bouncing back safely when bad things happen as well as minimising the likelihood of an operational disruption occurring in the first place.

Priorities for FMI operational resilience ahead of March 2025

Less than a year out from the March 2025 deadline, there is still a lot of work for FMIs and us as regulators to do. Over the past few years, the Bank has been engaging with FMIs to understand their progress towards meeting this regulatory deadline. We are encouraged by some progress that has been made, however there is still considerable work to be done for many FMIs.

When thinking about how FMIs implement the operational resilience policy, we consider the wider business model and company structure they operate within. The FMIs that the Bank regulates are often subsidiaries of large groups – sometimes internationally active groups. In these cases, the Bank supervises the subsidiary that provides FMI services that are systemic to the UK financial sector, such as clearing and settlement.

While the FMI subsidiary may produce only a fraction of the group revenues, the FMI's services are systemically important to the UK's financial sector, so the continuity and stability of these FMI's services are vital to the UK's markets and financial stability.

FMIs and their parent companies need to ensure that appropriate investment and resources are being directed, within the group, to the UK 'FMI' subsidiary, so that the UK 'FMI' subsidiary can meet our expectations for operational resilience.

Whilst the March 2025 deadline represents a significant milestone, it is also not the end of the story and should not be seen as a 'one off' event – after the deadline, FMIs will need to continue to monitor and improve their operational resilience as risks and technologies evolve.

Cyber threat actors who seek to harm the financial system will not stop developing their techniques, so FMIs need to remain vigilant to the changing threats they are exposed to.

FMIs need to make sure that they are both addressing known vulnerabilities and taking into account changing or increasing risks, for example from increasing digitalisation and the emergence of new technologies – such as Cloud services, Artificial Intelligence (AI), or Distributed Ledger Technology (DLT).

Whilst these emerging technologies can bring efficiencies and improved risk management, FMIs also need to be aware of and manage the risks when these technologies are introduced to their ecosystem– risks from either adoption of these technologies within their businesses or use by customers and suppliers. Some technologies may also heighten threats from malicious actors – such as AI or quantum computing being leveraged to make cyber-attacks more powerful.

What we expect to see over the next year from FMIs

Over the next year, as we approach the March 2025 deadline, we expect to see FMIs accelerating their efforts to ensure that they have calibrated their tolerance for negative impacts on their important business services, and mapped the key people, processes, technology, facilities, and information needed to deliver these services.

FMIs should then be fully testing their ability to remain within impact tolerances for 'extreme but plausible' scenarios – ensuring that response plans and capabilities are robust, and where not, that strategic investment is being made. This is a key requirement.

For the calibration of impact tolerances, we expect to see greater engagement than we have seen thus far between FMIs, their participants, and the wider market. When designing impact tolerances, FMIs should ensure they are considering the impact of disruption to their services on the market they serve – recognising that, where an incident is not contained within a short period of time, this could cause contagion and additional risks to crystallise.

Another area that still requires significant work is the approach and method FMIs use to test disruption to important business services. How FMIs design the scenarios used to test their ability to respond to and recover from an incident, is critical to ensuring FMI's capabilities are adequate.

For example, FMIs should be asking themselves the following questions: are the scenarios extreme enough? How many scenarios are sufficient to ensure the risk has been looked at from several angles? Do the scenarios 'think the unthinkable'? We need to see FMIs prevent incidents where they can, but we also need to know they know what to do when things do go wrong and 'the worst' – so to speak – does indeed happen.

Mature scenario testing requires depth and consistency of approach across scenarios and the design needs to be really clear: the cause of the disruption (for instance is it a cyber-attack or an internal system issue?), the scale of the disruption (how many important business services, participants or transactions are impacted and for how long) and the key risk factors and vulnerabilities that are being tested are clearly set out.

We also expect to see FMIs working to ensure that the 'extreme but plausible' scenarios they have planned for directly link to the risks and vulnerabilities they face and have mapped. This is not an off the shelf set of scenarios.

It's important that the scenarios chosen are indeed of an 'extreme but plausible' scale. What could these be? Well, loss of an important third-party provider, or a severe cyber-attack impacting multiple data centres at once could be a couple of examples.

Testing for these kinds of scenarios helps ensure FMIs are thoroughly testing their response and recovery capabilities. It also means FMIs are challenging assumptions they may be making about the suitability of their response and recovery plans, especially over what will happen over longer timeframes or within heightened impact scenarios.

FMIs need to do further work to improve on the sophistication of their testing approaches, looking for testing methods in addition to tabletop and desktop exercises. Testing types and methods should be as realistic and sophisticated as possible, covering recovery of all critical systems, services, and data – whilst also of course ensuring the testing itself does not introduce any additional risk.

Operational resilience testing should also consider the impact of disruption on the wider eco-system that the FMIs operate in, and FMIs should increase their efforts to involve critical third parties and their participants within their testing. This could be through industry wide tests such as Sector Simulation Exercise ('SIMEX'), as well as tests designed and tailored by the FMI, to test impact and recovery actions, both for themselves and their participants and wider ecosystem.

The Bank expects FMIs to prioritise their efforts on scenario testing over the next year so that they can identify vulnerabilities sufficiently early to remediate them before March 2025. We'll be continuing to look over the coming year for robust remediation plans from FMIs, with appropriate funding and resources dedicated to address weaknesses found during testing.

The speed at which vulnerabilities are remediated should reflect the potential impact to the financial sector that disruption, associated with that vulnerability, would cause.

The broader operational resilience picture

I've spoken at length about our expectations of UK FMIs, and as supervisor for these entities this is obviously a key focus for the Bank. But the broader operational resilience picture does not stop at UK FMIs – similar existing policies also cover their participants like banks and insurers, with the PRA supervising these firms. And we are increasingly also looking at key elements of the supply chain, the UK financial system as a whole, and building international standards.

The Financial Policy Committee (FPC) has set an impact tolerance at the system level for payments recognising how important payments are to the economy and to trust in the financial system. FMIs that provide payments services should consider the FPC's impact tolerance when formulating their own impact tolerances for payments.

The FPC has also recently published its macroprudential approach to operational resilience, which emphasises the vital foundation for system-wide resilience of firm-level resilience and, in support of that, that firms and FMIs should be considering their own roles in the wider system and the effects that their actions can have on financial stability. This is particularly important when firms and FMIs are identifying their important business services and designing their response and recovery plans.

Also, the UK authorities were recently granted powers under the latest Financial Services and Markets Act to create a regime for direct oversight of critical third parties (CTP) following the FPC's view that increasing reliance on a small number of third parties that provide vital services 'could increase financial stability risks', especially given the complex, interconnected financial sector in the UK.

Technology services such as cloud computing and data analytics can bring benefits – enabling digital transformation, catalysing innovation, and potentially providing greater resilience than firms' and FMIs' own technology infrastructure.

We want to ensure FMIs have access to these benefits in a safe way – so the CTP regime will provide the authorities with direct oversight of the third parties that pose the greatest potential risk to the financial system, so we can better ensure system-wide operational resilience.

While we consider this direct oversight to be an important part of our operational resilience toolkit – and a recognition that no single firm or FMI can adequately monitor or manage the systemic risks that certain third parties pose to financial stability – it is crucial to stress that FMIs are still responsible for their own operational resilience. The critical third party's regime in no way detracts from those responsibilities.

Operational threats can come from anywhere in the world and are not limited to jurisdictional boundaries, so we are also working closely with international regulators to share best practice and develop common approaches.

Most recently, the Bank has been closely engaged with the CPMI-IOSCO Operational Resilience Group, which is looking to bolster international understanding of third-party risks facing FMIs, to promote and facilitate the use of existing guidance on cyber resilience and identify emerging risks.

For cyber, the Bank of England has also recently published the results of its most recent CBEST Thematic Test for 2023. CBEST is a targeted cyber-threat intelligence-led assessment, carried out by focused penetration testing on firm and FMI technology infrastructures. It allows regulators and firms to better understand weaknesses and vulnerabilities to cyber resilience and take remedial actions.

The FPC also carries out a programme of system-wide cyber stress testing to build an understanding of the financial system's ability to absorb a significant operational (cyber) incident. The 2022 system-wide cyber test explored a hypothetical data integrity scenario affecting retail payments. The FPC will start the next cyber stress test in Spring 2024, with the findings expected to be published in the first half of 2025.

All these developments are important building blocks of operational resilience of the financial system. But ultimately, this operational resilience starts and ends with systemically important actors in the financial system understanding their responsibilities and ensuring they are prepared for the worst. And that's why I have focused on the operational resilience of FMIs.

Concluding remarks

So, to wrap things up, I'd like to leave you with five key messages. First, FMIs are important to financial stability. Firms providing infrastructure for the UK's financial markets and payments are critical to the resilience and safe functioning of those financial markets and the real economy – and so vital to the UK economic vibrance and growth.

Second, operational disruption will happen. It's very important to the financial stability of the UK's financial system and economy that FMIs are operationally resilient – while we expect FMIs to have processes in place to prevent their occurrence, bad events will inevitably occur, so firms need to be able to respond to and recover from these incidents.

Third, deep and collaborative thinking is required. In their work preventing and preparing to recover from incidents, FMIs need to really think deeply about (and test) their business' impact on both their participants and on the wider financial system. Testing of scenarios should involve FMI participants and critical third parties.

Fourth, detail is important. FMIs must ensure their testing of scenarios is sufficiently mature, incorporating granular and consistent testing approaches. Testing must provide precision around (a) the cause and scale of the incident, and (b) the key risk factors and vulnerabilities.

Fifth, test the unlikely. Think the unthinkable. Yesterday's 'unlikely' may be tomorrow's reality – and FMIs need to consider this when deciding what scenarios are extreme but plausible.

FMIs which do these things will meet our expectations and more importantly will help ensure that the UK financial system functions well in good times and bad – and that individuals' and businesses' confidence in the financial system is maintained. ■

Sasha Mills is Executive Director, Financial Market Infrastructure, at the Bank of England

I'd like to thank Roisin Brennan, Anthony Avis, Charles Gundy, Shane Scott, Simon Morley, Clare Ashton, Amandeep Rehlon, Justin Jacobs, Emma Butterworth, Sarah Breeden, Andrew Bailey, Charlotte Gerken, Andrew Carey, Sean Plumb, Duncan Mackinnon, Adrian Hitchens, Wai Keong Lock, Orlando Fernandez Ruiz, Philippa Cohen, Kiyan Mody and Jon Sepanski for their assistance in preparing these remarks. This article is based on a speech given at the London Institute of Banking and Finance, 30 April 2024.

Cyber risk and its implications for financial stability

The growing digitalisation of the banking sector has increased the exposure to cyber risk. Pablo Hernández de Cos calls for collaboration among financial institutions, regulators, and policymakers to protect against cyber risks

I would like to share with you some reflections on cyber risk[1] and its implications for financial stability. The relentless digitalisation of the economy and of society at large has made it a priority for us all, but especially for the financial sector and its supervisory and regulatory authorities.

The growing digitalisation of the banking business …

The financial sector is extensively digitalised. Banks depend on technology not only as a fundamental support for business processes, but also as a differential and competitive factor.

In recent years, digitalisation has accelerated, both to improve the efficiency of banks' internal processes and to offer their customers flexible, personalised and immediate services, accessible anywhere and via a range of devices.

The COVID-19 pandemic and the emergence of new competitors, such as bigtechs or fintechs, have reinforced these developments.

… has increased the exposure of financial institutions and their customers to cyber risk …

This digitalisation process has intensified the financial sector's exposure to cyber risks. Similarly, by broadening access to remote financial services, it has increased customers' exposure to cyber attacks and digital fraud.

The number of cyber incidents – especially of malicious cyber attacks[2] – has grown continuously in recent years, with the financial industry among the sectors most affected[3]. Moreover, not only are cyber attacks on the rise, but they are also more sophisticated and have a larger potential impact[4], regardless of whether they are strictly economically or geopolitically motivated. In particular, cyber attacks have become more frequent since Russia invaded Ukraine.

By type of cyber incident, cases of fraud that use social engineering, such as phishing[5], smishing[6] and vishing[7], have risen sharply, along with website and mobile app impersonation, among others.

The losses associated with cyber incidents are also significant. For instance, data leaks – one type of cyber attack that has become increasingly common – cost firms $4.45 million on average worldwide in 2023. In the financial sector, data leaks were not only more frequent, but the average cost was also higher ($5.9 million in 2023).

Banks and regulators have worked hard to protect the financial sector against cyber risk, but this must continue to be a priority area of attention in the coming years

Moreover, the possibility of transferring this risk is limited. The financial cover offered by cyber incident insurance policies does not usually extend to all the effects of such incidents[8]. In addition, the terms and conditions of these policies have tightened recently worldwide.

… and can affect financial stability

Financial institutions considered individually have extremely complex technological environments, where old software coexists with other programs that depend on newer technology, the result not only of transformation processes, but also, in some cases, of successive mergers and acquisitions. This complexity makes maintaining an adequate control environment challenging for banks. It also makes them more vulnerable to system failures and cyber attacks.

The financial sector as a whole is also a highly complex ecosystem, consisting of many participants who are very closely interconnected and interdependent. And we are not just talking about financial interconnections: there are also operational interconnections between industry participants, through market infrastructures, common service providers and even the provision of services between financial institutions.

In addition to these traditional interconnections, there are also the new financial service providers to be considered, as well as the growing dependence on technology providers. In many cases this dependence is highly concentrated on a relatively small number of providers.

This is true particularly for cloud service providers, some of which constitute single points of failure. In consequence, cyber incidents affecting these providers – even unintentional ones – can impact the entire sector and thus become systemic[9].

For this reason, macroprudential authorities in different jurisdictions – for instance, the European Systemic Risk Board (ESRB) – include cyber risk among the main sources of systemic risk in the world today[10].

From cyber risk to cyber resilience: towards a holistic approach

Given all the above, it is not surprising that financial institutions and prudential authorities alike are prioritising the implications of cyber risk and the possible mitigating measures.

In this respect, as we move towards a fully digital world in which cyber threats are becoming increasingly frequent and sophisticated, a paradigm shift becomes essential. As does the need to accept that, despite all our preventive efforts, at some point there will be a cyber incident that has an impact.

The idea of cyber resilience stems from the concept of cyber security and is understood as the ability of an organisation to continue pursuing its business, anticipating and adapting to cyber threats and other key changes in its environment, withstanding, containing and quickly recovering from any cyber incidents[11].

In turn, cyber resilience can be seen as an extension of the concept of operational resilience, understood as an organisation's ability to maintain its critical operations in adverse circumstances[12].

It is, therefore, a holistic approach, which is not exclusively focused on managing technology but attaches equal importance to an organisation's people and processes and ties in with more traditional concepts such as business continuity.

Banks are moving towards this holistic approach. In addition to continuously improving technical measures, they are making important efforts to educate and raise awareness about cyber security among their staff, seeking to ensure that they do not become access vectors for attackers.

In the same vein, in recent years banks' senior management have accumulated more knowledge and become more aware of cyber risk, while cyber risk management and audit functions have been strengthened. Banks are also striving to raise awareness about the importance of cybersecurity among their customers.

The necessary regulatory and supervisory response

A broad prudential supervisory and regulatory response is being implemented, globally, across Europe and in Spain, both at microprudential and macroprudential level. Globally, in 2021, the Basel Committee on Banking Supervision (BCBS) approved the Principles for Operational Resilience, which establish that banks should assume as a working hypothesis that disruptions will occur and should define their tolerance for disruption. The principles encompass both preventive and pre-emptive measures as well as those aimed at response and recovery when disruption to critical services occurs[13].

Noteworthy in the European Union is the Digital Operational Resilience Act (DORA), which seeks to mitigate the risks associated with digitalisation and bolster sector-wide resilience by means of:

- Information and communication technology (ICT) risk management and third-party risk management requirements for all financial institutions.

- The obligation to report ICT-related incidents to supervisors so that potential adverse events that may require some form of intervention by the authorities can be detected as early as possible.

- System resilience testing. The most advanced tests consist of simulating cyber attacks, using intelligence on the most likely attackers and their modus operandi. The aim is to assess financial institutions' technical, human and organisational capacities to detect and respond to an attack.

On the microprudential supervision front, European banking sector financial authorities have incorporated cyber risk as one of their supervisory priorities, as a result of which all ongoing monitoring and on-site inspections of banks and horizontal activities targeting cyber risk are being strengthened.

To this end, they have boosted their specialised resources and have established methodologies and working procedures adapted to cyber risk's specific features. The Single Supervisory Mechanism will conduct a cyber resilience stress test in 2024.

In 2017 the ESRB set up the European Systemic Cyber Group, a dedicated task force to study cyber risk's potential impact on financial stability. The different analyses performed show:

- The usefulness of working on cyber resilience stress test scenarios at a systemic level[14].
- The need for systemic cyber incident response plans and for them to be regularly reviewed and tested.
- The desirability of identifying the circumstances in which a systemic crisis may be triggered and establishing thresholds that enable a rapid response and preventive mitigation measures (systemic impact tolerance objectives (SITOs))[15].
- The need to create a pan-European systemic cyber incident coordination framework that fills the current gaps.

Meanwhile, from this systemic standpoint, DORA encourages information sharing between institutions and establishes arrangements for cooperation between financial and non-financial sector authorities. It also tasks the European supervisory authorities with assessing the feasibility of a single EU Hub for major ICT-related incident reporting for all European financial institutions and establishes an oversight framework for those ICT providers that are critical to the European financial sector.

Macroprudential tools to manage cyber risk: from financial capital to technological capital

In parallel, a discussion is under way on the most appropriate tools to mitigate cyber risks. Prudential capital's role in particular has been analysed. In this case, the survival of a bank affected by a ransomware attack that encrypts all of its critical systems would depend on whether or not it has the technical measures in place to allow it to recover. Capital would therefore not be the backbone of its resilience.

However, aside from higher solvency potentially meaning it is easier to fund the means required to recover from a cyber incident, the potential disruptive consequences for the financial system as a whole may sometimes warrant the release of macroprudential capital buffers so that banks can continue supplying credit to the economy.

For example, the systemic risk buffer could be used to distinguish between banks based on their technological systemicity. This could help limit the occurrence of systemic events and subsequent spillover effects.

The materialisation of cyber incidents with a significant financial impact may require financial instruments for crisis management to be used and even new ones to be developed. A bank could see its operations curtailed as a result of a cyber incident, leading to liquidity problems.

Central banks providing liquidity to solvent banks whose liquidity has dried up because of a cyber incident could allow such banks to continue their activity, helping to mitigate the potential risk to financial stability and allowing them to continue providing services to the economy.

In a similar vein, resolution and recovery plans, while not specifically designed for such situations, may be adapted to ensure the continuity of the critical functions of the banks potentially affected by the incident.

In any event, the cyber resilience afforded by a certain amount of capital could probably be achieved more efficiently and effectively by building up technological resources (software, hardware, know-how, specialists, etc.) to render cyber incidents less likely and lessen their financial impact. For example:

- The introduction of circuit-breakers that suspend processes in the event of simultaneous technological and financial crises, limiting their spillover effects.

- Collective support arrangements whereby banks share technological capital, enabling system-wide collaboration between banks to get processes back on track should one fail or, similarly, providing access to data compromised in an isolated cyber attack. Specifically, data vaulting strategies include offline and offsite storage of the information that a bank needs to operate its critical services.

  The most advanced example is Sheltered Harbor, which involves and is supported by the main US banking associations. Participating banks send their encrypted data in an agreed format to shared facilities so that their data can be recovered and processed on a recovery platform.

Future outlook: artificial intelligence and quantum computing

The possibilities provided by new technologies will afford new opportunities to defenders and attackers alike.

In the case of artificial intelligence (AI), content generation capacities will facilitate identity theft and make social engineering attacks much more credible. AI can also help create malware and optimise attacks.

Conversely, it may enable banks to identify cyber threats early through pattern recognition based on large volumes of near real-time data. It may also enable the response to be partially automated, thus complementing the work of analysts and substantially shortening response times.

Meanwhile, it is estimated that quantum computing could mean that many of the current encryption systems will be breached in the medium term. This will affect both the confidentiality and the integrity of the encrypted data, potentially leading to the manipulation of legal history via tampering with signed documents or the creation of validly signed falsified documents.

However, work is already under way to create quantum-safe cryptographic algorithms and to plan the migration of hardware, software and services using potentially vulnerable cryptography to such algorithms.

Conclusions

In sum, although some studies suggest that the financial industry is one of the key sectors best prepared against cyber risk, in part because of its high degree of regulation and supervision, the acceleration of digitalisation, the development of new technologies, the sector's systemic nature and the complexity and growth of technology risk all mean that cyber risk should remain a key focal point over the coming years, and that recent efforts should even be stepped up.

This adaptation will require organisations to recruit the necessary technical profiles. As a result, attracting and retaining talent will remain a challenge for the sector and for the authorities.

Sharing information on cyber threats and cyber incidents will also be key to improving collective defence capabilities. In this respect, the central role of the authorities, which under DORA will receive reports on cyber incidents from banks, will allow them to give the sector useful feedback.

Similarly, cyber resilience stress tests and sectoral crisis management exercises, involving critical providers and other sectors with which there are operational interdependencies, will be essential over the coming years.

In addition, further progress is needed on the quantification and understanding of cyber risks for financial stability and the potential role of macroprudential policies in mitigating them. ■

Pablo Hernández de Cos is Governor of the Bank of Spain

Endnotes

1. Cyber risk is understood as the combination of the likelihood of cyber incidents occurring and their impact. In its Cyber Lexicon, the Financial Stability Board (FSB) defines cyber incidents as events, whether or not malicious, that compromise the confidentiality, integrity or availability of information or interconnected information systems.

2. Cyber incidents include those caused by natural disasters, human error or accidental system failures. Indeed, non-malicious cyber incidents are the most frequent, although the impact of cyber attacks, despite being less frequent, is usually greater.

3. In 2023 the National Cryptology Centre registered 107,777 cyber incidents, compared with 42,997 in 2019. The recently published 2023 National Security Report (Informe Annual de Seguridad Nacional 2023) cites the financial sector, along with energy, ICT and transport, as those with the highest volume of cyber incidents in recent years.

4. We are seeing cyberattacks in the form of transfer fraud, cryptocurrency theft and ransomware (where the attackers demand a ransom in exchange for returning and not disclosing their victims' encrypted information).

5. Phishing attacks are those in which the attacker fraudulently tries to obtain confidential information (passwords, bank details, etc.) from legitimate users by supplanting the digital identity of a trusted bank.

6. Smishing is a technique that consists of a cybercriminal sending an SMS to a user, pretending to be a legitimate bank, with the aim of stealing private information or charging their account.

7. Vishing is a type of social engineering scam over the telephone in which the caller supplants the identity of a trusted company, organisation or person in order to obtain personal and sensitive information from the victim.

8. For example, compensation payments cannot cover some of the most important impacts of a ransomware attack, such as the shutdown of a bank's infected systems, which must be resolved as swiftly as possible.

9. For instance, a large-scale attack on a great number of banks or against a critical provider, or failures in software commonly used in the sector, would be scenarios with a potential systemic effect.

10. See, for example, ESRB. (2020). Systemic cyber risk.

11. See the FSB's Cyber Lexicon.

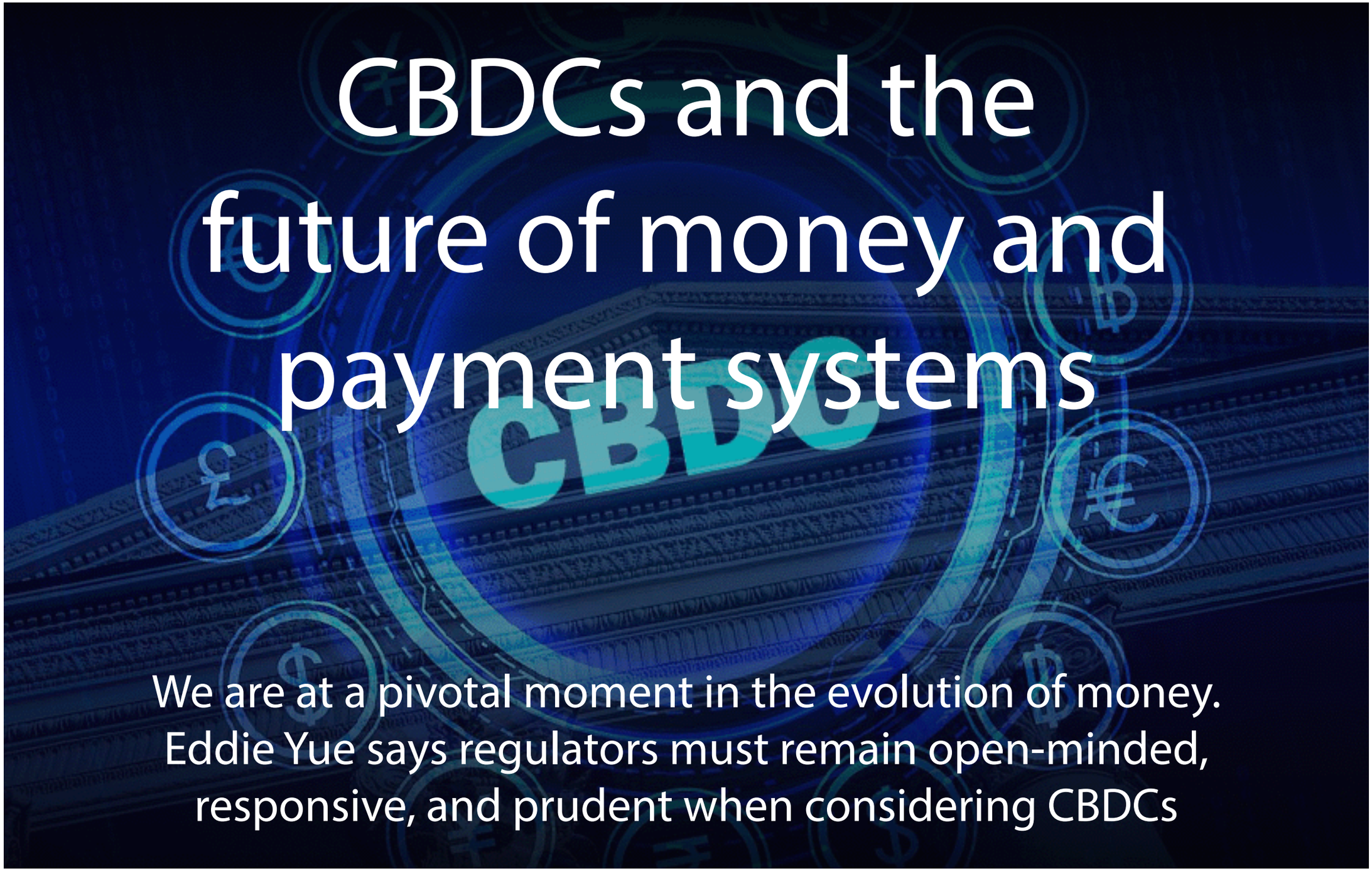
12. BCBS. Principles for Operational Resilience.
13. These principles cover: governance; operational risk management; business continuity planning and testing; mapping interconnections and interdependencies; third-party dependency management; incident management; and ICT, including cyber security.
14. DORA also encourages authorities to organise crisis management and contingency exercises involving cyber attack scenarios, with a view to gradually enabling an effective coordinated response at European Union level.
15. A SITO could be defined for a specific economic function based on the number of transactions affected, their value in euro, the duration of the cyber incident and the number of banks and jurisdictions affected.

This article is based on a speech delivered at the Punto de Encuentro Financiero FINANZA, organised by Elkargi, Bilbao, 12 April 2024.

Finance21 is pleased to announce that Euro Exim Bank has been awarded the Best Trade Finance Bank 2024.

The Finance21 awards celebrate achievement, innovation and excellence across several fields of endeavour. Our award programs are tailored to provide a comprehensive analysis of the very best in each market.

The Finance21 awards are recognised as the principal indications of professional conduct and excellence. The selection panel took into account product innovation, on-going customer support and best practice criteria as well as a continuing commitment to deploying the best possible solutions for the benefit of their clients.



CBDCs and the future of money and payment systems

We are at a pivotal moment in the evolution of money. Eddie Yue says regulators must remain open-minded, responsive, and prudent when considering CBDCs

Introduction

I would like to share my thoughts on central bank digital currencies, or CBDCs, and what it means for the future of money and payment systems. Throughout history, the evolution of money and its institutional foundations have closely followed the advancements in technology. Today, we are witnessing a significant shift in how we view and use money in our increasingly digital world.

The way we make payments is changing, with a decline in cash transactions and a growing trend towards digital payments for goods and services. The developments in the online space are even more interesting, with an explosion of new asset classes and transaction protocols.

In response, some 130 central banks have been conducting research on CBDCs as an advanced representation of central bank money for the digital economy. The HKMA is an early mover in this respect, having started CBDC explorations since 2017 with Project LionRock, which has since evolved into Project mBridge, one of the more advanced explorations of a multi-CBDC platform globally.

Digital money in the tokenisation era

A key technique that has underpinned and enabled this continued innovation is tokenisation, a process of transforming the rights to an asset into a digital token, represented on say a blockchain.

This technique has the potential to enable the interweaving of money and assets, whether in the physical or digital world. This could in turn lead to a major shift in the way we perceive and interact with money and assets.

The HKMA is well aware of how emerging technologies and innovations like tokenisation can potentially enhance the efficiency, transparency, and security of the transactions for financial assets. We have, for example, assisted the Government in issuing two series of tokenised green bonds in the past two years.

We have also been working closely with the industry to see how technological innovation can improve the functionality of money. We invited the industry last year to explore potential applications for an e-HKD, a new form of digital money.

The first phase of the e-HKD Pilot Programme had attracted enthusiastic responses, and we will continue to experiment with the industry on different innovative use cases in the next phase that we have launched last month.

As central banks embrace these new forms of money, it is crucial for us to enable the right financial innovations to flourish while ensuring that the core financial systems that underpin our economy remain stable, trustworthy, and resilient

We observed that tokenisation, alongside programmability, had the potential to transform how money could better serve our financial needs, and that it could create new opportunities for consumers and businesses to transact in new spaces and new ways. These can potentially unlock new avenues for wider inter-industry settlements, facilitate financial inclusion, enable seamless crossborder transactions, and enhance the efficiency of capital markets.

The topics of tokenisation and programmability have since reverberated throughout the industry, with pilot participants proactively sharing their insights, designs and frameworks on how a digital money ecosystem could look like in future, which could encompass retail and wholesale CBDCs in co-existence with other forms of money issued by private firms. And by 'digital money', I intend to also cover different types of tokenised money, capable of supporting the settlement of tokenised assets and transactions.

Publicly and privately-issued money

With this future monetary system, a key question to consider is what role central banks should play in providing 'digital money' to the public, like individuals and corporates. In fact, some research studies show that over 90% of money used in modern economies is private money, issued digitally by private entities like commercial banks.

Digital money should be designed with public interest in mind. Any digital money should enable fast, convenient, seamless, and around-the-clock transactions between consumers and businesses. Over the last decade, we have observed digital money move beyond isolated systems and pure accounting ledgers in banking. What was once confined to basic units of transfer has evolved into something far more sophisticated and powerful, enabling end-to-end integration with different industries.

We envision this close-knit relationship between the public and private money to continue. It goes without saying that the interplay between public and private money is crucial in this regard, especially if we are to realise the full benefits of tokenisation, as well as the full potential of CBDCs and other forms of digital money.

Just as cash and deposits exist harmoniously today, the tokenisation of our financial system will likely involve a complementary relationship between a digital public money, like CBDCs, and other forms of digital private money.

As central banks embrace these new forms of money, it is crucial for us to enable the right financial innovations to flourish while ensuring that the core financial systems that underpin our economy remain stable, trustworthy, and resilient. Sound regulations are essential. And in some areas, we may also need to consider establishing shared infrastructure for public goods, just as we have done with payment systems in the old world.

So far as regulations are concerned, a delicate balance has to be struck. Too stringent an approach may stifle financial innovation, as the industry becomes overly risk-averse or reliant on the central bank to dictate the direction of innovations.

An overly lenient approach, on the other hand, could potentially undermine the central bank's effectiveness as a stable monetary anchor for common citizens, businesses, and financial institutions. This could result in a digital economy dominated by a few closed, privately-run platforms that prevent the emergence of more affordable and efficient solutions.

It is therefore crucial to strike the right balance between enabling innovation and maintaining stability. Central banks should continue to play their role in the provision of public money serving as a trusted monetary anchor, enabling the private sector to confidently develop innovative, value-adding financial products for consumers and businesses.

In this connection, a CBDC should be viewed in the context of these functions of the central bank in the monetary system. To build on the current two-tier structure, central banking institutions like the HKMA should underpin the foundation of our payment systems.

This is a model that we have explored together with the BIS Innovation Hub in 2022 under Project Aurum, where wholesale CBDCs will be used by regulated financial institutions, and customer-facing activities will be assigned to the private sector, just like today under a two-tier system.

Keeping in mind these developments, through the lens of the regulator, it is clear that we need a holistic digital money framework that will facilitate the healthy co-existence of public and private money to support our future digital economy.

A digital money framework

This digital money framework will encompass different forms of tokenised money, which consumers and businesses will likely see and adopt in the years to come. Our intention is to provide the necessary guidance and creative freedom for the industry to experiment with new technologies and develop innovative solutions.

This framework is the collation of our previous work in close partnership with the industry. This guidance is critical in an environment where debates on different forms of digital money are proliferating not just in Hong Kong, but across the globe.

We envision that this digital money framework will comprise three key components from the outset. These components are retail CBDCs, regulated stablecoins, and tokenised deposits.

Retail CBDC

Today, individuals and corporates alike only have access to digital money issued by private entities. As we are transitioning to a digital economy where most transactions

are conducted online, there may be a case for the HKMA to provide a digital form of physical cash, that is a retail CBDC, for everyday payments.

From a more technical perspective, a retail CBDC could serve as a fundamental layer to facilitate interoperability and interlinking between various entities participating in our future digital economy. It could act as a bridge between different types of privately-issued digital money, and ensure all private money to be exchangeable with a public money on demand and at par. A dollar is a dollar, regardless of who is issuing the money, and that preserves the 'singleness of money'.

We also envision that a retail CBDC could be a potential 'backbone' and anchor, bridging a legal tender and digital assets, offering price stability and confidence needed to empower more innovations, and developing a vibrant sector and ecosystem for digital assets in Hong Kong.

Having said that, while a retail CBDC may represent a more advanced version of cash, it remains to be seen whether the benefits of its issuance would outweigh the risks. More research would also be required on how the introduction of a retail CBDC would impact the broader financial system. Given the generally efficient and competitive retail payment ecosystem in Hong Kong, we will continue to take a use-case driven approach in thinking about whether and when to introduce a retail CBDC.

Stablecoins

For those of you who are familiar with Hong Kong's monetary system, you would know that we have a relatively unique system, where the Government, through the HKMA, has given authorisation to three commercial banks to issue banknotes in Hong Kong. These banknotes are issued against payment to the Exchange Fund in US dollars.

The digital version of this arrangement has been explored as a CBDC-backed stablecoin under Project Aurum, and some economists have recently called this a 'reserve-backed token'. This, I would say, is a close cousin of a CBDC, and the difference would be whether the token is issued by the HKMA directly or through the commercial banks.

But of course, stablecoins are more generally issued by nonbanks. Given that stablecoins may be used in payments, the HKMA has been working to bring them into the accepted regulatory parameters. We have recently completed the consultation on a legislative proposal to implement a regulatory regime for stablecoin issuers. Under our proposal, stablecoin issuers would need to ensure full backing of reserve assets, which must be of high quality and high liquidity, so that stablecoin holders will be adequately protected.

Tokenised deposits

We should also not forget about the most commonly used digital money today, which is bank deposits. Deposits today are already digital in nature and are core to a well-functioning banking and payment system.

Deposits can be adopted for use in a tokenised financial system. Simply put, bank depositors may be allowed to convert their deposits into and out of their tokenised form, ie. tokenised deposits, which can circulate on distributed ledger technology platforms.

These tokenised deposits would represent a claim on the depositor's commercial bank, just like a regular deposit. Financial institutions are already using tokenisation to facilitate more efficient liquidity management across borders for corporates, and tokenised deposits would be a natural step forward.

Over the past year, the industry has been trying out the idea of tokenised deposits. Whilst the full potential of tokenised deposits has yet to be explored, we have observed strong interest in this area, and we will continue to work closely with the industry in developing commercially viable use cases.

A wholesale CBDC to support digital money

Finally, let me turn to the prospect of a wholesale CBDC, which could become the instrument that underpins all the different forms of digital money that I have mentioned earlier.

In a way, we already have a wholesale central bank-issued digital currency today, which are the banks' reserves in our Real Time Gross Settlement (RTGS) system. While synchronisation of ledgers achieved via enhancements to RTGS could theoretically deliver atomic settlement, we consider it important to explore a tokenised version of RTGS could support the growing market of tokenisation.

That is why we have announced the development of a wholesale CBDC sandbox under Project Ensemble. Our goal is to explore the technical viability of interbank settlement functionality between different institutions that have developed their own ledgers and systems for tokenised assets. We want to maintain the ability for wholesale financial transactions to settle in central bank money, the safest form of money.

This wholesale CBDC could potentially forge a new financial market infrastructure that bridges the existing gap between different forms of digital money, including a retail CBDC, regulated stablecoins, and tokenised deposits.

Closing

We are at a pivotal moment in the evolution of money, and indeed, in the evolution of our very concept of what 'money' is, and what benefits 'money' should deliver to keep up with the times. The presence of new opportunities provides incumbents with the impetus to innovate, and it can also open the door to better products and services. As regulators, we must remain open-minded, responsive, and prudent to ensure we continue to build on Hong Kong's position as an international finance centre.

As I have mentioned in other occasions, these projects are not and should not be merely technology projects. Their implementation would entail far-reaching implications on a wide range of issues relating to areas such as legal, regulatory, policy, financial stability, privacy, cybersecurity, and interaction with existing payment methods.

That is why the HKMA has been fostering public-private collaboration among the Government, industry and academia. We are conducting a number of research studies with the CBDC Expert Group, which was formed last year, and we hope that the results of some of these research studies will become available to the public later this year. ■

Eddie Yue is Chief Executive of the Hong Kong Monetary Authority

This article is based on a keynote speech at the International Conference on Central Bank Digital Currencies and Payment Systems, Hong Kong, 11 April 2024. The views expressed in this speech are those of the speaker and not the view of the BIS.

Innovation, integration and independence

As payments transition into the digital era, Piero Cipollone argues we need to take SEPA to the next level through the digital euro and private pan-European payment solutions

The Single Euro Payments Area (SEPA) was launched in 2002, aiming to address the fragmentation in non-cash payments that prevailed at the time. Payments between euro area countries were slower, more cumbersome and more expensive than domestic payments. And yet, many market participants questioned the merits of the project: will SEPA make payment services more efficient? Will it make the economy more competitive? And will it deliver real benefits to customers?

Fast-forward to today and it is clear that the initial scepticism was unfounded. We no longer differentiate between national and crossborder payments in euro for credit transfers and direct debits[1]. And people really appreciate the benefits of these two payment services for seamless money transfers across Europe.

However, SEPA has fallen short when it comes to digital payments that are even more central to our daily lives: there is no SEPA at the point of interaction, namely for in-store, mobile or e-commerce payments. Person-to-person (P2P) solutions also remain fragmented.

Most European retail payment solutions are focused on national markets, covering only some use cases and lacking pan-European reach. Because of this fragmentation, crossborder transactions within the euro area have become dependent on a very small number of non-European market players. This hampers competition, innovation and resilience.

Moreover, the digitalisation of payments is undermining the crucial role cash plays in financial inclusion. After all, it is the only means of payment that has legal tender status and can be used by anyone, anywhere in the euro area, free of charge.

As a result, we are once again at a crossroads. And just like in the past, the added value of taking SEPA to the next level is now being questioned: do we *really* need a Single Euro Payments Area at the point of interaction? Do we *really* need a digital euro?

The answer, much like two decades ago, is an unequivocal yes. We cannot afford to settle for the status quo. And we should ask ourselves some hard questions: why aren't European retail payment solutions and platforms able to compete at the global level?

We now stand at a crossroads as payments transition into the digital era, with the risk of crowding out public money, and European providers fail to be competitive on a pan-European, let alone global, scale

Today, the market capitalisation of the largest European bank is several times lower than that of the dominant international card schemes. European payment solutions struggle to compete with these non-European payment providers even within Europe, while in the United States new retail payments companies succeed in scaling up rapidly[2].

In my remarks I will argue that this has to do with the difficulty European payment service providers (PSPs) have in reaching pan-European scale. And I will advocate a comprehensive vision encompassing both public and private retail payments.

Our goal is clear: to further integrate European payments with a view to supporting competition and innovation, while reducing excessive dependencies. Payments offer significant scope to deepen the Single Market in the interest of users and to enhance the competitiveness of European financial services[3].

To emulate the success we had with the launch of the SEPA project, we need to resist the temptation to preserve the status quo. Instead, we must act, relying on the combined knowledge, expertise and efforts of both public authorities and private intermediaries to achieve a single area for retail payments in euro. The benefits in the medium and long run will be much greater than the initial investment costs. The ECB is calling on the payments industry to redouble its efforts.

Retail payments remain fragmented and dominated by a few non-European players

Despite the integration of the euro retail payments market over the past 15 years[4], today's ecosystem is facing three major challenges.

Fragmentation along national lines

First, European payment solutions remain fragmented along national lines. Currently, European solutions for payments at the point of interaction, whether in physical shops, mobile or e-commerce, are scarce and mostly confined within national borders[5].

And we do not have a European digital solution for P2P payments covering the entire euro area. Instead of joining forces and sharing resources to develop pan-European solutions, national communities have often preferred to preserve the legacy of investments made in the past.

Consequently, citizens who live, work, travel or shop online in another euro area country find themselves reliant on very few, non-European solutions. And small companies that consider expanding their business across borders or online may be more reluctant to do so given the need to rely on those solutions and bear the associated costs.

We are thus in a paradoxical situation: the fragmentation of European payment solutions along national lines stands in the way of deepening the Single Market and further digitalising the economy. But efforts to reduce barriers to trade and accelerate digitalisation within the EU generate additional revenue for the few non-European players that currently make it possible to pay in shops and online across Europe, entrenching their dominant position.

Some of the benefits of digitalisation and market integration are thus at risk of not reaching European consumers and instead growing the rents of non-European players.

Limited competition at the point of interaction

Second, the failure of European payment solutions to achieve pan-European scale, and often to even go beyond their domestic market, has resulted in limited competition at the point of interaction. This issue is particularly

pronounced for card payments, which, in terms of value, now account for the majority of retail payment transactions[6].

Their share in the total number of digital transactions has also been increasing, while that of credit transfers and direct debits has receded (Chart 1).

According to the most recent data, international card schemes account for close to two-thirds (64%) of all electronically initiated transactions with cards issued in the euro area[7]. And 13 out of the twenty euro area countries rely on them entirely due to the absence of a national card scheme (Figure 1).

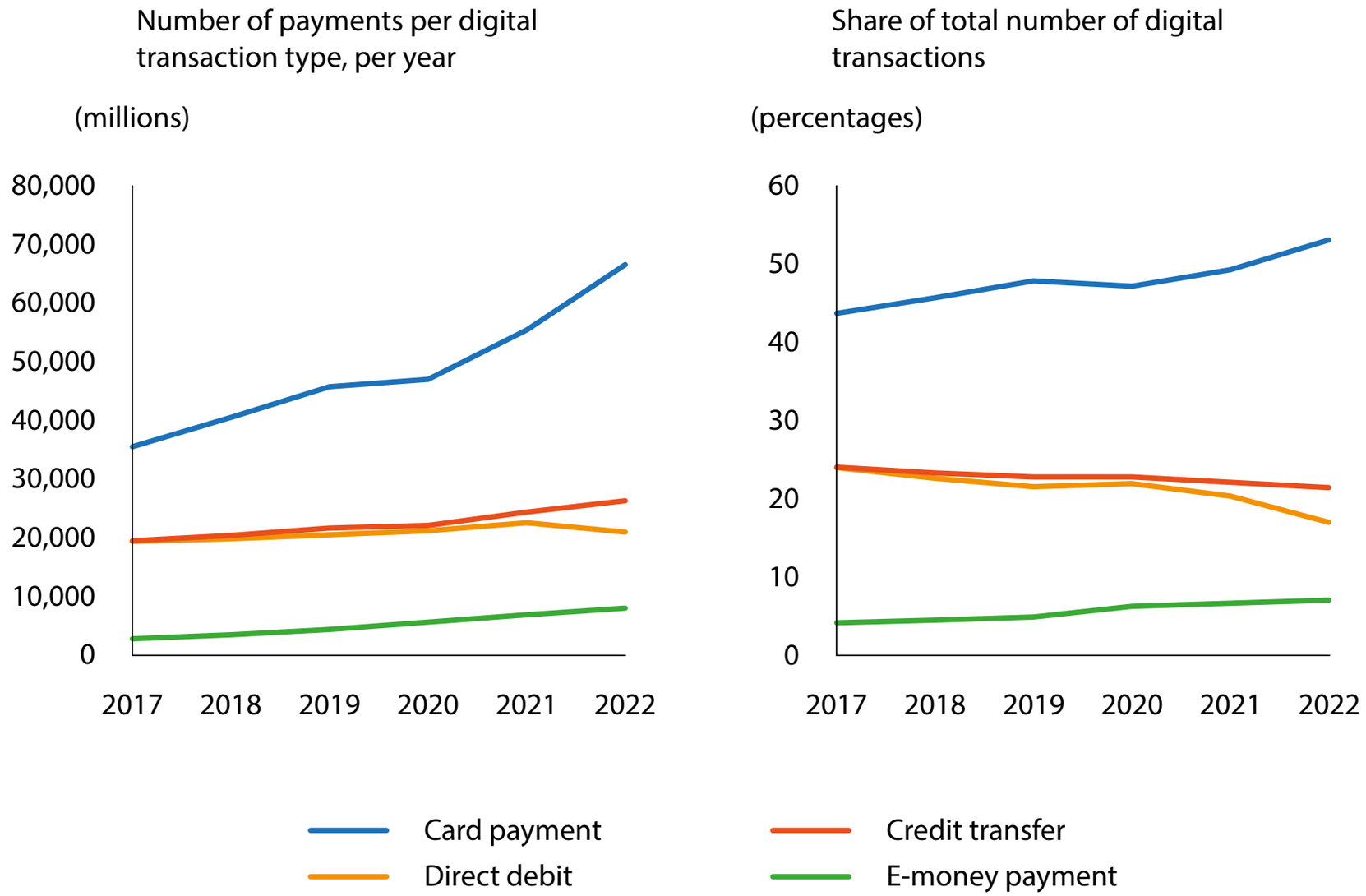
The share of international card schemes is likely to grow further[8], as even the largest domestic card schemes are losing market share[9]. The latter should be wary of this development: while for the time being they maintain a steady revenue stream as card transaction values and volumes increase, this may well change once the market matures.

Competition is also hampered by barriers to entry, which hinder the emergence of new competitors. For instance, in the case of contactless transactions, which are rapidly becoming the new norm in card payments, potential new entrants face the challenge of costly and time-consuming terminal updates.

The lack of a widely available European open near-field communication (NFC) kernel further compounds this issue, forcing new entrants who want to offer contactless payments with mobile devices in stores to either abandon their efforts or depend on existing kernels provided by competitors.

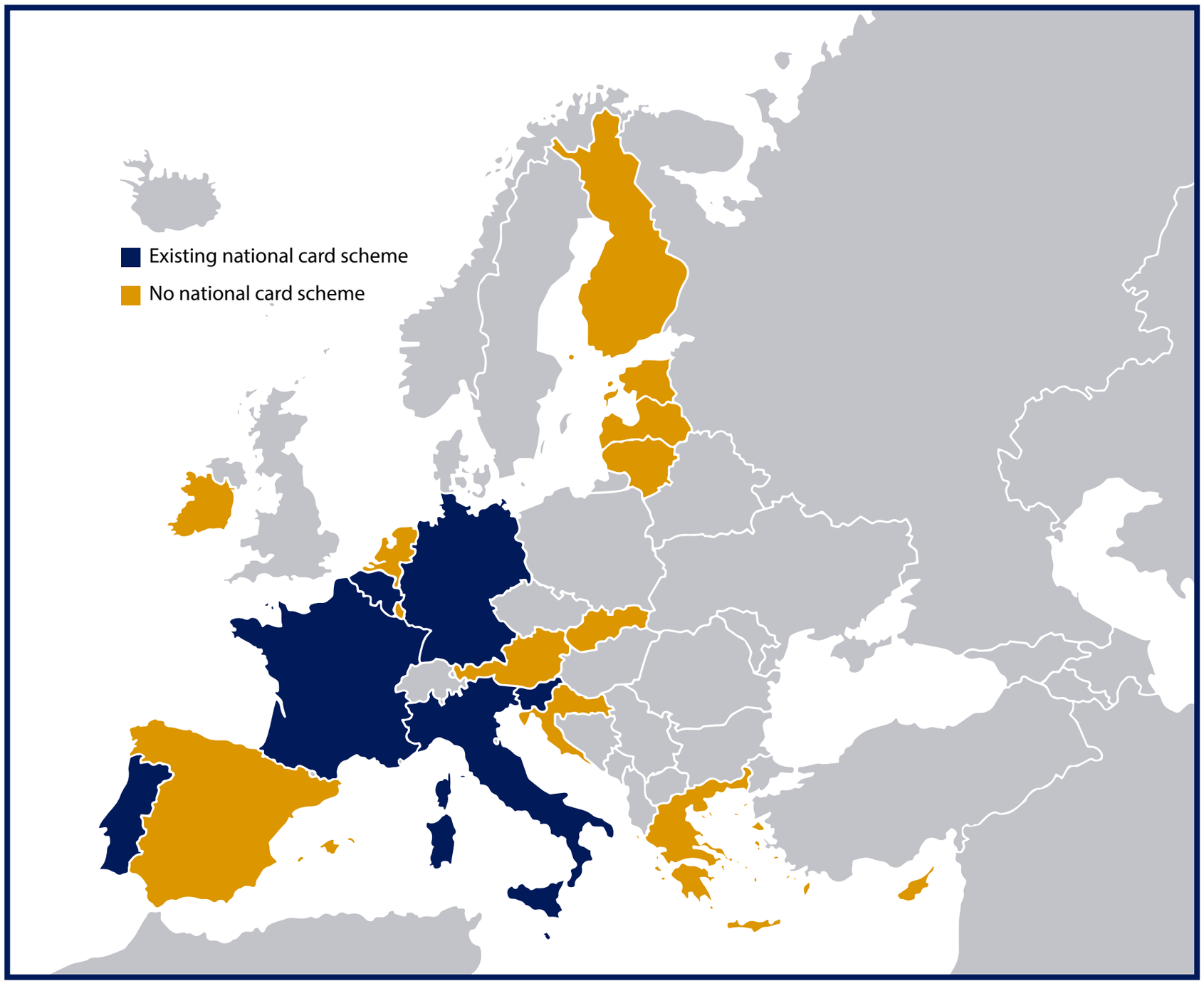
Limited competition in card payments translates into higher fees. According to a recent study by the European Commission, the average net merchant service charges applied by card schemes in the EU almost doubled between 2018 and 2022 (from 0.27% to 0.44%)[10], resulting in significant additional costs for merchants[11].

Chart 1. Payments per digital transaction type in the euro area



Source: ECB statistics.

Figure 1. National card schemes



Source: ECB data.

When those are passed on to consumers, the effect is similar to a consumption tax, albeit one that does not benefit European governments. Furthermore, European merchants criticise the complexity and opacity of card scheme fees, which make it difficult to understand why they are charged so much.

The lack of competition is a problem in other segments, too, such as e-commerce, mobile and P2P payments. While some national initiatives have seen success in specific use cases[12], they fall short when competing with global players on a pan-European scale.

Moreover, big techs entering payments creates further risks, as they could leverage their dominant positions in neighbouring markets and their closed ecosystems. For instance, Apple's significant market power in smart mobile devices and its dominant position in mobile wallet markets on the iOS operating system have raised concerns about anticompetitive behaviour[13].

It led to the opening of a formal antitrust investigation in connection with Apple Pay, the only mobile wallet solution that Apple allows to access the NFC antenna on iOS devices[14].

Dependence on non-European payment providers

The dependence on non-European players is the third major challenge for euro area retail payments. Openness to global competition is essential for fostering innovation. But without a genuine pan-European alternative to international card schemes, payments are more expensive for consumers and merchants. And an overreliance on non-European providers makes our payments and financial system more vulnerable to external disruptions.

European alternatives would improve the resilience of the euro area and the Single Market to such disruptions. And it would increase Europe's ability to set its own standards, rather than depending on standards established elsewhere. Europeans should have more control over an asset as crucial to our economy and society as payments.

The Eurosystem's response: our retail payments strategy and digital euro project

To tackle these challenges effectively, we must take decisive action to move away from the status quo. And I would like to thank Commissioner McGuinness and the European Commission as a whole for their continuous support and legislative ambition in this regard.

At the ECB we envisage a future where retail payments are faster, cheaper, easier and more resilient, thanks to a diversity of pan-European means of payment using European infrastructure. And we do not want this to happen ten years from now, but much sooner.

An old proverb says: *"the best time to plant a tree is 20 years ago, the second-best time is now."* Digitalisation and geopolitics are not standing still. This is why our strategy aims at fostering integration, innovation and independence, all for the benefit of users.

Our proposal encompasses two complementary transformation policies, mirroring the dual pillars of the financial system: public money and private money. These policies are not contradictory by nature; rather, they complement each other and enhance the overall functioning of the European retail payment system.

Our policy on public money

On the public money side, the Eurosystem maintains a steadfast commitment to issuing cash. Our pledge[15] is to ensure that cash remains widely available and accepted as both a means of payment and a store of value. Therefore, the ECB strongly supports[16] the establishment of rules on the legal tender status of euro banknotes and coins across the euro area.

Banknotes have played a crucial role in integrating payments within the euro area for over two decades, by providing a simple and universally accepted payment method. As we transition into the digital age, it is imperative that we preserve the same level of integration and ensure that our currency remains future proof.

There is no reason why public money should not go digital in keeping with all the other forms of payment. We need to adapt to evolving consumer preferences, which are increasingly digital. So the status quo is no longer a viable option. This is why we have launched the digital euro project, currently in its preparatory phase[17]. A digital form of cash holds the promise of preserving the pivotal role of central bank money (Figure 2).

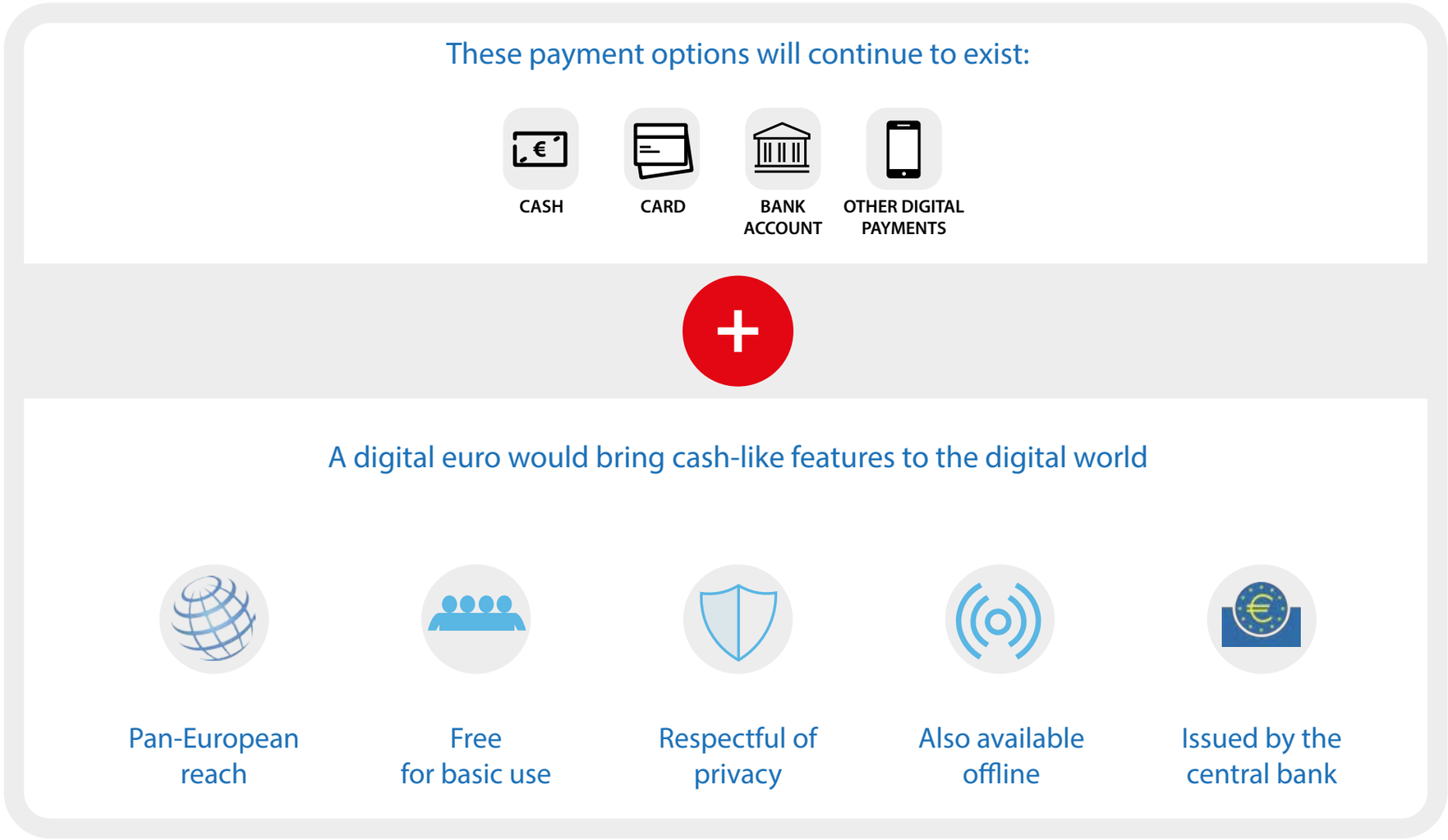
First and foremost, a digital euro would provide unparalleled pan-European reach, ensuring that payments can be conducted seamlessly anytime, anywhere within the euro area, for all types of digital payments (Table 1).

Moreover, as a public good, a digital euro would be provided to citizens free of charge for basic use. Crucially, a digital euro would uphold stringent privacy and inclusion standards, safeguarding user data and rights in the digital age.

Furthermore, a digital euro caters to a diverse range of payment scenarios, covering everything from online transactions to in-store purchases and P2P payments, both online and offline. The offline digital euro will provide a level of privacy very close to cash, while also contributing to resilience and inclusion.

Unlike existing payment methods, a digital euro would offer a comprehensive solution that aims to meet all – rather than just a few – of the evolving needs of modern consumers. Besides covering point-of-sale, e-commerce and P2P payments across the euro area, it would offer seamlessly integrated online and offline functionalities, ensuring that transactions would not be interrupted - even in a situation of limited network coverage or a power outage. No existing payment method offers all these benefits at once.

Figure 2. Digital euro, a digital form of cash



Source: ECB.

Table 1. Availability of the digital euro in all retail payment scenarios in the euro area

Comparison of the availability of the main retail payment methods across retail payment scenarios

	Cash		National schemes (card or account-based)		International schemes (card or account-based)		Digital euro	
	Domestic	Euro area	Domestic	Euro area	Domestic	Euro area	Domestic	Euro area
P2P payments	Yes*	Yes*	Some	No	No	No	Yes	Yes
POS payments	Yes	Yes	Yes**	No***	Yes**	Yes**	Yes	Yes
E-commerce payments	No	No	Some	No***	Yes**	Yes**	Yes	Yes

*Notes: *Only proximity transactions, unless mailing cash. **Where accepted. ***Only through co-branding with international schemes.*
Source: ECB.

Finally, the digital euro will leave no one behind. Promoting digital financial inclusion is a fundamental principle underlying the concept of a digital euro, as also reflected in the relevant draft regulation[18]. The Eurosystem is thus committed to offering a digital euro app in an inclusive and accessible way for people with low digital and financial skills and resources, as well as those with disabilities or functional limitations and the elderly.

While private intermediaries will be able to integrate digital euro services into their own banking apps and wallets, a digital euro app offered by the Eurosystem would not only support accessibility – a feature that is important to consumers[19] – but also ensure that public money remains tangible for people by ensuring a harmonised, baseline user experience across the euro area.

It would be made available in at least all official languages of the EU and be designed such that everybody will immediately recognise the digital euro, just as everybody can recognise euro banknotes today. And smaller PSPs that lack the resources to develop their own front-end solutions would be able to distribute digital euro services through the digital euro app. This app is thus essential for achieving the objectives of the digital euro.

At the same time, the app offered by the Eurosystem would not impinge on the relationship between PSPs and their customers: it would merely provide a uniform point of entry allowing users to interact with their PSP via a smartphone, for example to display information or initiate payments[20].

Moreover, PSPs will be free to provide customised value-added services in their own apps and wallets, going beyond the basic payment functionalities supported by the digital euro app.

Our policy on private money: the retail payments strategy

While the digital euro complements private solutions by giving citizens an additional option for digital payments, it alone cannot resolve all the challenges facing European payments today and in the future.

That's why our vision on payments entails a strategy[21] centred on fostering the development of *privately* operated, European-governed, pan-European payment solutions at the point of interaction.

The Eurosystem supports market-led initiatives that meet a set of requirements it has defined for a European solution at the point of interaction[22]. The ECB therefore welcomed the European Payments Initiative (EPI), which has recently made further significant progress towards a European-grown instant payment solution, including the establishment of EPI Company, development of its brand, completion of acquisitions[23], and the pilot of instant P2P

transactions. The ECB encourages EPI to continue its progress and to expand its geographical coverage to achieve pan-European reach.

Furthermore, the ECB views initiatives by mobile payment solutions and third-party providers favourably, recognising that they may enhance competition at the point of interaction. For instance, a recent collaboration involving three national mobile payment solutions[24] seeks to achieve interoperability in P2P transactions as a first step - and potentially also interoperability at the point of interaction in the coming years. Interoperability could be viewed as an intermediate step towards merging into a single payment solution.

While these European initiatives demonstrate market vitality, we need to avoid fragmentation. The division of consumers and merchants along geographical lines – with national communities joining solutions that cover only parts of the euro area – runs counter to the Eurosystem's vision of *pan-European* reach. This fragmentation would also prevent payment solutions from taking advantage of the sheer scale of the Single Market.

So how can we avoid this undesired outcome and move towards a win-win situation for all payment providers? One solution is to further develop the interoperability between conceptually different solutions. But we would need to see that this approach generates sufficient resources to sustain a common governance, shared functions and product innovation. Furthermore, some countries have national solutions with low market shares while others have none at all.

Therefore, while working to make progress on their current plans, private initiatives and national communities could consider joining forces to create strong integrated solutions and aim for pan-European reach within a reasonably short time horizon.

Although this has not materialised so far, it could be short-sighted to stick to positions taken in the past rather than grasp the opportunities offered by a landscape in transformation.

Our policy on the complementarity of public and private money

The public sector can facilitate such initiatives to reach pan-European scale. In particular, the digital euro could play a key role in shaping open standards. This could allow intermediaries to optimise their implementation strategies and unlock both technological and monetary benefits.

To ensure a seamless implementation of the digital euro and a consistent payment experience across the euro area, the Eurosystem is actively working on a digital euro rulebook. This rulebook will implement common standards in the EU acceptance network. It will be designed to leverage existing standards while also preserving ample flexibility for the market to innovate and develop additional solutions[25].

The digital euro rulebook, along with a robust infrastructure provided by the Eurosystem, would allow private providers to reach pan-European scale with their own payment solutions, achieving cost efficiencies and contributing to an integrated European payment market.

This infrastructure would serve as a catalyst for innovation, enabling the development of new value-added services tailored to customers' needs emerging in the digital age. We envision the digital euro infrastructure as being like a unified European railway network, where various companies can operate their own trains delivering additional services to their customers.

Imagine the following possibilities: innovative front-end solutions designed for conditional payments, functionalities enabling effortless bill-splitting among friends and family, or micropayment applications making it easier to buy online content and services.

These innovations could enhance the overall user experience, enable new business models and drive greater convenience in our day-to-day transactions. Some of these innovative and value-added services are already present in some countries, but it is very expensive for the providers to expand their services across the entire euro area.

On the one hand, the possibility to leverage the open digital euro infrastructure would ensure the necessary standardisation, the lack of which currently hinders innovation; on the other, it would enable private retail payment solutions to launch new products and functionalities immediately on a broader scale. This would give users access to a wider array of services and result in greater competition and innovation on a continental scale. It would also mitigate our current dependence on a few non-European providers.

For example, an open digital euro infrastructure would allow a Belgian citizen to open an account with a payment service provider in Spain that may offer services not yet available on the Belgian market. And new services could be developed such as automatic refunds when rail journeys or flights are delayed.

In addition, private sector players could also review and enrich the portfolio of their payments products in private money, providing customers with new options like rewards, bonuses and subscriptions. Or they could explore the cross-selling of core business products. Also, the competitive rush towards the use of friendly and secure technologies could improve customer experience.

Beyond the point of interaction: strengthening the 'classic SEPA'

But our retail payments strategy extends beyond the point of interaction. The second major goal of the Eurosystem' strategy is to strengthen the 'classic SEPA' framework, which provides the backbone for innovative European payments. SEPA has been a joint undertaking since the beginning. Both private and public clearing and settlement mechanisms have contributed to pan-European reach and resilience.

A key priority within this framework is the full deployment of instant payments. However, their roll-out has so far not enabled users to take full advantage of the important benefits that instant payments could generate. For instance, instant payments give consumers a clearer picture of their finances.

And for businesses, they reduce the amount of money locked in processing, allowing for better cash and liquidity management and reducing the need for overdraft facilities. Instant payments can also trigger faster deliveries and real-time reconciliation of payments, as well as increasing the digitalisation of corporate supply chains.

The ECB welcomes the recent regulation on instant payments[26], which aims to address obstacles such as the fragmented adherence of PSPs to the scheme and limiting transaction fees for payers.

Additionally, ongoing initiatives such as the SEPA Payment Account Access (SPAA) scheme contribute to enhancing independence and innovation. The scheme leverages 'open banking' principles in payments. For a fee, participants can exchange data related to payment accounts and initiate payment transactions with premium features.

SPAA-based payment solutions can provide a variety of account-to-account payment options as an alternative to cards at the point of interaction. The ECB welcomes this innovative European road to 'open banking' and encourages market players to join the scheme.

However, the effectiveness of new schemes can sometimes fall short of expectations. For instance, PSPs implement proprietary solutions instead of using the newly designed SEPA schemes. Or key users discard core features of the instant payments scheme intended to offer functionalities equal to or better than those of traditional cards.

Beyond these particular cases, clarifying the reasons for such shortcomings may be worthwhile. The Eurosystem stands ready to help reflect on how to improve payment-related schemes.

Preparing for the transformation

Addressing these challenges and fulfilling our common vision for the future of retail payments requires industry readiness. The oft-voiced argument that resources are constrained points to the need for a shift in perspective. While past resource allocations may seem entrenched, we are entering a new phase in payments that demands additional efforts, re-allocation of resources and proper planning from industry stakeholders.

Moreover, confining these investments to national level is neither efficient nor sustainable, especially in the light of the significant influence wielded by non-EU players in the digital landscape. Opting for joint EU investments and leveraging economies of scale would enhance the efficiency and effectiveness of our common efforts. And aligning these investments with the introduction of the digital euro would further maximise outcomes.

But it is crucial to recognise that this is not solely about costs. Just as SEPA facilitated the adoption of new standards and honed skills in the payments sector, the integration of private and public solutions at the point of interaction can foster innovation and resilience and benefit the economy.

Conclusion

Innovative, integrated and independent retail payments are crucial components of our monetary system. Indeed, as Tommaso Padoa-Schioppa cautioned over two decades ago, *"Public confidence in the currency could be endangered if retail payment [instruments and] systems were inefficient, impractical for users or unsafe."*[27]

The efforts made since then including the establishment of SEPA and the widespread acceptance of cash as a universally accepted payment method, were crucial for achieving a higher level of integration and efficiency in European retail payments.

However, we now stand at a crossroads as payments transition into the digital era, with the risk of crowding out public money, and European providers fail to be competitive on a pan-European, let alone global, scale.

To address these challenges, we have set up two transformation strategies: the digital euro and the retail payments strategy, the latter focusing on private pan-European payment solutions at the point of interaction.

The digital euro will not only give European citizens more freedom of choice and the ability to pay with a secure solution that is widely accepted throughout the entire euro area. It will also establish a common infrastructure with pan-European reach, on which private intermediaries can build to offer competitive and innovative private payment solutions across Europe.

Today we need to take SEPA to the next level, at the point of interaction, through the digital euro and private pan-European payment solutions. Public-private cooperation can achieve greater integration, innovation and independence in payments, to the benefit of consumers and payment service providers. Together we can recapture the original spirit of SEPA. ■

Piero Cipollone is a Member of the Executive Board of the European Central Bank

Endnotes

1. Direct debits allow customers to authorise companies or organisations to take money directly from their payment accounts to pay their bills.

2. For instance, Stripe – which was established by Irish entrepreneurs – scaled up rapidly in the United States and is currently valued at USD 65 billion.

3. Letta, E (2024), "Much More than a Market – Speed, Security, Solidarity: Empowering the Single Market to deliver a sustainable future and prosperity for all EU Citizens", high-level report on the future of the Single Market.

4. The Payment Services Directive 2007/64/EC entered into force in 2009, creating a harmonised legal framework for the provision of payment services in the European Union.

5. There are notable exceptions of European providers going beyond the country of origin, such as Bluecode, Satispay and a few others.

6. The SPACE 2022 study shows that in terms of value of payments, cards (46%) accounted for a higher share of transactions than cash payments (42%). Card payments were used in 34% of point-of-sale transactions, up from 19% in 2016 and 25% in 2019. For payments over €50, cards were the most frequently used method.

7. Volume share of international card schemes in total electronically initiated card payments with cards issued in the euro area and transactions acquired worldwide for the first half of 2023. Based on data collected under Regulation (EU) No 1409/2013 of the European Central Bank on payments statistics (ECB/2013/43), as amended.

8. Börsen-Zeitung (2024), "Visa is investing billions to increase acceptance", 14 February.

9. Le Parisien (2024), "CB, Visa et Mastercard: la guerre des cartes bancaires fait peur aux commerçants", 22 March.

10. European Commission (2024), Study on new developments in card-based payment markets, including as regards relevant aspects of the application of the Interchange Fee Regulation - Final Report, February.

11. A rough estimate suggests that card schemes' revenues from merchant services charges in the euro area may have increased by more than €7 billion between 2018 and 2022, reflecting both the increase in fees per transaction and the increase in the value of card transactions. In 2022, card payments amounted to €2.74 trillion in the euro area, compared with €1.8 trillion in 2018 (source: ECB Payments Statistics).

12. For instance, Bizum in P2P payments, iDEAL in e-commerce.
13. European Commission (2024), "Antitrust: Commission seeks feedback on commitments offered by Apple over practices related to Apple Pay", press release, 19 January.
14. For details on the investigation, see the Commission's website. See also the letter sent by Piero Cipollone to Commission Executive Vice-President Margrethe Vestager and his letter to Commissioner Thierry Breton regarding the commitments offered by Apple over access restrictions to NFC technology.
15. Further details on the Eurosystem cash strategy can be found on the ECB's website.
16. Opinion of the European Central Bank of 13 October 2023 on a proposal for a regulation on the legal tender of euro banknotes and coins (CON/2023/31).
17. Further details on the digital euro project can be found on the ECB's website.
18. Proposal for a Regulation of the European Parliament and of the Council on the establishment of the digital euro.
19. BEUC (2023), "Digital euro: BEUC's recommendation on the legislative framework for the digital euro", Position paper, September.
20. See ECB (2023), "A stocktake on the digital euro", 18 October.
21. ECB (2023), "The Eurosystem's retail payments strategy - priorities for 2024 and beyond".
22. (1) Pan-European reach and customer experience; (2). convenience and cost-efficiency; (3) safety and security; (4) European brand and governance; (5) global acceptance (in the long run).
23. ECB (2023) "ECB welcomes the EPI's progress on building a European payment solution", MIP news, 25 April.
24. SIBS, BANCOMAT and BIZUM (2023) "Leading European mobile payment solutions MB WAY, BANCOMAT Pay, and BIZUM establish a partnership for interoperability", 14 December.
25. For details on governance and stakeholders, see the ECB's website.
26. Regulation (EU) 2024/886 of the European Parliament and of the Council of 13 March 2024 amending Regulations (EU) No 260/2012 and (EU) 2021/1230 and Directives 98/26/EC and (EU) 2015/2366 as regards instant credit transfers in euro.

27. BIS (2003), Policy issues for central banks in retail payments, March.

This article is based on a speech delivered at the ECB conference on 'An innovative and integrated European retail payments market', Frankfurt, 24 April 2024.

Payments innovation, technical standards, and the Fed's roles

To foster payment system safety and efficiency, Christopher Waller advocates an open and transparent standards development process, where standards are technically sound and supportive of business

From its founding, the Federal Reserve has played a central role in payments. We seek to foster a payment and settlement system that is safe and efficient. Such a system promotes a vibrant economy, since it allows for participation from a broad range of individuals and businesses without requiring them to invest heavily in bilateral arrangements.

Technical standards are the common language used by payment system participants like businesses, financial institutions, payment system operators, and payment service providers. Technical standards can help contribute to payment system safety and efficiency, benefitting the full range of participants.

Technical standards for payments are in no way new—we have had them in place for decades, and those standards make it more efficient to process a check, access cash at an ATM, and carry out other functions less visible to consumers but essential to the payment system.

As payments technology advances at a rapid pace, technical standards must also evolve. And they must do so in a way that supports innovation while promoting safety, efficiency, and interoperability among a growing number of payments products and technologies. I will share some perspectives on technical standards, particularly in the context of evolving payments technology.

Everything old is new again

But before getting to that, I would like to provide some historical context since the benefits of standardization within an evolving payment system predate this current wave of innovation.

Let's start with some early examples from US payments history. The free-banking era from 1837 to 1863 experienced significant fragmentation and corresponding frictions in the payment system as competing banknotes circulated across different geographical regions, requiring costly and inefficient trading of various forms of currency and featuring high levels of fraud and counterfeiting[1].

Collaboration among a broad range of private and public stakeholders can help to establish standards for integrating the new technologies into the payment system

Later, during the national banking era from 1863 to 1913, checks were becoming a predominant form of payment at the same time that the population of the United States was moving westward, requiring a nationwide method to clear checks. The check clearing infrastructure at the time, however, was regional, creating a patchwork network reliant on correspondent banks to move and clear checks across regions, often at high cost and with long delays.

Congress created the Federal Reserve in 1913 and assigned the Fed a number of key roles including issuance of Federal Reserve Notes and, in short order, establishment of a national check-clearing network. These are examples that predate electronic payments where standardization improved the efficiency of payments.

We can also call on examples where we have seen rapid changes in the technology of payments, and technical standards supported innovation by providing a common language for systems to communicate with each other. Immediately after 9/11, planes were grounded, and check payments, which were still largely paper-based, could not be processed[2].

This highlighted inefficiency in the clearing of paper checks. In response, Congress passed the Check 21 Act, which enabled banks to process check information electronically using digital images of an original check. The Check 21 Act created a framework for the substitute check to be provided under generally applicable industry standards and also served as a bridge until industry acceptance of fully electronic checks became ubiquitous[3].

Electronic check processing at such scale necessitated common check image exchange standards, which were developed and adopted through industry and Fed collaboration. And, as I mentioned earlier, the payment system once again is experiencing a technology-driven revolution[4].

Shifting consumer preferences and new technologies have led to new financial products and services. This includes the introduction of instant payments, advances in digital wallets and mobile payment apps, and the emergence and growth of digital assets. Innovators see an opportunity for enhancements in payments, whether through the introduction of new capabilities or finding ways to alleviate frictions such as speed, cost, and accessibility.

While the potential benefits of these advances are significant, disjointed or incompatible innovations could increase fragmentation in the payment system or exacerbate frictions at different points of the payment chain. Will platforms that use newer technologies, for example, be able to talk with one another and traditional payment rails, or will they lead to siloed, parallel networks—so-called 'walled gardens?'

Supporting compatibility or interoperability is important because payment systems are similar to other networks in that greater participation in the network usually means greater value for its users. The network effects in turn can drive a more vibrant economy, since a well-functioning payment system can promote commerce across a broader range of consumers and businesses. While achieving goals like network scale and interoperability is a complex problem, technical standards are one part of the solution to support consistent communication among existing and emerging technologies.

Perspectives on technical standards

A technical standard is simply a common specification or guideline. For example: when I drive across the country, I do not need to worry about whether pumps at different gas stations will fit my gas tank. That is because there is an agreed-upon standard. Or, in the example of payments, financial institutions adopt messaging standards to allow systems and networks to communicate with one another.

If done well, the development and adoption of technical standards can help increase efficiency across payment chains and promote consistency in how disparate systems communicate. This requires standards that are technically sound, implemented consistently, and timed correctly to meet a market need.

Technical standards—such as those developed within the International Organization for Standardization (ISO)—are voluntary, open, and consensus-based. The collective market can identify frictions where standards would be beneficial, and then individual firms can choose whether and when to adopt them based on perceived benefit.

The process to develop the standards is open and transparent, and it requires consensus at every step. This process establishes standards that are seen as credible and legitimate.

Considerations and lessons learned

It is important to note that simply having a standard in place will not automatically translate to economic benefits. To achieve network effects, standards must be widely adopted. Developing *consensus-based* standards implies that all relevant participants are at the table and advocating for their interests, which is not a given, but is critical if standards are to be widely adopted.

Timing is also important. In the early stages of technological advances, standards should not be overly prescriptive to stifle innovation and should be flexible and adaptable as the market changes. Conversely, waiting too long to adopt a technical standard could lead to inefficiencies and high costs to comply after the fact.

Standards also need to be implemented consistently, which requires industry coordination to align on common practices. Building on previous lessons learned, the ISO 20022 messaging standard was developed to be a common

standard for multiple payment rails, with a structured data format, and aims to align messages across payment systems.

Nonetheless, there has been variation in ISO 20022 implementation, and further industry coordination over time will be necessary to align on common practices.

All of this—identifying the need and timing for standards, incorporating a breadth of industry perspectives, and aligning on implementation approaches—requires an immense amount of coordination. And much of this coordination must take place among competitors. So often a neutral convener is needed to achieve better outcomes.

That is why we have standard development organizations like ISO to provide a forum where participants from the private sector, academia, and the public sector can contribute expertise and collectively shape standards in a way that meets business and public sector objectives, providing some comfort that the needs of all market participants and consumers will be considered.

The Federal Reserve and payments technical standards

That brings me to the Federal Reserve's roles in this process—not role but *roles*. One such role is to serve as a convenor, with this week's conference being a typical example. Among standards development organizations, like the Accredited Standards Committee X9 domestically and ISO internationally, the Fed is known for impartial facilitation of efforts to work toward a common goal.

In another role, the Federal Reserve leverages its expertise as a payment system operator to inform future standards and broader improvements in the payment system. For example, in 2015, the Fed issued a call to action in its paper,

Strategies for Improving the US Payment System, asking stakeholders to come together in pursuit of a better payment system for the future.

The paper, and subsequent work, highlighted the role of standards to enhance payments security and efficiency at various points in the payment and ancillary processing flows, noting the importance of supporting industry efforts to develop and adopt standards, as well as common approaches to implementation[5].

The Federal Reserve itself adopts standards in its role as a payment system operator. For example, the FedNow Service, which went live in July 2023, was built using the ISO 20022 messaging standard. Incorporating this standard enables a number of operational efficiencies, including improved information exchange between financial institutions, support for straight-through processing, and reduction in errors and the need for manual processing steps.

We have also committed to adopting the ISO 20022 messaging standard for the Federal Reserve's wholesale payments service, the Fedwire Funds Service, by March 2025. We expect this adoption will improve the speed of wholesale crossborder payments by reducing operational costs and promoting consistent communication across global platforms. The richer data in ISO 20022 could also improve anti-money laundering and sanctions screening and support broader adoption of extended remittance information.

In serving as a neutral convenor, providing payments expertise, and improving payments in our role as a payment system operator, the Fed extends its long history of engaging in the development of payments technical standards.

Conclusion

As we navigate the latest wave of technological innovation in payments, the fundamental payment system dynamics and role of technical standards are not new. However, the pace of change is a lot faster now than in the 1910s, when paper checks were used.

Collaboration among a broad range of private and public stakeholders can help to establish standards for integrating the new technologies into the payment system.

Advocating for an open and transparent standards development process, where participants adopt standards that are technically sound and supportive of business and public objectives, is one way the Federal Reserve can help to foster payment system safety and efficiency, as well as support responsible innovation at its present rapid clip. ■

Christopher J Waller is a Member of the Board of Governors of the Federal Reserve System

Endnotes

1. See Alan Greenspan, "Our banking history" (remarks at the Annual Meeting and Conference of the Conference of State Bank Supervisors, Nashville, TN, May 2, 1998), and Daniel Sanches, "The Free-Banking Era: A Lesson for Today? (PDF)," Economic Insights (Third Quarter 2016): 9–14.

2. See Christopher J Neely, "The Federal Reserve's Response to the September 11 Attacks," Regional Economist, January 01, 2002.

3. See section 3 of the Check 21 Act, Pub. L. No. 108-100 (2003).

4. See Christopher J Waller, "Reflections on Stablecoins and Payments Innovations" (speech delivered via webcast at the 2021 Financial Stability Conference, co-hosted by the Federal Reserve Bank of Cleveland and the Office of Financial Research, Cleveland, OH, November 17, 2021).

5. See Federal Reserve System, "Strategies for Improving the U.S. Payment System (PDF)," last modified January 26, 2015, and Federal Reserve System, "Strategies for Improving the U.S. Payment System: Federal Reserve Next Steps in the Payments Improvement Journey (PDF)," last modified September 6, 2017.

Thank you to Alex Sproveri, Hannah Kim, and Cy Watsky of the Federal Reserve Board and Ainsley Hargest of the Federal Reserve Bank of Minneapolis for their assistance in preparing this text. The views expressed here are my own and not necessarily those of my colleagues on the Federal Reserve Board. This article is based on a speech delivered at the International Organization for Standardization Technical Committee 68 Financial Services 44th Plenary Meeting, hosted by the Federal Reserve Bank of Minneapolis, Minneapolis, Minnesota, May 17, 2024.

Modernising the trains and rails of UK payments

Amidst the prospect of significant technological change in payments, Sarah Breeden sets out how the Bank of England seeks to deliver trust and support innovation, both as a provider and as a regulator of retail and wholesale money

Innovation in money and payments

The money we use to support economic and financial activity, and the payment systems we use as we do so, are fundamental to the Bank of England's role to maintain monetary and financial stability. Money and payments have been no stranger to technological innovation over many centuries[1]. But the pace, breadth and depth of technological change we see now suggests even more radical change may be ahead of us.

My aim is to set out how the Bank of England is seeking to ensure that we will be able both to capture the benefits of these advances and to ensure they are safe. I'll discuss how, as we consider the payments landscape in its entirety, we are increasingly focused on innovation in wholesale payments and the importance of payments innovation by banks.

We'll publish a *Discussion Paper* on these areas this summer to draw on input and collaboration from the private sector, as a complement to the work we've been doing on stablecoin regulation and retail central bank digital currency (CBDC).

Looking back nearly five years ago, Facebook and others proposed a Libra stablecoin that aimed to apply technologies pioneered in the cryptoasset ecosystem to retail payments in the real-world on a global scale. That was somewhat of a bolt from the blue[2].

It reminded us in the UK that, notwithstanding an apparently quick and seamless payments experience for many people (with near ubiquitous contactless card payments and new wallet apps like Apple Pay), technological innovation in money and payments was far from over.

Today, we see examples of interbank retail payment systems around the world – the Unified Payments Interface in India, Pix in Brazil, Swish in Sweden – that are used alongside cards for retail payments in a way that doesn't yet happen in the UK.

Individuals can pay retailers in-store or online out of their bank accounts without going via cards – offering savings particularly for small businesses[3]. And people increasingly make payments between different banks using only the recipient's mobile phone number or a QR code.

I am firmly of the view that the technology revolution will hit – indeed, is hitting – finance in the way that it has hit other sectors of the economy

As we look ahead, we appear to be on the cusp of widespread, more fundamental technological change. Technologies loosely grouped under the broad heading of 'tokenisation' – distributed ledger technology (DLT) and the potential for atomic settlement and programmability, as pioneered in cryptoasset markets – have the potential to offer greater efficiency and functionality for 'real world' retail and wholesale payments. That innovation matters for our role in maintaining monetary and financial stability, and also has the potential to offer significant benefits for customers and businesses, economic activity and growth.

In retail payments (the high volume, low value payments made between households and businesses), our experiments with the private sector and the London Centre of the Bank for International Settlements' Innovation Hub have already shown how use cases for these technologies can embed payments more deeply, automatically and efficiently into our increasingly digital lives[4].

Such technologies can enhance online shopping by enabling a buyer's funds to be reserved at time of purchase and automatically released to the seller only once physical delivery of goods is confirmed. That could enable greater competition in online retail if consumers are more confident to shop online with a new merchant or platform.

Similarly (and perhaps I'm playing to the gallery for those who've travelled long distances today), they could allow commuters to purchase train tickets and be refunded immediately and automatically if the train arrives late, without a need for separate forms and payment instructions. In the context of supply chains, automatic payment upon delivery could help alleviate the perennial challenge of late payments to small businesses[5].

And finally, these technologies could meaningfully reduce the cost of retail payments. Libra was proposing to enable crossborder payments like remittances at a fraction of the 6-7% average cost at the time. And lower cost payments could make micropayments (those for very small amounts) more economical, so that, for example, I pay for the article I read rather than needing an entire newspaper subscription.

I highlight these examples not as a comprehensive description of what these technologies have to offer. Rather, they simply highlight to me that there is genuine potential for these technologies to improve our everyday lives.

Such technologies also offer the opportunity to enhance wholesale payments and settlement – low volume, high value activity between financial and corporate institutions, which includes interbank payments as well as the payments that settle financial market activity.

Here, while the execution of financial market transactions today takes place at speed, its clearing and settlement is far slower, as the multiple entities involved (central securities depositories, central counterparties, custodians, banks and end-investors) each update their records.

DLT allows a single, definitive database to be shared and updated simultaneously across all network participants, rather than each party maintaining its own records. It thus enables all participants in a network to have the same, single source of truth, without the need for manual reconciliation. This can facilitate faster, more efficient processes with fewer intermediaries, and increase the speed at which a transaction can settle – all of which could reduce the risks and costs involved.

Tokenisation could also increase the liquidity of a wider range of financial assets (for example private assets, units in investment funds, or even real estate), enabling them to be held, traded or perhaps used as collateral by a broad set of players, and 'fractionalised' so that investors can hold a portion if they can't afford the whole[6].

The Bank of England's role in payments innovation

I like to think of money and payment systems in terms of trains and rails – money is the train, and the payment system is the set of rails by which that asset travels from sender to receiver. The Bank of England's role for each is two-fold – both provider and regulator.

First, we provide trains and rails of our own. We issue our own money (cash and central bank reserves) and run the Real-Time Gross Settlement (RTGS) infrastructure (the system by which reserves are transferred between financial institutions). These sit at the heart of the economy and financial system, and together deliver trust in money and payments, which underpins monetary and financial stability[7]. They are also an important foundation on which the private sector can provide services and innovate.

Second, we regulate the trains and rails provided by the private sector – the money issued by commercial banks, and the interbank payment systems and card schemes through which that money is transferred[8]. Again, our aim is to deliver trust in money and payments and so monetary and financial stability – by ensuring this private sector activity is safe. In so doing, we support safe innovation.

The technological innovation I've described requires us to look ahead to identify what changes in our role as provider and regulator might be needed. I'd highlight three reasons for that.

First, some of these technologies can reduce risks in the financial system. For example, collapsing the myriad of post-trade clearing and settlement activities and intermediaries into a potentially instantaneous smart contract could reduce, or even remove, counterparty credit and settlement risks.

This could bring some new challenges to manage (the ability to net transactions before settlement is reduced, for example). But the direction of travel towards more efficient post-trade processes in financial markets is clear.

Second, as the central bank, we issue the safest form of money in the economy and that plays a crucial foundational role for monetary and financial stability. And so, as technology advances and new players introduce welcome competition, we need to ensure that we provide central bank money and payments infrastructure with the requisite functionality so that its role as an anchor of confidence in different types of money in the economy is not eclipsed.

Third, a regulated financial system that isn't delivering at the frontier in terms of outcomes for the real economy and financial market participants is vulnerable to new players growing quickly outside of the regulatory perimeter. Such players might quickly get to systemic scale and, in so doing, create risks that are hard retrospectively to address – a phenomenon we've seen as technological innovation has hit other sectors of the economy[9].

Our work on retail innovation

Stablecoin regulation

We were reminded starkly of this final point by the Libra announcement in 2019. We faced the prospect of Facebook harnessing its network of (then) over two billion users to drive adoption both of a new form of money for retail payments (a stablecoin) and of new DLT rails for it to travel on. On both fronts, the risks were not adequately addressed by the regulatory frameworks then in place. And that meant trust in money and payments would not be assured.

While the Libra/Diem[10] proposal came to naught, the possibility of stablecoins coming to be used at scale for retail payments (including by harnessing large firms' existing user bases) remains. PayPal launched a dollar stablecoin with Paxos last year. And Visa is experimenting in the US, enabling merchants to receive payments in Circle's USD Coin.

So in November, we published a *Discussion Paper* proposing a regulatory regime for stablecoins used at systemic scale in retail payments, while the Financial Conduct Authority (FCA) proposed a regime for stablecoins more broadly[11]. We've received valuable feedback from a range of stakeholders in the crypto, payments and banking sectors, as well as from academics and civil society.

Our proposals are focused on ensuring the new train (the stablecoin itself) and the new rails (the system that transfers the stablecoins from payer to payee) meet our and the public's expectations for them to be as safe as those currently used for retail payments (commercial bank money and interbank payment systems respectively).

As with commercial bank money, safety comes from the financial resources stablecoin issuers hold to reduce the likelihood of failure and the resolution arrangements in place in case they do. It will be impossible to provide deposit insurance-style protections for stablecoins, since there isn't a broader industry among which to share the costs. And so that means tighter requirements on financial resources than for commercial bank money.

We have proposed that stablecoins used at systemic scale for payments be backed 100% by central bank deposits. And that these would be unremunerated, as stablecoin issuers would not be engaged in lending or in money market activity, and so not expected to play a role in monetary policy transmission.

In the absence of revenues from backing assets, revenues for stablecoin issuers would have to come from other sources. This could be from fees for the use of the payment rails themselves, as is the case for interbank payment systems, including card companies, today.

Or it could be through providing ancillary services, in the way that Open Banking firms today neither issue money nor provide payment rails, but rather offer users services such as budgeting tools based on access to payments data. (At the risk of straining my metaphor, such firms offer neither the train nor the rail, but rather the onboard concessions delivering services to passengers.) But my fundamental point is that business models wanting to offer a systemic means of payment and earn revenues from maturity and liquidity transformation should be regulated as banks.

In feedback to the *Discussion Paper*, some respondents said these requirements would challenge stablecoin issuers' business models and so might effectively bar use of stablecoins at systemic scale. We'll of course closely consider all feedback received and we will then consult further on a draft rulebook – guided always by the principle that the same risks in different business models need to be addressed to achieve the same regulatory outcome.

Exploration of retail CBDC

At the same time as considering how to regulate new tokenised money for retail payments issued by the private sector, we are also exploring whether to provide such money ourselves. Hence, the Bank of England and HM Treasury (like many other countries around the world) are exploring whether or not to issue a retail CBDC in the UK.

Use of cash in UK retail transactions has already fallen from around two-thirds of transactions twenty years ago to 14% in 2022[12]. And the possibility of large technology companies issuing money risks bearing down further on the role of central bank money in retail payments. To be clear, we will issue banknotes for as long as people wish to use cash. But these trends do raise concerns for the Bank of England's objectives that may warrant a response.

A core tenet of monetary and financial stability is that the public can be confident in the value of money, regardless of its form and issuer, and that all types of money can be easily exchanged, delivering so-called uniformity of money. Wide access to banknotes – alongside bank regulation and provision of RTGS – has provided the anchor for our financial system and our economy for generations - since it ensures different forms of private money (like bank deposits) can always be converted into financially risk-free public money.

The fall to date in the use of cash hasn't yet challenged this uniformity of money. The risk that it might do so in future is hard to quantify and may be remote. But this uncertainty and the severity of the impact should it crystallise is an important motivation for our exploration of a retail CBDC. I firmly believe that, as technological innovation

takes place, we must not forget the contribution that retail central bank money has made to monetary and financial stability.

There is also a risk that new players issuing money may impede other firms' ability to provide wallets and other payment services for it (so-called 'walled gardens'), or issue money in such a way that users cannot transact with users outside of their platform.

This could challenge uniformity of money, since this money might not always be easily converted into other types of money. Such dynamics would also challenge competition in the payments market – and we've of course seen in other sectors of the economy how digital, highly networked markets can present challenges to competition regulation[13].

Given that context, we consulted with HM Treasury last year on a retail digital pound, and we published in January our response to the over 50,000 responses we received[14]. Many people raised important concerns around ensuring privacy and control over how they spend their money should a digital pound be launched[15].

We take these concerns seriously. In response, the Government has committed to introducing primary legislation to Parliament before a digital pound could be launched (and to consult further before doing so). And that legislation would guarantee users' privacy and control of their digital pound payments.

Let me emphasise – we have not taken any decision on whether or not to issue a digital pound. But preparatory work is prudent to ensure the option to issue is available if needed. Our work over the next two years or so will be focused on making a robust and objective assessment of potential benefits and costs, including operational and technical feasibility.

To do that, we will consider the potential design of a digital pound in greater detail, informed by technology experimentation and proofs of concept with the private sector. We will continue to engage widely with external stakeholders. And importantly, we will be informed by how innovation in the wider payments landscape evolves.

Indeed, key inputs into any decision to proceed with a retail CBDC will be the nature and scale of innovation in other forms of retail money (in particular, commercial bank money), as well as how our own wholesale payments infrastructure might evolve in support. Let me cover both in the rest of this speech, starting with our work supporting innovation in wholesale payments and settlement – an area on which we're increasingly focused.

Our increased focus on wholesale innovation

Settlement

I spoke earlier about how tokenisation could drive greater speed and efficiency in post-trade financial market processes, through less need for manual reconciliation and long chains of intermediaries. As regulator of financial market infrastructure (FMI), we're proactively encouraging firms to explore these technologies.

New legislation passed last year gave HM Treasury the power to establish so-called 'FMI sandboxes' – where we modify regulations to allow firms to experiment with new technology. Earlier this month, the Bank of England and the FCA consulted on our approach to the first such sandbox, the Digital Securities Sandbox[16].

This will enable the private sector to set up real-world trading venues and settlement systems using new technologies such as DLT. We will review feedback to our consultation closely, and plan to open the sandbox to applications this summer.

I want to emphasise the 'real world' aspect, because 'sandbox' as a moniker probably undersells the innovation we're making as regulators here. These won't merely be prototypes or experiments. Our aim is to design the sandbox

so that sandbox entrants and the digital securities in it are used in broadly the same way as in the conventional financial system[17].

So this is a really exciting opportunity for market participants and regulators to gain practical insights into the potential benefits of new technologies, while placing limits on activity to enable that innovation to take place safely.

The sandbox will last for five years, during which time the Bank of England, FCA and HM Treasury intend to learn from firms' activity and, subject to that, to create a new permanent regulatory regime for the trading and settlement of digital securities. The government has the tools to put that in place reasonably quickly. In that sense, the sandbox ought not be a 'bridge to nowhere', but rather a means to test and then embed durable and safe innovation.

Payments

As securities are transferred from buyers to sellers, so money needs to be transferred the other way. Our starting point here is that there are clear benefits to financial stability from settling wholesale transactions in central bank money. It is the ultimate risk-free asset and safest form of money in the economy, and it enables an unequivocal discharge of obligations between parties.

Of course, some wholesale transactions settle in commercial bank money today[18]. And commercial bank money is a much safer asset than it was at the time of the Global Financial Crisis (GFC), thanks to reforms to bank regulation and central bank liquidity insurance, as well as the establishment of resolution regimes.

That said, there is a tail risk from too much wholesale settlement taking place in privately issued money. Banks holding money with each other in order to settle payments creates interconnectedness – and this could amplify

a banking stress, as banks pull wholesale deposits from each other, leading to wider deleveraging and losses (dynamics that were clearly at play during the GFC).

Moreover, a payment system settling in commercial bank money would struggle quickly to switch to central bank money in a banking stress, impairing the economic activity being conducted through it.

Settlement of wholesale transactions in central bank money on the other hand can be an anchor of confidence and stability in a stress. We have a low-risk appetite at the Bank of England for a significant shift away from settlement in central bank money towards private settlement assets. And so, as the technology for trading, payments and securities settlement evolves, we need to ensure our wholesale payments infrastructure keeps pace.

We have already made significant headway in renewing the UK's RTGS system. It now operates using global messaging standards, to enable more automatic processing of payments domestically and crossborder. Later this year, it will transition to a more resilient and capable core settlement engine (including with capability to move to near 24/7 operation in future). And we are reviewing how best to expand further the number of financial institutions with direct access to RTGS[19].

At the moment, the tokenised securities transactions in the Digital Securities Sandbox would not be able to settle in central bank money. Rather, they would have to use privately issued settlement assets, such as tokenised bank deposits. We are therefore increasingly focused on how best our wholesale infrastructure should evolve to support the settlement of tokenised transactions, to ensure the continued role of central bank money in wholesale payments.

Already, we have established a new 'omnibus' account in the current RTGS system to allow private sector payment systems using DLT to offer settlement in a tokenised representation of central bank money[20]. Indeed, the first such payment system entered its initial phase of operations at the end of last year under limits from the Bank of England.

Returning to my trains/rails framing, such a solution in effect allows the private sector to offer a train with the safety of central bank money, travelling on privately operated payment rails using the latest DLT technology – even though the underlying central bank money continues to employ a traditional centralised ledger.

More broadly, the Bank of England is exploring with the industry the benefits of extending the renewed RTGS system in future to offer synchronised settlement in a variety of assets, by linking the traditional centralised RTGS ledger to other ledgers, including those using DLT[21].

Money could move on our RTGS system at the same time as tokenised securities move in financial market transactions. Or money in RTGS could move at the same time as money on another central bank's distributed ledger in foreign exchange transactions. Such 'atomic settlement', with the simultaneous movement of money and assets on different ledgers, can effectively extend Delivery versus Payment or Payment versus Payment to a greater range of use cases.

Indeed, in a project with the BIS Innovation Hub last year, we explored how such synchronisation might work in practice to enhance housing purchases. Movement of funds in RTGS could automatically take place at the same time as the change in home ownership is recorded on a digitised title deed, meaning less need for costly and risky chains of intermediaries[22]. (And all this was coordinated by a 'synchronisation operator', which – now probably stretching my analogy to breaking point – is akin to managing the train timetable so that money and assets move on different rails at the same time.)

Meanwhile, some other central banks are experimenting with going further, tokenising central bank money themselves on a separate, distributed ledger – so-called wholesale CBDCs[23]. The question we are keen to consider is how these different models for central bank infrastructure compare, including whether there are important differences in the payment use cases they can support.

The evolution of wholesale payments will also be relevant to our consideration of a retail CBDC, since both could help support the future uniformity of money in the UK. Today, the equal value of money issued by different banks is supported not only by being able to convert them one-for-one into cash, but also by payments between those banks settling across the books of the central bank, in reserves, through RTGS[24].

Similarly, as the technology underlying privately issued money evolves (whether through tokenised bank deposits or regulated stablecoins), this uniformity of money could potentially be bolstered not just by retail CBDC, but also by enhancing the Bank of England's wholesale payments infrastructure so that it can continue to play its key role in supporting settlement between these new forms of private money.

Payments innovation by banks

I want to finish on what the coming wave of technology innovation in payments means for commercial bank deposits – which of course support the vast majority of retail payments in the UK at present.

As Andrew Bailey said in his call to action at Mansion House last year, we don't want the central bank to be left as the only game in town when it comes to payments innovation[25]. In addition to our exploratory work on retail CBDC and our proposed regime for regulated stablecoins, we want to encourage more thinking and, crucially, action by banks in payments innovation.

That needs to cover trains and rails – both how tokenisation might be applied to bank deposits to enhance their functionality across the full range of retail payments use cases, and what interbank payment rails would be needed to support this.

I am firmly of the view that the technology revolution will hit – indeed, is hitting – finance in the way that it has hit other sectors of the economy. It strikes me that, for banks (as the incumbents), payments innovation is both a first-tier opportunity (given the potential benefits for them and their customers of additional payments functionality) and a first-tier threat (given the risk of disintermediation by new players). That in my view demands urgent action.

For the Bank of England's part, we'll support this both as a regulator by setting out clear expectations[26], and as a provider of the wholesale infrastructure via which interbank payments ultimately settle. On the latter, we are experimenting with six other central banks, the Bank for International Settlements and the private sector on 'Project Agorá', which will explore how tokenised commercial bank deposits could be integrated with tokenised wholesale central bank money to enhance payments efficiency and functionality, including crossborder[27].

To pause for a moment on this international dimension before I conclude, central banks all around the world are facing questions similar to those I've set out today. Projects like Agorá will help us pool our resources in tackling them.

But importantly, they will also help ensure we tackle them in a way that is coherent, helps reduce rather than add to frictions in crossborder payments, and importantly (and definitely a subject for a speech on another day) helps us understand the implications for the international monetary system should digital money lead to more frictionless cross-border money holdings and payment flows[28].

Conclusion

New technologies offer the prospect of enhanced retail and wholesale payments functionality in the UK. Our immediate focus at the Bank of England following Facebook's Libra announcement in 2019 has rightly been on establishing regulatory regimes for stablecoins to be used safely for retail payments, and on considering whether a retail CBDC might be needed in future.

But we are increasingly complementing that work with a focus on how we can best support innovation both in wholesale payments and financial markets (including through modernisation of the Bank of England's wholesale payments infrastructure) and in banks' retail payments offerings.

To do that will require input and close collaboration from the private sector, and so we are planning to set out our thinking in these areas in a *Discussion Paper* this summer[29]. ■

Sarah Breeden is Deputy Governor for Financial Stability at the Bank of England

Endnotes

1. Indeed, my colleagues David Rule and Iain de Weymarn have written recently about the lessons we can draw from the late seventeenth century (when paper money first began to circulate alongside silver coin in the UK) for the innovation in money we're seeing today. New money, old money – Bank Underground.

2. As my predecessor Jon Cunliffe described in his final speech as Deputy Governor for Financial Stability in October: Money and payments: a 'black ships' moment? - speech by Jon Cunliffe | Bank of England.

3. Small businesses in the UK pay on average almost five times the fees paid by the largest retailers to accept card payments - close to 2% of the transaction value. See Chart B.3, The digital pound: a new form of money for households and businesses? Consultation Paper (bankofengland.co.uk).

4. Project Rosalind: developing prototypes for an application programming interface to distribute retail CBDC (bis.org).

5. Time is Money | FSB, The Federation of Small Businesses; Prompt payment and cash flow review - GOV.UK (www.gov.uk).

6. Unlocking the power of securities tokenisation - UK Finance.

7. In his Mansion House speech last year, Andrew Bailey set out two important foundations underpinning public trust in money – the uniformity (or 'singleness') of money (wherever we hold our money – in bank accounts, notes and coins etc – we can be assured that it all has the same value) and 'finality of settlement' (when we pay for something, we can rest assured that it actually has been paid for). Speech given by Andrew Bailey at the Financial and Professional Services Dinner, on Monday 10 July 2023 (bankofengland.co.uk)

8. In the UK, the Bank of England supervises systemic payment systems, and their systemic service providers, as part of its financial stability objective. The Payment Systems Regulator (PSR) also supervises UK payment systems to promote competition, innovation and the interests of their users.

9. Londoners overwhelmingly against TfL decision to ban Uber, analysis of social media posts reveals | London Evening Standard | Evening Standard.

10. Libra was rebranded as Diem in December 2020.

11. FCA and Bank of England publish proposals for regulating stablecoins | Bank of England
12. UK Payment Markets 2023 | Policy and Guidance | UK Finance.
13. Unlocking digital competition, Report of the Digital Competition Expert Panel - GOV.UK (www.gov.uk).
14. Bank of England and HM Treasury respond to digital pound consultation | Bank of England
15. People's use of digital payments today also generates personal data, and there is a robust legal regime in place around the access that the public and private sectors can have to that data.
16. The Bank of England and the FCA issue joint consultation and draft guidance on the Digital Securities Sandbox | Bank of England
17. Provided entrants have completed the requisite testing and engagement with supervisors.
18. This is due to a range of reasons, including some firms not having access to the central bank balance sheet or having cost-effective arrangements with private settlement banks. That said, the barriers to entry for direct access to the UK wholesale payment system CHAPS (such as the cost of technical infrastructure required) has reduced in recent years and, we expect, will continue to decrease. Hence, the Bank of England is reviewing whether the set of financial institutions with access to RTGS should be widened further - Reviewing access to RTGS accounts for settlement | Bank of England.
19. The Real Time Gross Settlement service: an open platform to drive innovation - speech by Victoria Cleland | Bank of England
20. Bank of England publishes policy for omnibus accounts in RTGS | Bank of England
21. Synchronisation is one of the key priority areas for further work as part of the Future Roadmap for RTGS once the new core ledger and settlement engine go live later this year. We are now working closely with industry to assess business cases and define the high-level design of priority features before deciding which features to implement and in what order. Future Roadmap for RTGS | Bank of England
22. Project Meridian: innovating transactions with synchronisation (bis.org).
23. For example – the Banque de France, the Swiss National Bank, the Monetary Authority of Singapore, the Reserve Bank of India, the Central Bank of Brazil, and the Hong Kong Monetary Authority.

24. Uniformity of money is also supported by having robust regulatory regimes for private issuers of money (to ensure that banks are robust enough to always honour their commitment to convert deposits into cash one-for-one).
25. Speech given by Andrew Bailey at the Financial and Professional Services Dinner, on Monday 10 July 2023 (bankofengland.co.uk)
26. Letter from David Bailey, Nathanaël Benjamin and Vicky Saporta on 'Innovations in the use by deposit-takers of deposits, e-money and regulated stablecoins' | Bank of England
27. Project Agorá: central banks and banking sector embark on major project to explore tokenisation of crossborder payments (bis.org).
28. The Bank of England is also closely involved in efforts coordinated by the G20 to enhance crossborder payments. G20 Roadmap for Enhancing Cross-border Payments: Consolidated progress report for 2023 - Financial Stability Board (fsb.org).
29. This will also support HM Treasury's work this year in setting out a National Payments Vision.

I would like to thank Michael Yoganayagam for his assistance in drafting these remarks. I would also like to thank Martin Arrowsmith, Andrew Bailey, Paul Bedford, Emma Butterworth, Diana Carrasco Vime, Victoria Cleland, Amy Lee, Sasha Mills, Ali Moussavi, David Rule, Danny Russell, Vicky Saporta, Cormac Sullivan and Danny Walker for their helpful input and comments. This article is based on a speech given at the Innovate Finance Global Summit 2024.

Opportunities and risks of fintech and AI

Recent developments in technology and AI provide huge potential for innovation and productivity growth. Randall Kroszner discusses the opportunities and risks of financial technology and AI

It's a pleasure to discuss a topic that has the power to transform the financial services landscape and, with it, the way we think about financial stability risks. I have approached this subject with my two Bank of England hats on: as an external member of both the Bank's Financial Stability Committee (FPC), charged with identifying, monitoring and mitigating systemic risks; and the Financial Market Infrastructure Committee (FMIC), which supervises financial market infrastructures.

Each Committee has a role in protecting and enhancing financial stability in the UK. Both are alert to the opportunities and risks presented by financial technology and artificial intelligence (AI). I intend to talk about how I think both Committees should approach developments in these areas.

My main point is that the UK needs to embrace opportunities for innovation and productivity growth. This means taking seriously the secondary objectives for each Committee: for the FPC that is supporting the economic policy of the government; and for FMIC it is facilitating innovation in the provision of FMI services – in addition to our financial stability responsibilities.

Productivity matters for all of us. Higher productivity means stronger economic growth, higher real wages, increased profitability and a boost to tax revenues[1]. The United Kingdom's (UK) weak productivity growth in recent years has long been discussed.

In the decade from 2012 to 2022, for example, the growth rate of output per hour averaged 0.5% in the UK, whereas it was double that rate in the US and the OECD as a whole[2]. Since the start of 2023, output per hour has grown by an average of 0.6% per quarter in the US, whereas it has contracted by an average of 0.1% per quarter in the UK[3].

Recently, much has been made about the role technology investment and innovation can play in explaining differences in productivity performance – and this is something we as regulators and policymakers should take seriously[4,5].

Ensuring both financial stability and innovation, however, is particularly challenging when we are dealing with the potential for fundamentally disruptive innovation that AI could bring versus the more traditional case when innovation and change is more incremental.

The challenge is to develop a regulatory framework that fosters the flowering of creativity and innovation but takes into account the potential financial stability risks

In this article I first discuss the importance of innovation and the potentially fundamentally disruptive impact of AI. I then draw a distinction between the regulatory challenges for dealing with more traditional incremental innovation versus fundamentally disruptive innovation.

This discussion then led me to sketch a framework for thinking about AI. In particular, I will focus on two of the many issues related to AI, namely, interpretability of the models and the potential for misalignment. Large language models (LLMs) involve complex dynamic algorithms, interactions, and weightings that are often extremely difficult to interpret to be able to give an 'explanation' of how the model produced a particular result or outcome[6].

I draw an analogy to the 'invisible hand' of the market that acts as a type of discovery procedure that generates innovations in products and services that can be similarly difficult to explain.

Finally, even though much of the terrain here is new, and often the challenges can seem daunting or even difficult to contemplate that's not an excuse for inaction. For one thing, we can draw on the lessons of past experiences. In addition, there are existing areas where policymakers can act to ensure the landscape around technology and AI developments is one conduce to both innovation and financial stability.

The importance of innovation

There is much the Bank can do to support and promote the innovation and productivity gains exciting new technologies can bring. Indeed, the Financial Services and Markets Act 2023 introduced a new secondary objective for the Bank, through the FMIC, to facilitate innovation in the provision of central counterparties and central securities depositories services with a view to improving the quality, efficiency and economy of the services[7].

That gives me, as a policymaker and regulator, a clear aim to ensure firms, and the services they offer, are able to evolve with the world around them, while maintaining their resilience in line with the Bank's financial stability objective.

Change is already occurring in the financial services world with the widespread adoption of financial technology. According to the *2022 Machine Learning (ML) Survey* conducted by the Bank and FCA, 72% of financial services respondents reported using or developing ML applications.

Firms are predominantly developing or using ML for customer engagement (28%), risk management (23%), and support functions like human resources and legal departments (18%). Industry engagement suggests that firms, particularly large traditional FIs, are typically using ML to improve their overall efficiency and productivity.

There are many estimates of the boost AI and technology can give to productivity growth. A recent report by Goldman Sachs suggests that generative AI could raise annual US and UK labour productivity growth by just under 1.5 percentage points and raise annual global GDP by 7% over a 10-year period following widespread adoption[8].

McKinsey finds that generative AI could enable global labour productivity growth of up to 0.6 per cent annually by 2040 depending on the rate of technology adoption and how workers are redeployed[9].

The precise impact AI and technology will have on the economy therefore comes down to a question of speed and scale, with lots of uncertainty. It has often been noted, for example, that the steam engine was patented in 1769 yet it took another 60 years before steam was able to match water as a source of power in the British economy[10].

Chad Syverson, my colleague at the Booth School of Business, provides a useful perspective on the question of timing and measurement that I think is worth bearing in mind. He notes that if new technologies (like AI) create a significant amount of investment in intangibles then, given the way we usually measure things, there will be a particular pattern to measured productivity: it will understate true productivity growth early in the diffusion of the new technology and overstate it later. He and his co-authors call this the Productivity J-Curve (Figure 1)[11].

Syverson reaches the optimistic conclusion that the developments of the past couple of years suggest we may be at the point where – in the US at least – measured productivity growth might start to understate true productivity growth.

In the US, he notes that gross labour flows – hires plus separations as a share of employment – are about 10 percent higher than their 2015-19 average. And within separations, the ratio of quits to layoffs is at historic highs. This, Syverson argues, can be associated with productivity growth. That gives us something to monitor closely over the next couple of years.

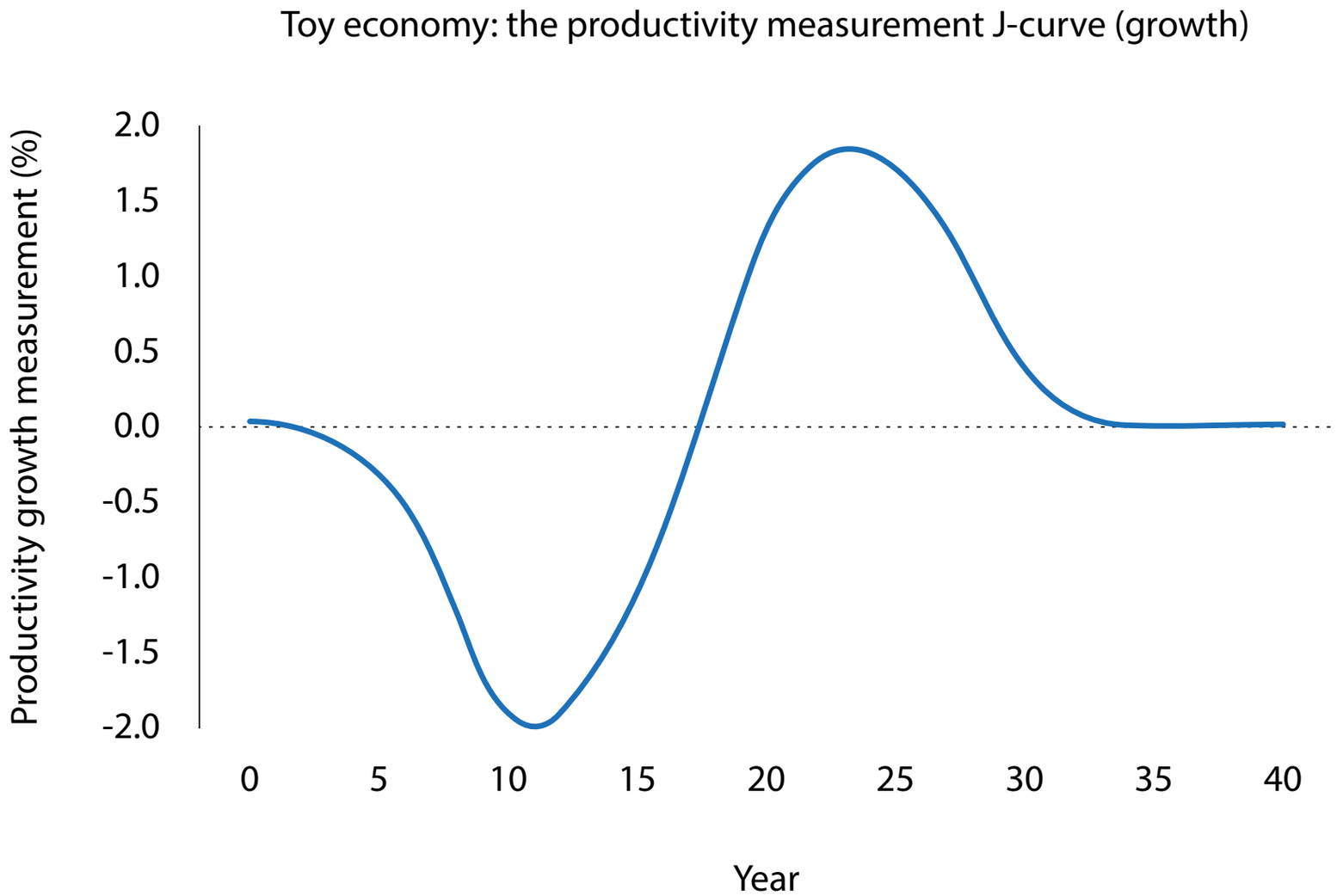
Fundamentally disruptive versus incremental innovation and change

So, if technology innovation and AI have the ability to unleash productivity growth – which is great for the FPC and FMIC's secondary objectives here at the Bank – where does that leave us in meeting our financial stability objective?

The first point I want to make here relates to the potential pace of change. The FPC was established following the global financial crisis (GFC) and charged with identifying, monitoring and mitigating systemic risks.

We can take action – such as increasing the UK's countercyclical capital buffer (CCyB) rate or intervening directly in markets like we recommended the Bank did during the Liability-Driven Investment (LDI) crisis – to maintain

Figure 1. The Productivity J-Curve, Stylized



Source: Chad Syverson (2023), 'Structural Shifts in the Global Economy: Structural Constraints on Growth': Remarks at the 2023 Jackson Hole Symposium.

stability. But it's a constant monitoring process, often examining relatively small changes in the data, to gauge if and when it's appropriate to take action.

When innovation is incremental it is easier for regulators to understand the consequences of their actions and to do a reasonable job of undertaking regulatory actions that align with achieving their financial stability goals. Of course, there is always the possibility of unintended consequences but feedback from market participants and industry will help make regulators aware of those.

When innovation is incremental, data from (recent) activity can provide some guidance, both for market participants and for regulators, about the likely impact of the innovation and allow at least a rough costs and benefits analysis of the regulation. In some sense, given that innovation is incremental, recent experience can provide a framework for discussion and debate – similar to how the FPC currently considers the appropriate setting of the CCyB.

But when innovation is disruptive it is much more difficult for regulators to know what actions to take to achieve their financial stability goals and what the unintended consequences could be for both stability and for growth and innovation.

Recent data thus may not be particularly illuminating. Perhaps there can be some analogies to past 'big' innovations (I've already made reference to the steam engine), but any framework would have much greater standard errors.

There might not be a common framework for either assessing the likely impact of the innovation or the consequences (intended and unintended) of regulatory action. In this state of the world, disagreements risk being more fundamental about how to achieve financial stability and the dialogue between firms, regulators and others can lack clarity and understanding.

Regulators, however, should be open to new approaches that might shape these frameworks. These can support safe innovation, as is the intention of the Digital Securities Sandbox (DSS) that we are consulting on along with the FCA. The DSS is a regime that will allow firms to use developing technology, such as distributed ledger technology, in the issuance, trading and settlement of securities such as shares and bonds. The DSS lasts for five years and will help regulators design a permanent technology friendly regime for the securities market.

This initiative is a great way to help provide a glidepath to a potential new technology friendly regime in this area. But fundamentally disruptive innovations - such as ChatGPT and subsequent AI tools – often involve the potential for extraordinarily rapid scaling that test the limits of regulatory tools. In such a circumstance, a sandbox approach may not be applicable, and policymakers may themselves need to innovate further in the face of disruptive change.

Invisible hand of the machine: interpretability and misalignment

In the context of the debates about the opportunities and risks of fundamentally disruptive innovation of AI, a key concern relates to the 'interpretability' of models, namely understanding how and why a model generates the outcomes it does, and this may become increasingly difficult the more advanced AI gets[12].

AI expert Stuart Russell describes deep learning systems as *"black boxes – not because we cannot examine their internals, but because their internals are largely impossible to understand."*[13] If we can't fully understand the technology, what does this mean for financial stability?

In the way I approach the issue, this is analogous to the challenges of explaining the 'how and why' of many innovations that arise from market competition – the market as a 'discovery procedure' as Hayek famously described. Often the 'eureka' moment is a mystery: how was there a leap to something new?

Polanyi and Hayek underscore the tacit or inarticulate knowledge fundamental to market (and I would argue also in the non-market) discovery processes much like the 'tacit' or 'inarticulate' knowledge in the algorithms and data weights of the LLMs[14,15,16].

So I believe there is a parallel between the 'invisible hand' of the machine or LLM and the discovery process that generates new ideas and new products never previously conceived of. The difficult-to-interpret complexities and dynamics of the LLMs share elements of the tacit or inarticulate knowledge of market (and non-market) human processes as both solve problems and generate innovations in ways that may be challenging to explain[17].

As with the market, just because we cannot fully understand and explain the 'how and why' does not necessarily imply that there is a problem. Much innovation and productivity outcomes could be lost if we only permit results that come from models that we can fully interpret – much like we do not reject innovations where the 'Eureka' moment cannot be fully explained.

We should also acknowledge that explainable AI is a focus of significant research and what we mean by explainability may have to evolve from how we've thought about it in the era of causal effects and regression modelling. This is potentially a new era and regulators should be engaged in understanding these developments.

I also want to say a word about misalignment – that is a concern that as soon as AI systems can act and plan in accordance with some specific goals they may, no matter how benign they are initially, begin to become misaligned with humanity's needs and values in the pursuit of their key objective[18].

While misalignment is not always inevitable, it is clearly something the FPC, as a committee inherently focused on risks, should consider. Indeed, recently my FPC colleague Jon Hall highlighted the potential risks emerging from

neural networks becoming what he referred to as 'deep trading agents' and the potential for their incentives to become misaligned with that of regulators and the public good. This, he argued, could help amplify shocks and reduce market stability.

This issue of misalignment is one policymakers and regulators will need to grapple with. Jon makes one proposal to mitigate this risk, arguing that neural networks should be trained to respect a 'constitution' or a set of regulatory rules that would reduce the risk of harmful behaviour. I am relatively optimistic about our ability to approach this issue and am receptive to Jon's way of resolving this.

Indeed, in the context of the disruptive change mentioned above, perhaps his idea of a 'constitution' could be combined with, and tested in, a sandbox as way of shepherding new innovation in a way that supports financial stability. In the cases where fundamentally disruptive change scales so rapidly that a sandbox approach may not be applicable, a 'constitutional' approach may be the most appropriate one to take.

So, for me, at least some of the interpretability and misalignment challenges of the AI and the LLMs are not new but familiar territory but in a different context. Nonetheless, given the potential for rapid scaling and the changes that can engender, it still poses challenges regulators and markets must consider.

Operational resilience

One way we as policymakers and regulators can lay the groundwork now for future challenges is through operational resilience. By this I mean the ability of participants in the financial system to prevent, respond to, recover and learn from operational disruptions, such as cyber-attacks and internal process failures. Operational resilience is becoming more important to financial stability as AI and fintech play a greater role in the provision of financial services.

We can debate where exactly developments in financial technology and AI are taking us, but we can all agree that greater adoption of new technology leaves us all open to more risks. First, as Sasha Mills notes, some technologies may heighten threats from malicious actors – such as AI or quantum computing being leveraged to make cyber-attacks more powerful.

Second, a greater reliance on common technologies could cause multiple firms or financial market infrastructures (FMIs) to respond in the same way during an incident, and such correlation or herding behaviour could amplify the impacts.

Third, concentration risk arises when there is reliance on a small number of providers of a given service, which means that an incident in one provider could have a disproportionate impact on the system.

For me personally, the correlation and herding point is crucial here. A key lesson for regulators and policymakers is the importance of ensuring models don't all operate in the same way. To do so creates classic potential for the unintended consequences of regulation unwittingly to induce greater correlation and herding.

Hence, it is important in a 'constitutional' approach that provides guardrails that regulators continue to allow for competition and alternatives to avoid an unintended consequence of generating greater correlation and herding that could challenge financial stability.

In March, the FPC published our macroprudential approach to operational resilience, reflecting its increasing importance in our agenda. In this *Financial Stability in Focus* publication, we were clear in this that our approach is forward looking, recognising up front the inevitability of change in service provision and business models.

The need for ongoing dialogue

It is still relatively early days when it comes to considering all these issues. But what I am clear about is how the rise of new technologies means a thoughtful approach from the FPC and FMIC – we should remain alert, but also in listening mode, monitoring developments and keen to understand better, in line with some of the key principles set out in the Bletchley Declaration from the AI Safety Summit last year[19].

Specifically, for me: that AI has the potential to transform and enhance human wellbeing; that it should be designed, developed, deployed, and used, in a manner that is safe, in such a way as to be human-centric, trustworthy and responsible; and that all actors have a role to play in ensuring the safety of AI, including nations, international bodies, and academia.

We should find positive ways to discuss these changes together. History has shown that innovation triggers calls for regulation, which in turn triggers a negative reaction by those affected. There's nothing to suggest that AI will be any different. But we can be prepared to have that inevitable debate in a more thoughtful and informed way[20].

The key lesson for me is that building relationships, facilitating dialogue, and being open with each other is key. I note that following the Bank and FCA's AI Public-Private Forum (AIPPF) they are now considering establishing a follow-up industry consortium.

Conclusion

So where does that leave us? Productivity growth is crucial to boosting real wage growth and sustaining economic growth, particularly when the number of hours worked in an economy may be declining as populations are ageing and growing more slowly (or are, in some countries, declining). Innovation is a fundamental driver of productivity growth, which is why it is valuable to have promotion of innovation incorporated into the FMIC's objectives.

AI may be the answer to some of these challenges – but it could involve fundamentally disruptive innovation and change that brings both enormous upsides and potential risks. The challenge is therefore to develop a regulatory framework that fosters the flowering of creativity and innovation but takes into account the potential financial stability risks.

I believe an analogy of the 'invisible hand' of the LLM as being similar to the traditional human 'invisible hand' of the discovery process provides a useful lens through which to consider these issues and encourages us not to dismiss innovation out of hand because we can't fully understand and explain how it was generated.

Alongside that, I want to make sure FPC and FMIC as regulators and guardians of financial stability are properly equipped to deal with the challenges ahead – that means continuously and consciously deepening our understanding of the issues so we can take part fully in conversations about whether and how we should respond to developments.

In the meantime, there is plenty for us to do to continue to facilitate innovation and growth where we can while making sure, as far as possible, we have guardrails in place perhaps through a 'constitutional' approach to ensure that innovation takes place in a way that is conducive to financial stability. Achieving both of these objectives together won't be easy, which is why ongoing dialogue with stakeholders will be key. ■

Randall S Kroszner is an external member of the Financial Policy Committee (FPC) and of the Financial Market Infrastructure Committee (FMIC) at the Bank of England

Endnotes

1. Former MPC member Silvana Tenreyro also showed it is associated with better healthcare and wellbeing indicators as well.

2. Organisation for Economic Cooperation and Development (OECD)

3. US Bureau of Labor Statistics, May 2024 and Office for National Statistics, May 2024.

4. Yann Coatanlem (February 2024), 'Why Europe is a laggard in tech'.

5. Chad Syverson (2023), 'Structural Shifts in the Global Economy: Structural Constraints on Growth': Remarks at the 2023 Jackson Hole Symposium.

6. AWS Whitepaper, Interpretability versus explainability.

7. See Bank of England, The Bank of England's supervision of financial market infrastructures Annual Report 2023.

8. Briggs and Kodnani, (2023), 'The Potentially Large Effects of Artificial Intelligence on Economic Growth'.

9. McKinsey, (2023), 'The economic potential of generative AI: The next productivity frontier'.

10. Nicholas Crafts (2003), 'Steam as a General Purpose Technology: A Growth Accounting Perspective'.

11. Brynjolfsson, Rock and Syverson, (2021), 'The Productivity J-Curve: How Intangibles Complement General Purpose Technologies'.

12. See, for example, Zhao et al (2023). 'Explainability for Large Language Models: A Survey'.

13. Stuart Russell (2023), Stuart Russell Testifies on AI Regulation as US Senate Hearing.

14. Michael Polanyi, The Tacit Dimension.

15. FA Hayek, Competition as a Discovery Procedure.

16. See also Donal Lavoie, National Economic Planning: What is Left?

17. See also Manning, Zhu and Horton (2024), 'Automated Social Science: Language Models as Scientist and Subjects'.

18. See, for example, Yoshua Bengio (2023), 'AI Scientists: Safe and Useful AI?'. Of course, 'bad actors' can explicitly try to use AI as well as other tools in ways that are not aligned with legal and social goals.

19. See The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023.

20. Benedict Evans has a useful take on the back-and-forth between market participants and regulators and suggest there are generally three reasons people, or tech companies, generally say 'no' to new regulation. The first, which he describes as the default, is they just don't like it. Even though the change is possible, it may be awkward, inconvenient or expensive. So they push against it. The second reason is that the proposed change will have drastic unintended consequences which the regulators do not realise. The third reason he lists for saying no is that a proposal from a regulator may simply be technically impossible, even if it is desirable. (Benedict Evans, 2023, 'When tech says 'no''.

I am grateful to Maighread McCloskey for her assistance in preparing these remarks. I'd also like to thank Rachel Adeney, Anthony Avis, Andrew Bailey, Sandra Batten, Sarah Breeden, Lai Wah Co, Alex Gee, Bernat Gual-Ricart, Jonathan Hall, Jonathan Haskel, Adrian Hitchins, Owen Lock, Harsh Mehta and Michael Yoganayagam for their helpful comments and contributions. The views expressed here are not necessarily those of the Financial Policy Committee (FPC) or the Financial Markets Infrastructure Committee. This article is based on a speech given at City Week 2024.

www.euroeximbank.com

EURO EXIM BANK

Facilitating Global Trade

www.euroeximbank.com

HELPING YOU STAY AHEAD OF THE GAME

Letter of Credit | Standby Letters of Credit |
Bank Guarantees | Proof of Funds

We aim to accelerate your import & export business transactions and connect you to new opportunities

Representative Office (UK), Euro Exim Bank Ltd, Suite 1C, 1 Oaks Court, 1 Warwick Rd, Borehamwood, Hertfordshire, WD6 1GS, United Kingdom.

Registered Office (WI) Euro Exim Bank Ltd, Office 2, 1st Floor, La Place Creole Building, Rodney Bay, Gros Islet, St. Lucia, LC01 101, West Indies

info@euroeximbank.com
+44 208 207 2868 (UK)
+1 758 450 8349 (WI)

Maintaining the UK's leading global position in FinTech

The UK's fintech sector remains resilient. Roberto Napolitano discusses how this leadership in financial innovation can be continued

In 2023, despite economic challenges, the UK's FinTech sector remained resilient, securing over $5 billion in investment, second only to the US and more than all European countries combined[1]. This solidifies the UK's position as the second-highest recipient of global capital in FinTech, underscoring its leadership in financial innovation. Looking ahead, as technology advances, maintaining this leadership is crucial and collaboration among industry players, regulators, and government is key.

There are many initiatives Innovate Finance spearheads in supporting the growth of the FinTech sector, whether it is attracting capital, ensuring smarter regulation, or driving further diversity of talent.

An example is the recent launch of the Unicorn Council for UK FinTech (UCFT)™, a coalition of 'unicorns' aimed at accelerating growth in the sector. Co-chaired by Innovate Finance and industry leaders like Zilch, Revolut, and Clearbank, the UCFT wants to provide key policy recommendations to the government on UK regulatory environment, capital markets, R&D, investment schemes, VAT and other matters that are crucial to maintain the global leadership of UK FinTech.

Looking at the months ahead, we have worked with our members to produce a *General Election Fintech Manifesto* that outlines the strategy we need to implement to ensure the UK remains the best place for FinTech businesses to scale and prosper. The three areas of focus of the Manifesto are:

1. To become the world's first smart data economy by leveraging technologies like open data, AI, and blockchain, and pioneering smart data initiatives. Recent advancements in legislation, such as The Data Protection and Digital Information Bill, support this evolution.

2. To secure the UK's digital finance sector by combating payment fraud, which accounts for 40% of all UK crime and whose majority originate from social media platforms. Smarter regulations and collaboration among stakeholders are vital to ensuring consumer and business safety in the digital finance sector.

3. To embrace new technology and regulations: as AI continues to drive innovation, updating regulatory frameworks is necessary. Investments in growth capital and enhancing IPO markets are essential for ongoing innovation.

Innovate Finance continues to play a pivotal role in shaping the UK's FinTech landscape, ensuring it remains dynamic and inclusive, benefiting both consumers and businesses

Moreover, to strengthen the UK's FinTech ecosystem, Innovate Finance spearheads various initiatives that aim to drive further diversity and inclusivity of talent, essential for the UK FinTech sector to continue to prosper.

One example is our annual *Women in FinTech Powerlist* shining a spotlight on +250 women and 45 standout leaders across 8 categories that are making significant and impactful contributions to the FinTech ecosystem. In addition to this and given the success of last year, we had the privilege to launch our second *Pride in FinTech*™ to champion the incredible LGBTQIA+ community making an impact in UK FinTech with the support of leading organisations such as Google, Zopa Bank and Alloy.

This year, we are also collaborating with Zopa Bank on the first *Pride in FinTech Barometer 2024* that wants to evaluate the current state of affairs for the LGBTQIA+ community working in FinTech and provide actionable recommendations to FinTech companies to assess what more can be done to make a tangible difference.

In summary, sustained collaboration and innovation are vital for maintaining the UK's leadership in FinTech. Through initiatives like the UCFT, the *General Election Fintech Manifesto*, our leading initiatives in diversity and inclusion, we want to play a pivotal role in shaping the UK's FinTech landscape, ensuring it remains dynamic and inclusive, and benefiting both consumers and businesses with more transparent and more democratic financial services for all. ■

Roberto Napolitano is CMO at Innovate Finance

Endnote
1. Please see Innovate Finance's 'FinTech Investment Landscape 2023'.

Innovate Finance is the independent industry body for UK FinTech. Its mission is to accelerate the UK's leading role in the financial services sector by directly supporting the next generation of technology-led innovators to create a more inclusive, more democratic and more effective financial services sector that works better for everyone. Innovate Finance's membership and partnership community ranges from seed stage startups to scale up and high growth FinTechs; from multinational financial institutions to big tech firms; and from investors to global FinTech hubs. Innovate Finance supports our members and the wider financial innovation ecosystem by promoting policy and regulation that allows innovation to thrive, encouraging talent, diversity and skills into the sector, facilitating the scaling journey, fostering business opportunity, partnerships and domestic and international growth, and driving capital into UK FinTech. By bringing together and connecting the most forward-thinking participants in financial services, Innovate Finance is helping create a financial services sector that is more transparent, more sustainable and more inclusive.

More information at www.innovatefinance.com.

Innovation and the evolving financial landscape

Financial technology is evolving. Michelle Bowman discusses how technology can enhance financial services in a manner that is consistent with operating in the highly regulated banking industry

I will share my views on evolving financial technology and the importance of ongoing, responsible innovation, including the roles of both the public and private sectors in shaping the future. In my role at the Federal Reserve, I view financial innovation through the lens of the central bank's responsibilities—issuing US currency; conducting monetary policy; fostering a safe and efficient payment system; maintaining financial stability; promoting consumer protection; and supervising and regulating financial institutions.

Regulators often espouse the benefits of innovation—innovation can lead to greater efficiency, and it can promote competition in the market, which can lower the cost and expand the availability of products and services to consumers and businesses.

While regulators acknowledge these benefits, innovation is inevitably accompanied by risk. Innovation can exacerbate traditional financial risks or introduce new risks that must be carefully understood and managed.

And on a basic level, the 'newness' of innovation often leads regulators to reflexively resist these changes, whether the change takes the form of new technology, new ways of delivering products and services, new financial infrastructure that underlies how the banking system works, and new relationships within the financial system and beyond.

So how do regulators get comfortable with innovation? Some attending today may answer that 'they don't!' And I am sure that some have experienced this friction in trying to pursue innovation in a heavily regulated environment.

In fairness, sometimes 'no' is the correct regulatory response when innovation either does not solve an actual problem, or simply cannot be executed in a safe and sound manner and in compliance with applicable laws. But my

goal is to propose some building blocks that could help regulators get to 'yes' more often, and potentially smooth the rough path to successful innovation in the banking system.

Understanding innovation

As a first principle, I would offer 'understanding' as a necessary predicate to promoting innovation. Before we craft a useful public policy around innovation in banking, we need to understand the various dynamics involved with particular innovations.

My hope is that as we enhance our understanding, and we recognize the promise of new technology, we can achieve a banking system that welcomes innovation, and is stronger and more efficient as a result

We must consider how innovation can impact different financial sectors, from small banks to wholesale financial markets, and how those impacts will ultimately effect end users. For example, will 'tokenized' products and platforms duplicate existing bank deposits and payment rails, potentially creating parallel systems? If so, will the products and platforms that duplicate these functions provide the same legal protections for customers and the overall financial system?

The wide variety of technology and use cases can be a significant obstacle to getting to a place where regulators understand any one innovation. Innovation in the financial system can take many forms including new technology, new business models, and improvements to existing infrastructures.

As one form of innovation, distributed ledger technology (DLT), including blockchain, has inspired new ways of thinking about the exchange of assets and data, in addition to the role of intermediaries and trusted third parties. DLT combines a number of different design elements—like distributed data storage, cryptography, and consensus mechanisms—that support the transfer process, information visibility, and transaction recordkeeping[1].

DLT is an interesting example because it highlights the challenges presented to regulators in understanding innovation. There are many variations in the application of DLT—and ongoing research and development creating more variability over time—which ultimately may complicate understanding how the technology can be used, and how different functionality can be incorporated into DLT solutions for different use cases.

In short, innovation can take many forms. It often involves new uses for existing technology that enhance core business lines, but it can also include the introduction of transformational new technology or capabilities. Customer demands for cheaper products or for new and innovative products often motivate banks to explore how technology and innovation can meet these demands and provide products and services for the future.

While understanding technology is an important first step, it is not the only one. Regulators also need to understand the players who operate in the fintech space. Banks of all sizes have a long history of innovation, from the introduction of checks; to the advent of ATMs that allow easy access to cash and basic banking information; to online banking platforms and electronic payments that have simplified and streamlined the way many consumers and businesses manage and access their funds, receive credit, and conduct financial transactions.

Innovation helps banks upgrade their existing business lines to better meet customer needs, for example, by introducing instant payments or providing credit in innovative ways.

There are also many nonbank providers in this space, including core service providers that often hold the 'keys' to innovation at small banks when they facilitate the add-on of new technology to core systems. And fintechs that may focus on even more transformational uses of technology within all aspects of banking. This presents another layer of complication when it comes to innovation in banking.

Even with an understanding of technology and of the players involved, regulators still need additional feedback to understand the 'why' of innovation. What is the purpose of the innovation? What problem is it designed to solve? Are there tradeoffs policymakers would need to consider if the new technology were introduced or integrated in the existing financial system?

There are many examples across this spectrum of innovation. For example, take the use cases around DLT. Some banks are exploring the possibility of DLT as a way to help facilitate services like crossborder payments and financial market transactions.

Other businesses may see DLT as a way to improve supply chain management by tying payments to specific activities across the supply chain. Some financial institutions see DLT as a potential solution for existing processes

that require manual interventions and coordination across disparate systems. Businesses may look at a single technology—DLT—and see possible solutions to a wide range of problems.

When the financial system promotes innovation, we better enable firms to serve their customers' unique needs while also advancing the capability of the financial system.

Regulators and industry both have an important role to play in achieving this goal of understanding. Regulators cannot hope to craft effective public policy without understanding. And while industry's focus is rightly on developing innovative solutions, part of successful innovation in the banking space is promoting education and understanding for the regulators.

Openness to innovation: getting to 'yes'

As a second principle, regulatory openness is a critical ingredient to fostering innovation in the financial system. The first reaction of regulators to proposed innovation in the banking system is often not one of openness and acceptance, but rather suspicion and concern.

The use of emerging technology and innovation may require a change in policy or supervisory approach. It also very often requires regulatory feedback—sometimes before innovation is introduced, in others after it has been rolled out and is reviewed during the supervisory process.

As a financial policymaker and a regulator, I recognize that there are a number of questions we must answer before pursuing a change in policy, whether in the form of regulation, or in supervisory approach, to facilitate innovation in the financial system.

Regulators need to ask whether we have considered the intended and unintended outcomes of a new innovation, and do the benefits of the new technology outweigh the risks? Would the introduction of a new technology or innovation in the banking system require updating our regulatory framework to incorporate clear oversight? Who should have responsibility for oversight, and what roles do we see for different regulators at both the state and federal level?

Transformational technology requires clear, consistent, and transparent guardrails and expectations to govern the activities that are allowed into the regulated financial system. Where current regulation does not contemplate a new activity, should it be acceptable for financial agencies to regulate the activity through supervision and enforcement alone? Or should congressional action address the treatment of these activities in the financial system?

These are difficult questions, to be sure, but ones we must confront if we are to allow innovation to flourish in the banking system and the broader financial system. Ensuring an orderly and observable method for regulators to understand and provide a path for potentially disruptive or transformational technology could ultimately enhance the long-term stability of the financial system.

So, while the obligation to *promote understanding* may fall more heavily on industry, the obligation to *be receptive to innovation* falls more heavily on regulators. We must fight the temptation to say 'no' and resist new technology, and instead focus on solutions—how can we mitigate the risk of new technology? What benefits will technology bring to the financial system? How can we provide clear regulatory expectations?

The appeal of resisting innovation—resisting change—is that it preserves the familiar and known. But the cost of resistance to change may be great, in that it promotes stagnation and inertia. Ultimately, this could lead to a

banking system that may be safer and smaller, but also less effective at providing banking products and services and in supporting the US economy.

Innovation as a priority in banking

Understanding and openness can go a long way to promoting innovation, but I think there are opportunities to do more. Can we shift from a reactive approach to innovation, to an active one that facilitates innovation? This leads to my final principle, which is that regulators can do more to promote innovation. Regulators can do more than simply tolerate innovation, they can promote it through transparency and open communication.

Take, for example, the frictions associated with crossborder payments, including the speed and cost and the ability to transfer money. At first blush, this seems like a prime opportunity for innovation to come in with a new approach.

The challenge in crossborder payments historically has been achieving the goal of serving new types of customers and increasing the speed of payments, without losing or watering down important compliance safeguards that deter criminal activity. Some perceived payment frictions exist for specific policy reasons and do not stem from issues with existing technology. Therefore, new technology alone cannot solve the issue unless it also addresses the required safeguards.

Regulators can serve a valuable role in identifying where a technology solution may have an important 'gap', as in the case of crossborder payments, by identifying how a solution can meet the needs of both customers and regulators. Innovation and regulatory and legal requirements can coexist—providing both enhanced capability and regulatory compliance. Transparency can promote innovation.

Policymakers should strive to define a clear and sensible regulatory framework designed to meet policy objectives that also enables the private sector to innovate within the established guardrails. A clear regulatory framework

supports private sector innovators by providing clarity and consistency that encourages long term business investment in pursing innovation, while continuing to support today's products and services. A regulatory framework empowers supervisors to focus on safety and soundness and ensuring a safe and efficient payment system.

In the past, I have described an approach to innovation that solves specific problems and enables banks and other providers to meet the needs of their customers in a safe and sound manner[2]. This approach recognizes the role of the private sector and focuses policymakers on payment and financial system infrastructure while balancing the public policy objectives with the benefits provided.

In this construct, the same activities that present the same risks must be subject to the same regulatory expectations—regardless of what the product is called or where it is offered.

The active promotion of innovation has proven to be successful in the past. One example from 2019 highlights the benefits of this approach, specifically as it relates to the use of alternative data to support small dollar consumer loans[3].

Timely guidance and regulatory clarity in the interagency statement clarified that, with a customer's consent, a bank may use alternative data, like checking account balance activity to help evaluate the potential borrower's creditworthiness who might not have otherwise qualified for a loan.

This example did not require a change in technology, instead it used data in a responsible and innovative way through leveraging deposit account cashflow information to show credit worthiness, enabling banks to meet their customers' credit needs.

I think there is more we could do on this front. For example, I think there are opportunities for regulators to provide clear, actionable, and timely feedback on innovation proposals. And the principles I have described are complementary—the more regulators understand innovation, the more comfortable they will be in accepting it and promoting its adoption in the financial system.

My hope is that the Federal Reserve's ongoing research and engagement with industry will help us continue to understand and assess the potential benefits and risks presented by new innovation.

Closing thoughts

As I consider the ever-evolving financial landscape, I will continue think about how the regulatory framework can accommodate new technology and services, while applying the safeguards in place today that protect households and businesses, and guard against illicit activity, ensuring the safety and soundness and the stability of the financial system.

My hope is that as we enhance our understanding, and we recognize the promise of new technology, we can achieve a banking system that welcomes innovation, and is stronger and more efficient as a result. ■

Michelle W Bowman is a Member of the Board of Governors of the Federal Reserve System

Endnotes

1. See eg. Mills, David, Kathy Wang, Brendan Malone, Anjana Ravi, Jeff Marquardt, Clinton Chen, Anton Badev, Timothy Brezinski, Linda Fahy, Kimberley Liao, Vanessa Kargenian, Max Ellithorpe, Wendy Ng, and Maria Baird (2016). "Distributed ledger technology in payments, clearing, and settlement," Finance and Economics Discussion Series 2016-095. Washington: Board of Governors of the Federal Reserve System.

2. Michelle W Bowman, "Responsible Innovation in Money and Payments (PDF)," (Speech at the Roundtable on Central Bank Digital Currency, hosted by the Harvard Law School Program on International Financial Systems, October 17, 2023).

3. Board of Governors of the Federal Reserve System, Consumer Financial Protection Bureau, Federal Deposit Insurance Corporation, National Credit Union Administration, Office of the Comptroller of the Currency, "Interagency Statement on the Use of Alternative Data in Credit Underwriting," CA letter 19-11, December 12.

Thank you to Priyanka Slattery and Alex Sproveri of the Federal Reserve Board for their assistance in preparing this text.
The views expressed here are my own and not necessarily those of my colleagues on the Federal Reserve Board or the Federal Open Market Committee. This article is based on a speech delivered at The Digital Chamber DC Blockchain Summit 2024, Washington, DC, May 15, 2024.

Sharper supervision in an era of technology races

Cecilia Skingsley proposes a framework to help central bankers in embracing innovation and at the same time solve some of the challenges confronting the central banking and regulatory communities

People often say that time stops for no one. I like to say the same about technology: it won't slow down to wait for anyone. It is hard to avoid the impression that the world around us is getting ever more complex: the growth of new activities, for example, or how we explore and think about the role of AI. What to bring into the realm of regulation and when is also changing.

I would like to propose a framework to help guide us in embracing innovation and at the same time solve some of the challenges confronting the central banking and regulatory communities. I think this framework will be particularly useful in the interaction between the public and private sector.

A core question is: how, over the next decade, do we wish to respond to what is likely to be more and faster change, bringing with it significant opportunities and potentially some profound challenges?

I will argue that this framework must inform the actions we take if we want financial stability and safety to be integral to progress.

But before I go ahead, let me present you with some information about the BIS Innovation Hub. A joint venture between the central bank-owned Bank for International Settlements and a number of individual central banks, the Innovation Hub has been in operation now for nearly five years. As a laboratory aimed at developing public goods, we know a lot about what it takes to experiment and bring value in a public sector context. We also have a clear understanding of where the challenges are.

We develop projects to create proofs-of-concept prototypes of new platforms and technologies. In these five years, we have expanded from a handful of people to an almost 100-strong staff. We'll soon have seven centres around the world. We have started more than 30 projects.

Our work spans six focus areas: central bank digital currencies, next generation financial market infrastructures, green finance, open finance and the one I want to focus on – supervisory technology, or Suptech. It is about using technology to support global financial stability.

By the way, our work also includes regulatory technology (Regtech) and a new area that we call monetary policy tech, which focuses on helping central banks use technology for the core tasks of research and monetary policy.

So now that you know who we are, let's go back to how we should approach the challenges ahead.

A core question is: how, over the next decade, do we wish to respond to what is likely to be more and faster change, bringing with it significant opportunities and potentially some profound challenges?

Transparency and honesty

The first key point in my framework today is transparent and honest engagement with a broad range of stakeholders. Central banks and supervisors face challenges like modernising legacy systems and technology infrastructure and designing tools using new technologies to enable us to be fit for the future.

With these challenges, we have some hard questions. For those of us on the public sector side, are we willing to be transparent to a satisfactory level, knowing that our work is complex, challenging and sometimes very sensitive? As digital transformation accelerates, can we honestly assess and communicate where we stand and where we should be going?

This self-appraisal can be difficult for organisations at the best of times, and central banks and other authorities on the financial side of our societies are no exception.

Cooperation

Which brings me to the other key point in this framework: collaboration. We, as in the public sector, need to join forces with those in the private sector to build technology solutions that can shape the future of supervision and support financial stability.

The public sector may not always be the easiest of customers. Often having more vague objectives than you typically find in companies, we are also complex organisms with a multitude of specialisations and different priorities.

The public sector needs specific processes to ensure fair and systematic decision-making and to be transparent about how public money is spent. And so we are not the most profitable types of customers.

Suptech is also a small market. When the BIS Innovation Hub London Centre conducted a survey in this field last year, about half of the vendors we asked pointed this out. I will come back to this survey a few more times. We expect to publish its full results later this year.

At the Innovation Hub, we think SupTech is critical in supporting effective and efficient supervision, from achieving real-time risk alerts to supporting automated regulatory compliance.

We know our colleagues across the central banking and supervisory worlds think so too. Recent research we undertook with the Financial Stability Institute, also a part of the BIS, across 50 jurisdictions highlighted that over 90% of them have already deployed SupTech tools – most commonly for regulatory reporting, assessing risk and automating supervisory processes.

Examples include dashboards to monitor capital, liquidity, credit and market risk; natural language processing to monitor and analyse social media, news, financial statements, etc; and tools to automate supervisory tasks such as licensing.

And 80% of the institutions surveyed have a dedicated internal resource to build these tools, with many more solutions in development, including in emerging areas such as environmental, social and governance reporting and crypto asset monitoring. So we know that both the appetite and the technical capability are there. That's the good news.

However, the majority of these tools are at an early stage of maturity, focusing largely on digitalisation and automation of existing workflows. We also know that they are not being widely adopted inside institutions, as many solutions built in-house remain at proof of concept or prototype stage, thus limiting the potential value they could

bring. There is a lot of experimentation but little deployment. Clearly, the job is not yet done, and there is still a long way to go to achieve a technology-first approach.

The Innovation Hub is trying to accelerate this important change in two ways: through experimentation and collaboration. We are building proofs of concept and prototypes aimed at monitoring stablecoins; working to better understand the economic significance of DeFi and cryptoassets; monitoring electronic markets; building a regulatory reporting and data analytics platform; using large language models to enable climate risk analysis; and encoding policy and regulatory requirements into cross-border payment protocols.

Our portfolio of projects is multi-pronged, experimenting at the cutting edge of supervisory activity while also focusing on some of these entrenched challenges to supervisory effectiveness, addressing both present and future challenges.

A second defining feature of the Innovation Hub's work is how we use cooperation to support innovation. In our projects we bring in central banks as partners and make sure that we procure vendors that can deliver what we do not have ourselves.

Which brings us to what I would call the 'supply side' of SupTech. We do work with the private sector in our projects. But the current dynamic of the supply and demand sides of the SupTech market is not without its challenges.

Coming back to the London Centre survey, we find that breaking down the barriers and allowing further collaboration between financial authorities and SupTech solution providers is not easy.

Half of the SupTech vendors told us that *"lack of visibility into the prioritized needs of financial authorities"* is a key challenge for expanding their SupTech portfolio. This makes it hard to match our problems with their solutions. So financial authorities need to collectively think of ways to bridge this gap.

At the Innovation Hub we believe that one way of achieving this is to have more showcasing events. TechSprints and hackathons are invaluable, and we know that the private sector also values them as a way to engage with us[1]. By the way, we will be holding a SupTech TechSprint later this year. We need to continue to think creatively about how to better understand each other.

We also discovered that over two thirds of vendors complain that *"length or complexity of procurement processes"* affects their ability to successfully engage with financial authorities.

One critical aspect is that this is more likely to hurt smaller firms, as it increases uncertainty and costs. That could mean that only the large and established ones are able to engage, at the expense of the smaller start-up community where true innovation and cutting-edge thinking often happens first.

As I mentioned earlier, public institutions need these procurement processes. We are mandated by law to be transparent and fair when purchasing from the private sector, and we need to be able to prove it. But perhaps there is scope to improve our processes to make them truly inclusive for firms of all sizes.

In some ways, it is up to us in the public sector to be better customers. We need to better articulate our needs to the market and find ways to be more inclusive. In this way, we benefit from all the choices available in the SupTech marketplace and continue to be up to date with the latest technologies. We might be missing a lot. We don't know what we don't know.

Conclusion

Let me conclude. The BIS Innovation Hub aims to create public goods that promote global financial stability. As those of you who work in innovation know, it is hard work. It is risky and in many ways countercultural to how central banks traditionally work.

But in our view, Suptech innovation continues to be essential to safeguarding our financial system. All of us – central banks, supervisors and regulators, the private sector – need to focus on this important task. To best achieve it, we need to operate within a framework that combines recognising the needs of multiple stakeholders with working more effectively with the private sector.

We at the Innovation Hub are genuinely excited to play our part, and we invite you to join us in creating the financial infrastructure of the future. ■

Cecilia Skingsley is the Head of the BIS Innovation Hub

Endnote
1. SupTech vendors identified 'showcasing solutions' (28%) and hackathons (16%) as two key areas by which they engage with financial authorities.

This article is based on a keynote speech delivered at the Innovate Finance Global Summit, London, 15 April 2024.

Building confidence in the path ahead

Christine Lagarde outlines what needs to be done to become sufficiently confident to start dialling back the ECB's restrictive policy stance

Since the pandemic, monetary policymakers have been facing an exceptionally complex environment. As inflation rose, we were confronted with profound uncertainty about how far it would go and how widely it would spread across the economy. And even as inflation has eased, uncertainty about its persistence has remained.

The potential costs of mis-calibrating policy have been high, which is why we had to employ a policy framework that minimises the risk of mistakes. And we have done so by building our reaction function around three criteria: the inflation outlook, the dynamics of underlying inflation and the strength of monetary transmission.

Though we conceived these criteria when we had low visibility of future inflation, they have also helped guide our decisions as inflation has fallen and forecasts have become more accurate.

As Marie Curie once said, to thrive through the ups and down of life, *"we must have perseverance and above all confidence."* And our framework has indeed encouraged us to persevere when necessary and to build up confidence when needed. It has served as a reliable compass for calibrating policy through three phases of our current policy cycle.

First, it helped create robustness during our tightening phase when we were devising how far we needed to go to rein in inflation.

Second, it has helped us practice patience during the holding phase until the signals from our inflation projections and underlying inflation are more consistent.

Third, it will support us in building up sufficient confidence to begin the dialling-back phase in which we make policy less restrictive.

The tightening phase

In the early phase of our tightening cycle, our main priority amid surging inflation rates was to exit our accommodative policy stance as quickly as possible. While the policy challenge was immense, the policy path was relatively simple to calibrate.

To calibrate policy accurately we needed a framework for policy decisions that would work when we had low visibility and would mitigate heightened uncertainty

But as rates rose and approached restrictive territory, calibrating our policy stance became more complex. We first had to assess how much rates needed to rise until they were sufficiently restrictive, and then for how long they needed to stay at that level. But our assessment was blurred by much lower-than-normal visibility of the future.

Our forecasts repeatedly underpredicted inflation by large margins, even at shorter horizons. From 2021 to 2022 for example, the absolute inflation forecast errors in the staff macroeconomic projections, one quarter ahead, more than doubled, largely owing to volatile energy prices[1].

At the same time, the mix of shocks that emerged from the pandemic and its aftermath – rotations in spending, energy spikes, 'bullwhip' cycles in manufacturing, supply bottlenecks, tight labour markets, fiscal expansion and reopening effects – heightened the risk of inflation becoming more persistent.

We faced a highly unusual conjuncture of high inflation and declining real wages, but also rising employment. This combination essentially implied a multi-year catch-up process to make up for real wage losses. In turn, this process could have triggered what I referred to at last year's conference as a 'tit-for-tat' inflation dynamic[2].

And we faced uncertainty as to how quickly and forcefully our monetary policy response would succeed in bringing down inflation. The ECB had not been through a tightening cycle for more than a decade, and there were reasons to believe that the transmission of monetary policy to firms and households might have changed[3].

So, to calibrate policy accurately, we needed a framework for policy decisions that would work when we had low visibility and would mitigate heightened uncertainty. This is why we built our monetary policy response around the three criteria I referred to earlier: the inflation outlook, the dynamics of underlying inflation and the strength of monetary transmission.

This approach made our decisions more robust, as the inflation path we foresaw in our projections had to be validated by data we could observe in real time and extrapolate into the medium term. That, in turn, enabled us to take forward-looking decisions with a higher degree of confidence. And it served us well in practice.

The three criteria helped us to map out the remaining climb, allowing us to bring rates to sufficiently restrictive levels to break the persistence of inflation[4]. But also, by guiding us to carefully evaluate the strength of policy transmission, they acted as a cross-check against overtightening. This helped us reach the decision to stop rate hikes after last September.

The holding phase

We then entered the current phase of our policy cycle – the holding phase – during which we committed to keep rates at restrictive levels for as long as necessary. Since the start of this phase, inflation has been declining consistently and our projections have been showing inflation returning to our target over the medium term.

We now project inflation to average 2.3% in 2024, which is 0.4 percentage points less than projected in December and 0.9 percentage points less than September. We then expect inflation to decline to 2.0% in 2025 and 1.9% in 2026.

And unlike in the earlier phases of our policy cycle, there are reasons to believe that the expected disinflationary path will continue. First, for some time now inflation outturns have been broadly in line with our expectations. In 2023 we saw a reduction of about 70% in the average absolute error in our staff projections relative to 2022, one quarter ahead.

Second, we now see inflation returning to 2% earlier in our projection horizon than before, in mid-2025, and not exceeding our target for the remainder of the horizon.

Third, the composition of inflation is improving, as we now expect lower core inflation in the medium term. This suggests that the convergence to 2% is likely to be more durable and less beholden to assumptions about commodity prices, although the latter can always prove hazardous.

The other criteria are also becoming more consistent with this improved inflation outlook. The transmission of our monetary policy is unfolding in the right direction. Financing conditions have reacted strongly to higher rates, loan demand has weakened and, in turn, activity has slowed notably in the most interest-sensitive sectors of the economy.

And underlying inflation is generally easing. Nearly all the measures that we track are declining, and the range of readings between the different measures has narrowed from 4.1 percentage points at its peak to 2.4 percentage points today. Some of the measures of underlying inflation with the best leading indicator properties for future inflation have dropped steeply[5].

But, at the same time, domestic price pressures remain strong. Services inflation is still stubborn and hovering around 4%, while momentum increased somewhat in February. And our indicator of domestic inflation, which measures items with a low import content, stands at 4.5%, at the top of the range of underlying inflation measures that we monitor. This measure has also been found to have good leading indicator properties[6].

These pressures largely reflect robust wage growth as the catch-up process continues, as well as a tight labour market that has so far been resilient to a slowing economy. Employment grew by two million cumulatively during

2023, even as the economy stagnated, while firms continue to hoard labour. This pattern is mechanically lowering labour productivity and pushing up unit labour costs.

At this stage, it is difficult to assess whether these price pressures simply reflect the lag in wages and services prices and the procyclical nature of productivity, or whether they signal persistent inflationary pressures.

So, although we have made significant progress in all three of our framework criteria, we are not yet sufficiently confident that we are on a sustainable path towards our inflation target.

Building sufficient confidence to dial back policy

So the essential question is: what do we need to see to become sufficiently confident to start dialling back our restrictive policy stance? Put simply, we need to move further along the disinflationary path. And there are three domestic factors that will be decisive to ensuring that the inflation path evolves as we project.

The first of these is wage growth. Our forecast sees nominal wages slowing to 3% over the next three years, allowing real wages to fully catch up to pre-pandemic levels over the projection horizon, also including productivity gains[7].

But with the unemployment rate expected to remain very low at 6.6%, this wage path cannot be taken for granted. Sensitivity analysis by ECB staff shows that if there were an earlier full catch-up by the end of this year, inflation would rise to 3% in 2025 and only fall to 2.5% in 2026[8].

The second is profit margins. The compression of profit margins has allowed wages to catch up without further accelerating inflation. Unit profits accounted for more than 50% of the GDP deflator in the last quarter of 2022 but this figure fell to just 20% a year later.

But our sensitivity analysis shows that, if firms were to regain pricing power as the economy recovers and profit margins were to rise by an accumulated 1 percentage point more than we project until the end of 2026, inflation would be 2.7% in 2025 and 2.4% in 2026.

The third factor is productivity growth. We expect that a pick-up in demand, if accommodated by fully utilising hoarded labour, will lead to rising productivity growth and falling unit labour costs. We project labour productivity growth of 0.1% this year before it rises to 1.2% in 2025 and 2026. But the path of inflation could be different if, in a new geopolitical environment, productivity losses for European firms turn out to be partly structural.

Given the delays with which these data become available, we cannot wait until we have all the relevant information. To do so could risk being too late in adjusting policy. But in the coming months, we expect to have two important pieces of evidence that could raise our confidence level sufficiently for a first policy move.

First, we will have more data to confirm whether wages are indeed growing in a way that is compatible with inflation reaching our target sustainably by mid-2025.

The latest data point in this direction. Growth in compensation per employee edged down to 4.6% in the fourth quarter of last year – slightly below our March projection – from 5.1% in the third quarter. Negotiated wage growth, which accounts for the lion's share of compensation per employee growth, also decreased from 4.7% to 4.5% in the fourth quarter.

Similarly, the ECB's forward-looking wage tracker, which anticipates the development of negotiated wage growth in the euro area, is showing early signs that pressure is easing. Average wage growth in 2024 for all existing wage contracts[9] fell from 4.4% at the time of our January Governing Council meeting to 4.2% at the time of our meeting in March.

The coming months will help us form an even clearer picture. We will receive data on negotiated wage growth in the first quarter of this year at the end of May. And many wage negotiations are currently taking place in large sectors, the outcome of which will be entered into our wage tracker as soon as the negotiations are concluded.

Employees whose contracts ran out last year and have not been renewed, or will run out by March 2024, account for around one-third of those in our wage tracker.

Second, by June we will have a new set of projections that will confirm whether the inflation path we foresaw in our March forecast remains valid. These projections will also implicitly give us more insight into the path of underlying inflation.

We will have more visibility on the strength of the recovery and the likely direction of the labour market, and therefore on the consequences for wages, profits and productivity.

In addition, we will have had a longer window to assess whether inflation data continue to fall broadly in line with our projections. If they do, we can be more confident that our models are now better accurately capturing inflation dynamics.

And this confirmation will be particularly important for the more persistent components, such as services, so that we can trust these components will continue to decline in keeping with their typical lagging pattern.

If these data reveal a sufficient degree of alignment between the path of underlying inflation and our projections, and assuming transmission remains strong, we will be able to move into the dialling back phase of our policy cycle and make policy less restrictive.

But thereafter, domestic price pressures will still be visible. We expect services inflation, for example, to remain elevated for most of this year. So, there will be a period ahead where we need to confirm on an ongoing basis that the incoming data supports our inflation outlook.

This has two important implications for the policy path ahead. First, our decisions will have to remain data dependent and meeting-by-meeting, responding to new information as it comes in. This implies that, even after the first rate cut, we cannot pre-commit to a particular rate path.

Second, our policy framework will remain important to process the incoming data and calibrate the appropriate policy stance. At the same time, the relative weights assigned to the three criteria will have to be regularly examined.

Conclusion

I said after our last Governing Council meeting that, when it comes to the data that is relevant for our policy decisions, we will know a bit more by April and a lot more by June. I hope that my remarks today help you to better understand our analysis and logic.

In the coming months, we will receive more data, which will help us to assess whether we are sufficiently confident in the path ahead to move to the next phase of our policy cycle. ■

Christine Lagarde is President of the European Central Bank

Endnotes

1. Chahad, M, Hofmann-Drahonsky, A-C, Meunier, B, Page, A and Tirpák, M (2022), "What explains recent errors in the inflation projections of Eurosystem and ECB staff?", Economic Bulletin, Issue 3, ECB; Chahad, M, Hofmann-Drahonsky, A-C, Page, A and Tirpák, M (2023), "An updated assessment of short-term inflation projections by Eurosystem and ECB staff", Economic Bulletin, Issue 1, ECB; Chahad, M, Hofmann-Drahonsky, A-C, Martínez Hernández, C and Page, A(2024), "An update on the accuracy of recent Eurosystem/ECB staff projections for short-term inflation", Economic Bulletin, Issue 2, ECB, forthcoming (21 March).

2. Lagarde, C (2023), "The path ahead", speech at "The ECB and Its Watchers XXIII" conference, Frankfurt am Main, 22 March.

3. ibid.

4. Lagarde, C (2023), "Breaking the persistence of inflation", speech at the ECB Forum on Central Banking 2023 on "Macroeconomic stabilisation in a volatile inflation environment" in Sintra, Portugal, 27 June.

5. Bańbura, M, Bobeica, E, Bodnár, K, Fagandini, B, Healy, P and Paredes, J (2023), "Underlying inflation measures: an analytical guide for the euro area", Economic Bulletin, Issue 5, ECB.

6. ibid.

7. This exercise is based on compensation per employee deflated by HICP and productivity growth using a four-quarter moving average. Conducting the same exercise without considering productivity, the catch-up would take place in mid-2025.

8. The analysis is based on Arce, O, Ciccarelli, M, Kornprobst, A and Montes-Galdón, C (2024), "What caused the euro area post-pandemic inflation? An application of Bernanke and Blanchard (2023)", Occasional Paper Series, No 343, ECB.

9. Including one-offs.

This article is based on a speech delivered at The ECB and its Watchers XXIV Conference, organised by the Institute for Monetary and Financial Stability, Goethe University, Frankfurt am Main, 20 March 2024.

How new EU fiscal rules can succeed

The EU has enacted new rules that overhaul the Stability and Growth Pact. Lucio Pench considers the issues that need to be addressed to ensure the new fiscal rules succeed

Executive summary

The debate on the reform of the European Union's fiscal rules, the Stability and Growth Pact, has largely focused of their design. This nearly exclusive focus has distracted attention from the equally, if not more, important issue of implementation. The reform, completed in April 2024, left implementation unaddressed, or at least open to very different potential outcomes.

In particular, the reform failed to clarify the interplay between EU countries' medium-term fiscal structural plans (MTFSPs), which embody the new focus on debt sustainability, and the excessive deficit procedure (EDP), which remains the main enforcement tool under the rules. The need for clarification is urgent as several countries are set to enter EDPs for breaching the SGP's 3 percent of GDP deficit threshold at the same time as their first MTFSPs are endorsed in autumn 2024.

There is a risk that the adjustment paths prescribed by EDPs may be at least temporarily less demanding than the debt-sustainability requirements of the MTFSPs would normally imply. Even if consistency between EDPs and MTFSPs is ensured from the start, inconsistencies may arise over time and be resolved in a way that further postpones the necessary adjustment.

The main risk is that the 3 percent of GDP deficit might be perceived as the only target that matters for countries that enter EDPs in 2024, as repeated revisions of the MTFSPs undermine the cogency of the debt sustainability requirements. This scenario is likely to materialise if the countries are allowed to exit their EDPs upon bringing their deficits to or below 3 percent of GDP, while being still far from the necessary correction of the debt trajectory.

It is important to shape countries' expectations on the implementation of the upcoming EDPs in a way that is conducive to the immediate internalisation of the debt sustainability constraint implied by the new rules, rather than allowing it to be viewed as a distant objective.

This change in expectations could be achieved by clarifying that, even if a country has been placed in an EDP only for breach of the deficit criterion, it should also satisfy the debt criterion for the procedure to be abrogated.

There is a risk that implementation of the forthcoming deficit-based EDP could lead to an unravelling of the entire reform

1 Introduction

The rules meant to constrain government deficits and debt in the European Union, known as Stability and Growth Pact (SGP), consists of two 'arms'. Under the preventive arm, all countries are expected to stick to the same medium-term objective of keeping their budgetary positions close-to-balance or in surplus. Under the corrective arm or excessive deficit procedure (EDP), meanwhile, countries with deficits in excess of 3 percent of GDP (deficit criterion) or debts in excess of 60 percent of GDP that are not falling fast enough (debt criterion) are subject to specific adjustment requirements to remedy the situation[1].

Compliance with the preventive arm is backed by soft-law recommendations. The corrective arm is more intrusive and, for euro area countries, potentially backed by financial sanctions.

In April 2024, following protracted negotiations between governments and with the European Parliament, the EU enacted new rules that overhaul the SGP's preventive arm (Box 1)[2]. The one-size-fits-all balanced budget target is dispensed with, and more focus is put on debt sustainability in each country. Countries must pursue debt sustainability through EU-endorsed, so-called medium-term fiscal-structural plans (MTFSP), which set out their intended fiscal adjustment paths.

For high-debt countries, deviating from the adjustment path can trigger the opening of an EDP with ensuing adjustment requirements. In other words, for these countries, the (debt-based) EDP is repurposed as an enforcement instrument of the new preventive arm. Countries with debt below the 60 percent threshold and no plans to exceed it are essentially left alone by the new rules, unless they breach the 3 percent of GDP deficit threshold.

The reform has left the EDP for breach of the deficit criterion (deficit-based EDP) practically unchanged. This reflects its popularity. Fiscally conservative governments see it the only element of the SGP that reliably constrains the fiscal

Box 1. The new preventive arm of the SGP (Regulation (EU) 2024/1263)

The principal aim of the new preventive arm of the SGP is to 'de-risk' public debt.

Each country should submit a medium-term fiscal-structural plan (MTFSP) setting out an adjustment path, expressed in terms of net expenditure (primary expenditure net of discretionary revenue measures, cyclical unemployment expenditure and one-off and temporary items), and covering in principle the same period as the term of that country's legislature (four or five years). By the end of the adjustment period (ie. four years at a minimum with the possibility of an extension of up to a maximum of seven years, conditional on investment and reform commitments), debt should be on a plausibly downward path or staying below 60 percent of GDP, with the deficit remaining below 3 percent of GDP over the medium term (defined as the 10-year period after the end of the adjustment). This forward-looking requirement is verified at the time of the endorsement of the plan, based on projections with unchanged policies, carried out according to a European Commission debt sustainability analysis (DSA) methodology.

The budgetary targets implied by the requirements can be expected to differ substantially across countries, depending on the starting level of debt and the projected rates of interest and GDP growth.

While maintaining these common risk-based requirements, the April 2024 reform has introduced two further numerical constraints ('safeguards'), which apply to countries with both debt above 60 percent of GDP and deficits above 3 percent of GDP. Specifically:

A debt sustainability safeguard requires the projected debt-to-GDP ratio to decrease by a minimum annual average amount of 1 percent of GDP for countries with debt ratios above 90 percent, and by 0.5 percent of GDP for countries with debt ratios between 60 percent and 90 percent of GDP, over the adjustment period. However, if a country is subject to the excessive deficit procedure (EDP) on grounds of a deficit in excess of 3 percent of GDP, the requirement does not apply before the year in which the country is projected to exit the procedure.

A deficit resilience safeguard requires that the overall deficit should eventually reach a level of no more than 1.5 percent of GDP in structural terms. Countries with higher deficits are required to adjust by a minimum of 0.4 percent of GDP per year (or 0.25 percent of GDP per year, if the adjustment period is extended to seven years).

Compliance with these requirements is verified by the European Commission and the Council of the EU, which can, if necessary, ask for revisions. Once a plan is endorsed by the Council, the net expenditure path becomes the sole reference for assessing compliance with the fiscal rules. Positive and negative deviations from the net expenditure path are accumulated in a notional 'control account'. If the balance reaches a certain threshold (0.3 percent of GDP in one year or 0.8 percent in two years), countries with debt in excess of 60 percent of GDP are liable to be subject to the EDP for breach of the debt criterion, with the associated prescriptions and eventual penalties. Countries that keep debt below the threshold of 60 percent of GDP are not liable to consequences under the new rules, other than the possibility of warnings and soft-law recommendations, unless they breach the 3 percent of GDP deficit threshold[3].

behaviour of their more deficit-prone partners (hence the insistence of fiscal hawks on retaining other numerical fiscal constraints). High-debt, high-deficit countries, meanwhile, appreciate the bespoke nature of the procedure, including the opportunities it offers for renegotiation if the corrective path is not met.

Importantly, the relative effectiveness of the deficit-based EDP, which is confirmed by empirical evidence (De Jong and Gilbert, 2020; Caselli and Wingender, 2021), appears to depend more on the stigma associated with governments being made subject to the procedure, rather than on the hypothetical possibility of financial sanctions. Stigma may include the impact of the EDP on government debt risk premia (Diaz Kalan *et al* 2018).

The debate on the reform has focused on the design of the rules, in particular the tension between the original objective of medium-term debt sustainability and the numerical constraints ('safeguards') that have been added to the requirements of the plans. This has distracted attention from the equally, if not more, important aspects of implementation, which the reform has left unaddressed, or at least potentially open to very different outcomes.

In particular, the reform has failed to clarify the interplay between the EDP and the new preventive arm in important respects. This risks compromising the functioning of the new framework from the outset. Many countries currently have both high debts and high deficits. Addressing these is urgent, requiring smooth coordination in the deployment of the SGP's old and new tools.

Specifically, preliminary simulations by Darvas *et al* (2024) suggest that if the debt-sustainability based adjustment requirements are applied rigorously, in about half a dozen cases the MTFSPs will have to set an annual fiscal adjustment in excess of 0.5 percent of GDP (in terms of structural primary balance), to be sustained for as long as seven years, something for which there is hardly any precedent.

At the same time, about a dozen countries are expected to immediately enter the EDP and receive adjustment prescriptions because of deficits persistently in excess of 3 percent of GDP. Several countries, including in particular Italy, France, Spain and Belgium, will likely be affected by both sets of prescriptions.

Note that high-debt countries are not immediately exposed to the debt-based EDP, because under the reform, the procedure can be triggered only by an accumulated deviation from the adjustment path in the MTFSP, while the first cohort of MTFSPs will be endorsed by the EU more or less simultaneously with the opening of the EDPs, which therefore will be only deficit-based. Although there is still some uncertainty on the timing of the procedures, the expectation is that EDPs will be opened and MTFSPs endorsed in autumn 2024[4].

Last but not least, the interplay between the EDP and the new preventive arm cannot be properly understood if one neglects two essential contextual elements, which do not stem from the new rules as such but can be inferred from a systematic reading of the EU fiscal governance legal framework (Pench, 2024):

- In spite of the common-parlance distinction between deficit-based and debt-based EDPs, legally there is only one procedure. This means that, once the EDP has been opened based on one criterion, a second procedure based on the other criterion cannot be superimposed on the existing procedure. Conversely, the closure ('abrogation') of an EDP opened based on one criterion should or even must be subordinated to the satisfaction of both criteria.

- The wide discretion enjoyed by the European Commission and Council in setting, and, if necessary resetting, adjustment paths, including departures from the apparently rigid benchmarks in the corrective arm, as long as a country is subject to an EDP[5].

> A further implication is that the application of the provisions on fiscal-structural plans should not result in undue restraint on the operation of EDPs, and in case of apparent conflict the latter should prevail. This principle, which has been described as 'the primacy of the EDP' reflects the strength of the respective legal bases and is firmly established in surveillance practice[6].

Taken together, these elements point to serious risks that need addressing when launching and implementing the forthcoming EDPs. Hopefully, they also suggest possible responses.

Risk 1. Defining the initial corrective path for countries subject to the EDP

It stands to reason that, if EDPs are opened at about the same time as MTFSPs are endorsed by the Council, the prescribed fiscal adjustment path should be the same, at least as long as the periods covered by the two procedures coincide.

It is not sure, however, that adjustment paths will be fully in line with the debt-sustainability requirements of the new preventive arm.

The principle of the primacy of the EDP over the preventive arm suggests that the adjustment in the MTFSPs would have to be aligned to that prescribed in the EDP. This conclusion is confirmed by a provision in the new preventive arm requiring that the trajectories that should serve as a reference for the MTFSP show *"consistency with the corrective path"* in the applicable decisions under the EDP (Regulation 1263/2024, Article 6(d)). In turn, for deficit-based EDPs, the reformed EDP regulation specifies only a *"minimum annual structural adjustment of at least 0,5 percent of GDP as a benchmark"* [sic] (Regulation 1264/2024, Article 3(4)).

Moreover, for 2025-2027, the regulation contains an ad-hoc provision allowing a downward departure from the 0.5 percent of GDP benchmark adjustment[7]. The reading of the provisions is complicated further by the fact that

the 0.5 percent of GDP benchmark adjustment is defined in terms of total structural balance, while the individual adjustment path prescribed to the countries by the EDP and the MTFSPs should be in terms of net expenditure, that is, approximatively, in terms of structural primary balance.

Bearing also in mind the wide discretion enjoyed by the Commission and the Council in setting the individual adjustment path under the EDP, there is reason to be concerned that that the adjustment paths in the forthcoming EDPs will focus on the deficit target of 3 percent of GDP, while falling short of the adjustment required, on an annual basis, to satisfy the debt sustainability requirements of the new preventive arm.

This would paradoxically result in more favourable treatment of the countries subject to an EDP, relative to countries that have already brought their deficits below 3 percent of GDP. Nor would demanding that the countries make up for the gap relative to the debt sustainability requirements after they have brought their deficits below 3 percent of GDP be a satisfactory solution, because of the obvious questions of credibility it would raise.

The logical way to address this risk, from both the economic and legal points of view, would be to clarify that the minimum benchmark adjustment under a deficit-based EDP (including the temporary exception for 2025-2027) should not be interpreted as allowing for individual adjustment paths inconsistent with the debt-sustainability requirements of the preventive arm.

Risk 2. Divergence from the MTFSP during the implementation of the corrective path agreed under the EDP

The consistency in principle between individual adjustment paths under the EDP and in MTFSPs set out at start of the process does not mean that inconsistencies might not arise, for two reasons.

First, the narrow focus of the deficit-based EDP on bringing down the deficit to 3 percent of GDP introduces a 'nominal bias' in the working of the procedure: a government that is on its way toward the nominal target of 3 percent of GDP does not have to face demands for budgetary correction, irrespective of whether or not it has delivered on the prescribed structural adjustment included initially in the MTFSP. Specifically, as long as a country achieves its nominal deficit targets, escalation of the EDP – potentially leading to sanctions – is not an option[8].

It is therefore not difficult to imagine a scenario in which a country complies with the EDP recommendation – or, more precisely, it cannot be penalised for departing from it – while deviating from the adjustment path, for example, through recourse to temporary measures, or thanks to windfall revenues.

This may be less of a problem than it seems, at least as long as it does not lead to the country exiting the EDP (see Risk 3). If the EDP covers several years, which is bound to be the case for countries starting from high deficits, it is anyway not very likely that a country will hit nominal deficit targets year after year without a corresponding structural adjustment.

A more serious reason why the initial structural adjustment may fall by the wayside is existence of another bias in implementation of the procedure, as distinct from its design. This is a 'no-escalation bias', referring to the reluctance of the Commission and the Council to escalate the EDP even when a country deviates from the structural and the nominal adjustment path.

Instead, the practice has been to issue a revised EDP recommendation with an extended deadline. While the adoption of a single indicator should make it easier to determine whether the adjustment has been delivered or not, incentives to fudge would persist, especially given the heavy penalties, both direct and indirect, that could accompany the escalation of an EDP (Box 2).

Since the no-escalation bias essentially reflects a tendency to depart from what the rules would prescribe when this would lead to politically awkward consequences, it is hard to tackle the problem by looking only at the rules and suggesting a different interpretation.

The reputational cost of refraining from escalation could be increased, for example by asking the European Fiscal Board (EFB), the independent advisory body established by the Commission, to provide its advice, a possibility introduced by the reform.

Note however that, since the right to ask for an EFB opinion is limited to the Commission and the Council, this reputational risk could be safely ignored if the Commission and the Council agreed to pretend that there has not be deviation from the structural adjustment path, and to extend the deadline for correcting the excessive deficit via a new EDP recommendation.

Another possibility would be that of reducing the potential fines attached to the escalation of the EDP to symbolic amounts, to give them a purely reputational effect.

The reform is silent on what would happen to an MTFSP in case the EDP recommendation is revised. The principle of the primacy of the EDP suggests that in case a revised EDP recommendation is made, an MTFSP, including in particular the adjustment effort, should be revised correspondingly, in spite of the provisions in the preventive arm that are meant to discourage changes in the structural adjustment path.

This would restore consistency between the MTFSP and the EDP, at the price of a delayed adjustment. One could note that what matters for debt sustainability is the size of the total adjustment; simulations suggest debt dynamics would hardly change if the fiscal adjustment took a few more years.

Therefore, it might be concluded that the problem is less serious than it at first appears, provided that pressure to adjust under the EDP is maintained. This proviso is crucial and coincides with the main risk to be addressed when launching and implementing the forthcoming EDPs.

Risk 3. Exit from an EDP might be based only on achieving the 3 percent of GDP deficit, turning the framework into a 'free for all'

Possibly the most important question that the SGP reform has left unaddressed concerns the conditions for exiting an EDP ('EDP abrogation'), in particular in cases when the EDP was opened based only on the deficit criterion (which, as noted, will be the case for all the forthcoming EDPs, including for high-debt countries).

The formulation of the relevant provisions in the new EDP regulation seems to suggests that, in case an EDP was not opened based on the debt criterion, ie. it was opened only for breach of the 3 percent of GDP deficit threshold, the procedure should be closed as soon as the deficit has been durably brought under 3 percent of GDP. The relevant provision (Regulation 1264/2024, Article 8(3)) is worth quoting more fully:

"A Council decision shall only be taken pursuant to Article 126(12) TFEU [Council decision abrogating decisions or recommendations under the EDP "to the extent that the excessive deficit … has in the view of the Council been corrected"] where the deficit has been brought below the reference value and is projected by the Commission to remain so in the current and following year and, where the excessive deficit procedure was opened on the basis of the debt criterion, the Member State concerned respected the corrective net expenditure path set by the Council in accordance with Article 3(4) or Article 5(1) of this Regulation" [Council recommendation for the correction of the excessive deficit or Council decision to give notice to take measures for deficit reduction (escalation of the EDP in case of no effective action in response to the recommendation)].

Box 2. Escalating an EDP; a 'nuclear option' that will never be exercised?

The EU Treaty envisages the possibility of sanctions, including fines, only after the repeated failure by a country subject to an EDP to take effective action to correct the excessive deficit (Article 126(11) TFEU). The SGP was initially limited to specifying the amount of the potential fines.

To strengthen the enforcement of the fiscal rules, the 2011 'six-pack' reform of the SGP introduced further sanctions, of 0.2 percent of GDP, at an earlier stage in the procedure – after the finding that the country had not taken effective action in response to the initial adjustment recommendations received under the EDP.

Moreover, the sanctions were expected to be triggered 'automatically' by the Commission once the Council had established, based on a Commission proposal, that the country had not taken effective action under the EDP.

The new enforcement provisions were tested first in 2015, following a substantial apparent deviation by France from the adjustment recommended under the EDP. On that occasion the Commission resorted to a double-negative formulation – *"Overall … the available evidence does not allow to conclude on no effective action"* – to avoid proposing to the Council that it should establish that France had not taken effective action (European Commission 2015).

This episode provided the background to the controversial statement by then Commission President Juncker, who said the apparent breach of the fiscal rules by France was ignored by the Commission *"because it is France."*[9]

The second test of the enforcement provisions occurred in 2016, when the Commission and the Council found Portugal and Spain liable for no effective action with respect to their respective EDP recommendations. The Commission however evaded the obligation to trigger the imposition of a fine of 0.2 percent of GDP by recommending that the Council simply cancel the fine (Council of the EU, 2016, 2017a). Critics of the decision presented the episode as the 'death' of the six-pack reform[10].

The enforcement provisions have not been invoked since, though they remain in force and were left largely untouched by the 2024 reform[11].

In addition to the above sanctions, which apply only to euro area countries, the decision to escalate the EDP has become the potential trigger for the suspension of funds under the European Structural and Investment Funds (ESIF) and the Recovery and Resilience Facility (RRF). It could also result in the loss of eligibility for government security purchases by the European Central Bank under the Transmission Protection Instrument.

The decision to escalate the EDP may therefore be seen as a 'nuclear' option, a perception that would likely reinforce the observed no-escalation bias.

The previous sections have highlighted the risk that the forthcoming EDPs may fail to achieve compliance on the part of high-debt countries with the debt-sustainability requirements that are central to the reform, either because the initial adjustment path may not be sufficiently stringent (Risk 1) or because the countries may deviate from the adjustment path without facing serious consequences (Risk 2).

Our analysis of the second risk concluded, however, that even if the required adjustment was less than fully complete and was delayed relative to the initial timeline, the EDP should eventually be able to put the debt dynamics on a safe path.

An exit from an EDP based only on achieving the 3 percent of GDP deficit would undermine this reassuring conclusion.

If, as the formulation of the provisions on abrogation may lead them to expect, high-debt countries are able to exit the EDP solely by bringing the deficit below 3 percent of GDP, irrespective of where they stand relative to the debt-sustainability requirements, there is reason to be afraid that the debt-sustainability requirements will never be enforced.

One might argue that a high-debt country that exits a deficit-based EDP with an accumulated deviation relative to the adjustment path in its MTFSP (although likely revised from its initial version, to reflect intervening revisions in the EDP), should immediately face the (re-) opening of the EDP based on breach of the debt criterion.

Evoking this scenario explicitly is equivalent to showing its implausibility. Against the unpalatable prospect of revolving-door EDPs, the temptation will be simply too strong to exploit all the leeway available under the rules.

In particular, the Commission and the Council might leverage the ambiguities in the provisions allowing for the revision of MTFSPs and the resetting of the 'control account' in which the deviation from the adjustment path is recorded, to make a 'clean slate' of all past deviations when countries exit the EDP. The Commission and Council might ask those countries simply to submit an entirely new MTFSP.

A clean-slate scenario for high-debt countries after achieving the 3 percent of GDP deficit is likely to have substantial negative ramifications for the working of the entire framework.

Besides constituting a source of moral hazard, it would exacerbate the inequality of treatment across countries. Compared to those starting from lower deficits, the countries initially made subject to an EDP for breach of the 3 percent of GDP deficit threshold would effectively be granted a potentially much longer adjustment to satisfy the debt-sustainability requirements of the preventive arm.

It is not difficult to imagine that, to mitigate this inequality of treatment, the Commission and the Council would adopt a comparably lenient attitude toward deviations from the adjustment path by countries with deficits below 3 percent of GDP.

Specifically, such deviations might never be considered sufficient reason for opening debt-based EDPs against these countries, even if their debts exceeded the 60 percent of GDP threshold. In this connection, one should recall that, although the EDP regulation contains a presumption that, for countries facing *"substantial public debt challenges"*[12], a deviation from the adjustment path in the MTFSP should lead to the opening of the debt–based EDP, the presumption is far from absolute.

The Commission and a Council continue to enjoy wide discretion in assessing all the *"relevant factors"* before deciding the opening of the EDP for breach of the debt criterion. This stands in contrast with the 'quasi-automatic' opening of the EDP for breach of the 3 percent of GDP deficit threshold.

The outcome of this scenario would be an early disintegration of the SGP reform into a 'free for all', with the 3 percent of GDP deficit threshold remaining, effectively, the only fiscal rule that countries would need to care about.

While assessments of the likelihood of this scenario might differ, it should be noted that it retraces almost exactly the history of the first attempt to operationalise the debt criterion of the EDP by the 'six-pack' reform: the adoption of a transitory regime before the full application of the newly introduced 'debt rule' ended up in the rule never being applied (Box 3).

There is however a solution that would avoid the risk of an early degeneration of the reform. It would require a clarification that, even if a country has been placed in an EDP only for breach of the deficit criterion, it should also satisfy the debt criterion of the EDP for the procedure to be abrogated.

At first sight this specification would seem to contradict the provisions quoted above, which seem to entitle a country to exit the procedure once it has brought its deficit below 3 percent of GDP, if the EDP was opened based on the deficit criterion.

However, the provisions could be read as implying that, for countries with debt in excess of 60 percent of GDP, the deficit condition should be considered as necessary, but not sufficient for the abrogation of the EDP. An argument supporting this reading is that it would be fully in line with the specifications on the abrogation of the EDP agreed by the Council in the aftermath of the 'six-pack' reform, which posed for the first time the question of the interplay between a deficit-based and debt-based EDP.

To make the proposed solution work, two further questions would need to be addressed.

The first concerns the adjustment path that that deficit-based EDPs should prescribe to countries with debt in excess of 60 percent of GDP. If the adjustment is to be conducive to satisfying also the debt criterion, then it would seem evident that, even if opened for breach of the 3 percent deficit threshold, an EDP should cover the entire adjustment period under the MTFSP, the rationale of which is to achieve debt sustainability[14].

This would reinforce the conclusion that the adjustment path under an EDP should not be inconsistent with the debt-sustainability requirements of the preventive arm (Risk 1).

The second question is how to define the condition for abrogation of the debt-based EDP, which would have to apply to all countries subject to an EDP, if their debt exceeds 60 percent of GDP. The abrogation provisions quoted above make sufficiently clear that a debt-based EDP cannot be abrogated if the country does not respect the structural adjustment ('net expenditure') path. The provisions are not equally clear, however about the length of the period during which compliance with the net expenditure path should be ensured, before the EDP can be abrogated.

If it is accepted that the period covered by the EDP should in principle coincide with that covered by the MTFSP, this would seem a natural reference for verifying that the abrogation condition has been satisfied. Two further specifications could be added to complete the conditions to qualify for abrogation: allowing for early abrogation of the EDP in case the sustainability requirements have been achieved (an unlikely scenario for high-debt countries); and requiring a minimum period of uninterrupted adherence to the net expenditure path before abrogation.

Note that the latter specification was explicitly envisaged in the original Commission proposal for reforming the EDP regulation[15]. It could be reintroduced in the revision of the Code of Conduct or analogous specifications following the entry into force of the reform.

Finally, it would be desirable that the conditions for escalating the EDP should be clarified, in particular whether the above-described 'nominal bias' (Risk 2) should continue to constrain the operation of the debt-based EDP.

Conclusion

There is a risk that implementation of the forthcoming deficit-based EDP could lead to an unravelling of the entire reform. This could happen if the implementation allows high-debt countries: i) to undertake, at least initially, less fiscal adjustment than they would be required to under the new preventive arm, that is, if they were outside an EDP; ii) to further postpone the adjustment in the course of an EDP without having to face consequences; and, above all, iii) to exit an EDP purely based on the achievement of the 3 percent of GDP deficit threshold. This may not be the intention of the prescriptions that countries will receive when the EDPs are opened. It may well be the outcome by the time the EDPs are closed, based on past experience with the implementation of the SGP.

Preventing this requires a common understanding between the Council and the Commission that:

1. The adjustment path prescribed under the deficit based-EDPs should be consistent with the debt-sustainability requirements of the preventive arm.

2. A deviation from the initial adjustment path during should not result in a revised path that moves further away from the debt-sustainability requirements and does not trigger any penalties under the EDP.

3. Even if a country has been made subject to an EDP only for breach of the deficit criterion, it should also satisfy the EDP debt criterion for the procedure to be abrogated. The conditions for abrogation and escalation of the debt-based EDP should be clarified.

A common understanding on these points could be put forward by the European Commission and endorsed by the Council.

One might object that the understanding proposed under 1), while fully in line with the overall logic of the reform, would run, in the author's view, against an apparent tacit understanding reached at the time of the adoption of the reform allowing for some backloading of adjustment (as reflected in particular by the temporary relaxation of the normal adjustment requirement under the EDP).

Irrespective of the existence or the value of such a tacit understanding, allowing any temporary deviation from the debt-sustainability requirements would make even more important to affirm the understandings proposed under 2) and, crucially, 3).

Enforcement has consistently proved the weakest link in the system of EU fiscal rules. The 2024 reform will be judged a success not for having managed to achieve a fragile consensus on new rules, but if the new rules are shown to improve the incentives for countries to avoid potentially unsustainable debt trajectories. ■

Lucio Pench is a Non-Resident Fellow at Bruegel

Box 3. The lack of enforcement of the 1/20th rule: a cautionary tale

One of the main features of the 2011 'six-pack' reform was the so-called '1/20th rule' – a requirement for countries with debt above 60 percent of GDP to reduce it by an annual average of at least 5 percent of the difference between the debt level and 60 percent. Countries that failed to make this minimum adjustment were to be placed in a debt-based EDP.

The question was how to treat countries that had been placed in the EDP on the basis of the deficit criterion before the entry into force of the reform. It was decided that these countries would be given a three-year transition period, during which they would not be liable for a debt-based EDP, provided that they made sufficient progress towards compliance with the benchmark. The Commission was even tasked with producing a numerical indicator to gauge progress towards compliance[13].

Effectively, however, once countries exited the deficit-based EDP, non-compliance with the debt criterion, either in its transitory or permanent formulation, never resulted in an EDP being opened based on the debt criterion.

Even when the Commission clarified that compliance with the preventive arm of the SGP would be considered a key relevant factor in assessing compliance with the debt criterion (effectively sidelining the debt-reduction benchmark), no debt-based EDP was activated, irrespective of the persistent lack of compliance with the preventive arm, in particular, by countries with the highest debt ratios (Commission, 2020).

Endnotes

1. Each arm of the SGP corresponds to an EU regulation. The EDP is based on specific Treaty provisions (Article 126 of the Treaty on the Functioning of the EU), the application of which is further specified by Regulation (EC) No 1467/97. The rules of the preventive arm meanwhile were first formulated through Regulation (EC) No 1466/97, based on the general Treaty provisions on the coordination and surveillance of economic policies (Article 121 TFEU). SGP reforms in 2005 and 2011 amended both regulations. The 2011 reform (included in the so-called 'sixpack' package) sought in particular to strengthen the enforcement of the fiscal rules and implement the hitherto ignored EDP debt criterion (see Boxes 2 and 3). The April 2024 reform of the SGP replaces the preventive arm regulation with Regulation (EU) 2024/1263 and amends the EDP regulation.

2. See Council of the EU press release of 29 April 2024, 'Economic governance review: Council adopts reform of fiscal rules'.

3. For further details and an overall economic assessment of the new rules, see Pench (2024) and Darvas et al (2024). See also Jeromin Zettelmeyer, 'Assessing the Ecofin compromise on fiscal rules reform', First Glance.

4. The European Commission (2023c) stated that the EDP would be reactivated in spring 2024, following the official release by Eurostat, by end-April 2024, of the deficit and debt outturns for 2023 (see Eurostat release of 22 April 2024, 'Euro area government deficit at 3.6% and EU at 3.5% of GDP'). According to the EDP regulation, this would normally imply opening of EDPs within the subsequent four months. However, the final text of the new preventive arm contains a provision on the submission of the first MTFSPs, according to which countries "should submit their medium-term fiscal-structural plans by 20 September 2024" (Regulation 1263/2024, Article 36(a)), instead of the normal deadline of 30 April. Considering that the same regulation envisages a maximum of six weeks for the Commission to assess the plans, and that presumably initial consistency will be ensured between the adjustment path in the EDPs and the MTFSPs, it seems reasonable to expect that the opening of the EDP and the endorsement of the MTFSPs will take place at the same time in autumn 2024.

5. Possibly most graphic example of departure from the rules of the corrective arm in the individual prescriptions addressed to countries under an EDP concerns the EDPs opened in 2009-10 in the wake of the Great Financial Crisis, when

countries were urged to temporarily accommodate the increase in public deficits and even adopt expansionary policies, in contrast to the immediate fiscal consolidation in principle envisaged by the rules once a country is subject to the procedure (European Commission, 2010).

6. The primacy of the specific prescriptions of the EDP over the general provisions of the preventive arm is illustrated by the practice whereby, for countries subject to an EDP, the 'annual fiscal recommendation', part of the annual country-specific recommendations in the context of the European Semester (based on Article 121(3) TFEU), is limited to simply stating that they should respect the EDP recommendation that they have received (regardless of the provisions of the preventive arm that would apply otherwise).

7. Regulation 1263/2024, Recital 23. The smaller adjustment is meant to take into account the ongoing increase in the average interest rate on debt (making a 0.5 percent of GDP adjustment in terms of overall balance more demanding than the same adjustment in terms of the primary balance) and "not to compromise the positive effects of the Recovery and Resilience Facility."

8. This conclusion is reached by recursive reasoning starting from the observation that, in the case of an EDP covering a single year, if the country has brought the deficit below 3 percent of GDP, the deficit criterion of the EDP has been satisfied and the country cannot continue to be subject to the EDP on grounds of the deficit criterion, irrespective of whether or not the prescribed structural adjustment has been delivered. This has been consistently interpreted to imply that, for an EDP covering more than one year, the procedure cannot be escalated as long as the country can be considered to be on its way to eventually achieve the 3 percent of GDP deficit. Intermediate nominal targets were introduced to operationalise the otherwise ambiguous notion of being on the way toward the 3 percent of GDP. For reasons of symmetry in the operation of the procedure, intermediate deficit targets equally apply to debt-based EDPs. This approach was confirmed explicitly by the Code of Conduct of the Stability and Growth Pact endorsed by the ECOFIN Council (Council of the EU, 2017, p. 15): "For legal reasons, a deficit-based EDP cannot be stepped up if the Member State achieves its intermediate headline deficit target, even when the recommended change in the structural balance is not achieved. At the same time, though, a careful analysis should still be conducted to better understand the nature of the underlying budgetary developments."

While intermediate deficit targets are no longer specifically mentioned in the reformed EDP regulation, and the Code or analogous specifications will have to be revised to reflect the reform of the SGP, it is difficult to see how the 'nominal bias' could be eliminated, since it is a consequence of the role of the 3 percent of GDP deficit threshold in the EDP, which remains unchanged.

9. Francesco Guarascio, 'EU gives budget leeway to France 'because it is France'' – Juncker', Reuters.

10. Daniel Gros, 'The second death of the Stability Pact and the birth of an inter-governmental Europe', CEPS Commentary, 28 July 2016.

11. Regulation 1264/2024 (Article 12) reduces the sanctions envisaged by the Treaty (Article 126(11) TFEU) as an ultimate consequence of repeated non-compliance with EDP decisions from a minimum of 0.2 percent of GDP per year to 0.005 percent every six months. However, the sanctions introduced by the six-pack reform in connection with an escalation of the EDP remain unchanged at a default amount of 0.2 percent of GDP.

12. Regulation 1264/2014 Article 2(4). Substantial debt challenges are understood to refer to countries classified as 'high-risk' according to the medium-term sustainability risk classification of the Commission Debt Sustainability Monitor (European Commission, 2023a). In the 2023 Monitor, this category included Belgium, Greece, Spain, France, Croatia, Italy, Hungary, Portugal and Slovakia.

12. Regulation 1264/2014 Article 2(4). Substantial debt challenges are understood to refer to countries classified as 'high-risk' according to the medium-term sustainability risk classification of the Commission Debt Sustainability Monitor (European Commission, 2023a). In the 2023 Monitor, this category included Belgium, Greece, Spain, France, Croatia, Italy, Hungary, Portugal and Slovakia.

13. The indicator was meant to measure the distance between the current structural position of the country and the position consistent with the respect of debt reduction benchmark at the end of transition period (see European Commission, 2019).

14. A specific provision in the EDP Regulation 1264/2024 seems to confirm the possibility that an EDP opened on the basis of the deficit criterion may extend beyond bringing the deficit below 3 percent of GDP. Specifically, Article 3(4) of the

Regulation amending the EDP regulation prescribes that (emphasis added): "Where the excessive deficit procedure was opened on the basis of the deficit criterion, for the years when the general government deficit is expected to exceed the reference value, the corrective net expenditure path shall be consistent with a minimum annual structural adjustment of at least 0,5% of GDP as a benchmark" (sic).

15. The Commission proposal for revision of the EDP regulation (European Commission 2023b) required (Article 8(3)) that for the abrogation of a debt based EDP: "the Member State concerned respected the corrective net expenditure path set by the Council in accordance with Article 3(4) or Article 5(1) of this Regulation [Council recommendation for the correction of the excessive deficit or Council decision to give notice to take measures for deficit reduction (escalation of the EDP in case of no effective action in response to the recommendation)] over the previous 2 years and is projected to continue to do so in the current year on the basis of the Commission forecast."

References

Caselli, F and P Wingender (2021) 'Heterogeneous effects of fiscal rules: The Maastricht fiscal criterion and the counterfactual distribution of government deficits', European Economic Review 136(C).

Council of the EU (2016) 'Implementing Decision (EU) 2017/2350 of 9 August 2016 on imposing a fine on Portugal for failure to take effective action to address an excessive deficit'.

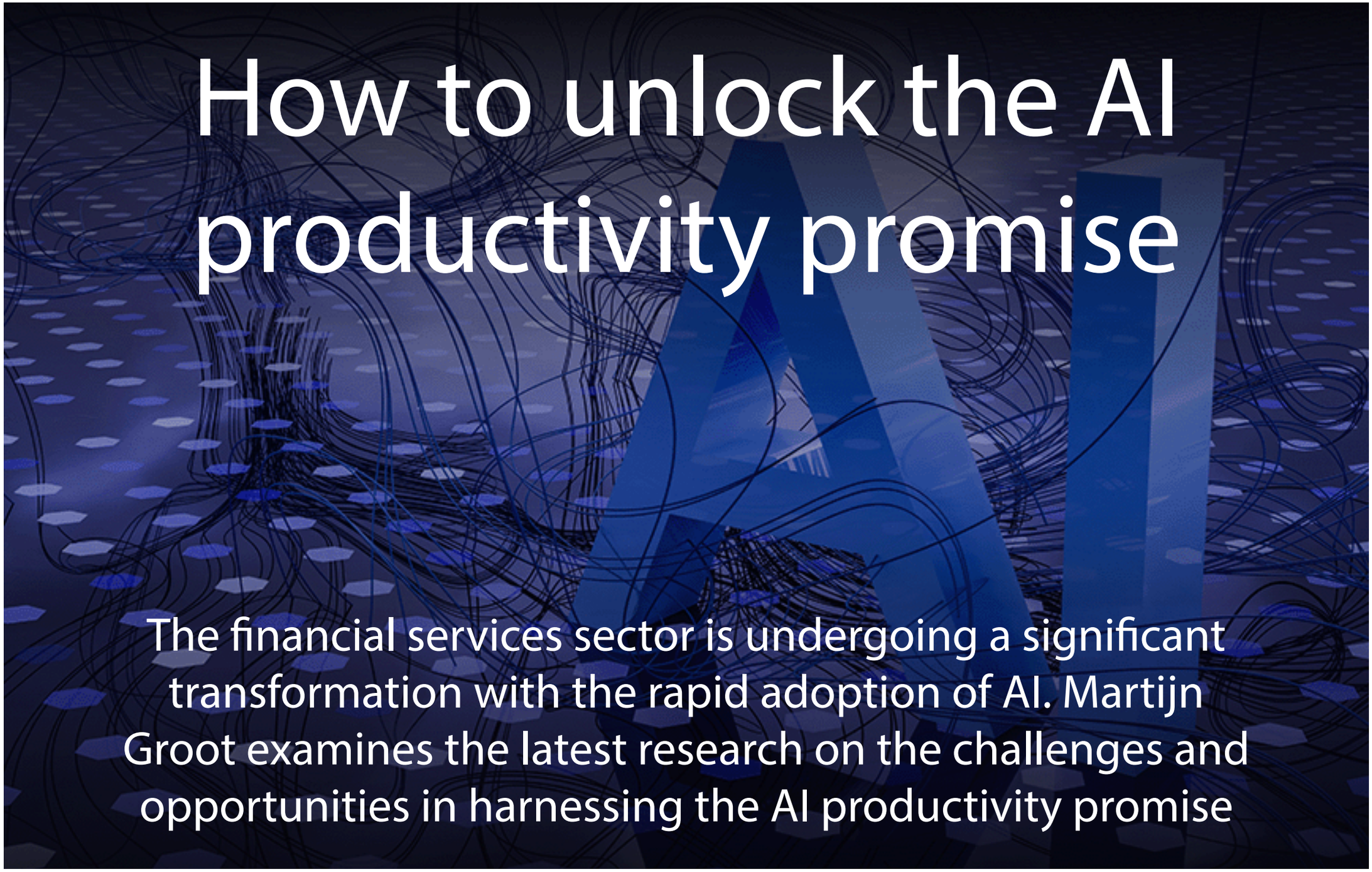
Council of the EU (2017a) 'Implementing Decision (EU) 2017/2350 of 9 August 2016 on imposing a fine on Spain for failure to take effective action to address an excessive deficit'.

Council of the EU (2017b) 'Specifications on the implementation of the Stability and Growth Pact and Guidelines on the format and content of Stability and Convergence Programmes' (Code of Conduct of the Stability and Growth Pact).

Darvas, Z, L Welslau and J Zettelmeyer (2024) 'Longer-term Term Fiscal Challenges Facing the EU', in F Fernández Méndez de Andés (ed) The Euro in 2024, Fundación Instituto Español de Analistas and Fundación ICO.

De Jong, FM and ND Gilbert (2020) 'Fiscal Discipline in EMU? Testing the Effectiveness of the Excessive Deficit Procedure', European Journal of Political Economy 61: 101822
Diaz Kalan, F, A Popescu and J Reynaud (2018) 'Thou Shalt Not Breach: The Impact on Sovereign Spreads of Noncomplying with the EU Fiscal Rules', IMF Working Papers 2018(087), International Monetary Fund.
European Commission (2015) 'Recommendation for a Council Recommendation with a view to bringing an end to the excessive government deficit in France', COM/2015/0115 final.
European Commission (2010) 'Public Finances in EMU – 2010', European Economy 4/2010.
European Commission (2019) 'Vade Mecum on the Stability and Growth Pact 2019 edition', European Economy Institutional Paper 101.
European Commission (2023a) 'Debt Sustainability Monitor 2022', Institutional Paper 199.
European Commission (2023b) 'Proposal for a Council regulation amending and Regulation (EC) No 1467/97 on speeding up and clarifying the excessive deficit procedure', COM/2023/241 final.
European Commission (2023c) 'Fiscal policy guidance for 2024', COM(2023) 141 final.
Pench, L (2024) 'The new Stability and Growth Pact: Innovation and Continuity in the light of Next Generation EU', in F Fabbrini (ed) Research Handbook EU Economic Law, Edward Elgar, forthcoming

This article is based on Bruegel Policy Brief Issue n°08/24 | May 2024.



How to unlock the AI productivity promise

The financial services sector is undergoing a significant transformation with the rapid adoption of AI. Martijn Groot examines the latest research on the challenges and opportunities in harnessing the AI productivity promise

The financial services sector is undergoing a significant transformation with the rapid adoption of artificial intelligence (AI). Recent studies indicate that AI is becoming an integral part of business operations across the industry. Alveo recently conducted a survey to poll the industry about the state of adoption, opportunities and challenges in adopting gen AI in enterprise data management.

Statista research from 2023 found that only 8% of financial businesses considered AI critical to their business in 2022 but by 2025, the expectation is that 43% will consider this to be so. Currently, nearly all financial services organisations are using AI in some capacity but there is a clear split between individual or departmental level experimentation versus more systemic adoption. 41% of firms have extensively deployed AI across different business operations, according to Alveo's research, indicating a growing trend towards more comprehensive adoption and embedding into workflows.

Benefits of AI in financial services

AI's implementation in the financial services sector offers numerous benefits, enhancing operational efficiencies, risk management, customer service, and product development. By automating routine tasks, AI allows financial institutions to process large volumes of data rapidly and accurately, reducing human errors and freeing up human resources for more complex tasks. This capability is particularly beneficial in areas such as operations but also in finance and risk management.

For example, AI-driven systems can handle vast amounts of transactional data, identifying discrepancies and potential fraud or money-laundering in real-time. This not only improves operational efficiency but also enhances security and compliance. Additionally, AI's ability to analyse customer data enables personalised customer interactions, improving customer satisfaction and loyalty. Financial institutions can offer tailored financial products and services, enhancing the overall customer experience.

AI also plays a crucial role in risk management by analysing vast datasets to identify patterns and potential risks that might be overlooked by human analysts. This predictive capability helps in proactive risk management and helping firms cope with increasing regulatory reporting requirements: in many banks a large portion of staff is needed to KYC and compliance and change budgets have necessarily been skewed towards regulatory compliance. AI adoption could help with more effectively KYC and reporting. Moreover, AI's ability to provide insights into market trends and customer behaviours can guide strategic decision-making, offering a competitive edge to financial institutions and increasing productivity by tailoring information collection and curation to specific user roles.

By taking a strategic, informed approach, financial institutions can harness the power of AI to drive efficiency, innovation, and growth

Furthermore, AI can enhance portfolio management by analysing market conditions and predicting asset performance, helping institutions optimise their investment strategies. By leveraging AI, financial firms can improve their decision-making processes, reduce operational costs, and increase their overall efficiency. The transformative power of AI lies in its ability to convert large amounts of raw data – both traditional market and reference data, as well as an increasing number of 'alternative' data sets - into actionable insights that drive business growth and innovation.

However, integrating AI into financial data management is not without its challenges. One clear impact seems to be an increasing premium on good quality data and data provisioning capabilities to feed the models. This will lead to increased data and technology cost which is only partially offset by a decrease of expected operations cost base; in Alveo's research sample, 63% of senior decision-makers in financial services anticipate an increase in data costs due to AI. Furthermore, 40% expect a rise in operational headcount, while 81% foresee increased IT spending in data management.

Challenges and barriers to AI adoption in enterprise data management

Despite the promising benefits, several barriers impede AI adoption in financial services data management. Technological limitations are identified by 50% of decision-makers as a significant barrier. Financial institutions often struggle with legacy systems that are incompatible with modern AI technologies. These outdated systems require substantial upgrades or replacements, which can be costly and time-consuming. Ensuring a seamless integration of AI into these systems necessitates a strategic overhaul, involving significant investment in new technologies and infrastructure.

Another major challenge is the ongoing lack of skilled personnel. Implementing and managing AI systems demands expertise in both new technology and the financial services domain. This combination of skills is scarce,

with 46% of respondents highlighting a shortage of skilled professionals as a critical obstacle to implementing AI in financial data management. Financial institutions need professionals who understand the deployment and integration into existing workflows of AI algorithms and can apply them to the financial services domain. Addressing this skills gap requires targeted training and recruitment strategies.

Data quality and licensing issues also loom large. ensuring high-quality data is vital for effective AI implementation, as AI systems rely heavily on accurate, consistent, and timely data. Poor-quality data can lead to incorrect predictions and decisions, undermining the effectiveness of AI applications. Additionally, licensing and compliance issues further complicate data management, especially with the advent of generative AI and the evolving legal frameworks around data usage. Financial institutions must navigate these complex legal landscapes to ensure they are using data ethically and legally.

Furthermore, the potential for AI bias and discrimination presents another significant challenge. AI systems learn from historical data, which can contain biases that are inadvertently incorporated into the models. This can lead to unfair outcomes, particularly in areas such as credit scoring and loan approvals. Regulatory frameworks on the use of AI in financial services are coming with the EU's AI act and its risk-based classification of risk levels for AI systems as the most salient example. Financial institutions must implement robust fairness and bias mitigation strategies to ensure their AI systems produce equitable and non-discriminatory results.

Addressing interoperability and data governance

To achieve the productivity increase that AI promises, financial institutions need to focus on interoperability between AI models and their existing workflows. This involves improving the way they provision models and broadening the traditional notions of data quality and data governance.

Traditional machine learning involved feature engineering, or preparing and tuning the data to *"give the models a hand."* The new models make for a very different, natural language-based interaction with business users which calls for training in prompt engineering or the natural language patterns to interact with models, as well as a good understanding of model limitations and risks.

A proactive approach to AI adoption emphasises the importance of improved data quality, provisioning, and governance. Financial institutions should invest in advanced data management technologies to support AI requirements. This includes data aggregation, cleansing, and validation systems to ensure data accuracy and relevance. Developing a skilled workforce is also essential. Targeted training and recruitment strategies are needed to bridge the skills gap, with institutions investing in upskilling existing employees and attracting new talent proficient in AI technologies and financial data management.

Enhancing data quality and management

A proactive approach to AI adoption emphasises the importance of improved data quality, provisioning, and governance. To optimise their use of AI, financial institutions should first of all, invest in data management technologies: Enhancing data management infrastructure to support AI requirements is crucial. This includes advanced data aggregation, cleansing, and validation systems to ensure data accuracy and relevance.

It is also important that financial services decision-makers collaborate with experts. Partnering with AI and data management experts can provide the necessary guidance and support to navigate the complexities of AI integration, ensuring a smoother transition to AI-driven operations.

The shift towards AI results in increased costs in data management and technology. Alveo's research indicates that 81% of firms expect a rise in IT spending due to AI. However, these costs can be offset by the long-term operational efficiencies and productivity gains AI brings.

Financial institutions should maintain transparency by clearly documenting the sources and training data for AI models to ensure accountability. Regularly reviewing and updating content licensing agreements to align with evolving legal landscapes and ensure compliance with data usage regulations is also crucial.

Additionally, financial institutions should implement continuous monitoring and auditing of AI systems to ensure they operate as intended and comply with regulatory standards. This involves establishing clear performance metrics and regularly evaluating AI models against these benchmarks. By maintaining rigorous oversight, financial institutions can detect and address any issues promptly, ensuring the reliability and effectiveness of their AI systems.

Furthermore, AI can enhance portfolio management by analysing market conditions and predicting asset performance, helping institutions optimise their investment strategies. By leveraging AI, financial firms can improve their decision-making processes, reduce operational costs, and increase their overall efficiency. The transformative power of AI lies in its ability to convert vast amounts of raw data into actionable insights that drive business growth and innovation.

The role of generative AI

Generative AI is set to revolutionise financial data management by producing synthetic data for various use cases, including model testing and scenario management. This capability allows financial institutions to test their AI models under a wide range of scenarios without risking real data. However, this requires clear guidelines on data usage and compliance to avoid legal and ethical pitfalls.

Financial institutions should maintain transparency by clearly documenting the sources and training data for AI models to ensure accountability. They should also regularly review and update content licensing agreements to align with the evolving legal landscape and ensure compliance with data usage regulations.

Generative AI also presents opportunities for creating new financial products and services. By leveraging synthetic data, financial institutions can explore innovative solutions that were previously not possible due to data constraints. This opens up new avenues for growth and competitive differentiation. However, the ethical use of generative AI must be prioritised to avoid potential biases and ensure fair and equitable outcomes.

To maximise AI's benefits, financial institutions need a strategic approach that combines investment in technology with a focus on human capital. This involves continuous learning and adaptation, establishing feedback loops for continuous improvement in data quality and model performance. Combining high-level strategic oversight with grassroots-level adjustments based on real-time data and user feedback ensures a comprehensive and effective AI integration.

Moreover, financial institutions should foster a culture of innovation and experimentation, encouraging employees to explore new ways of leveraging AI. By promoting a mindset of continuous improvement and adaptation, organisations can stay ahead of the curve and capitalise on the latest advancements in AI technology. This proactive approach will enable financial institutions to drive sustainable growth and remain competitive in a rapidly-evolving industry.

Embracing the AI-driven future

The journey towards AI integration in financial data management is challenging but essential for future competitiveness. By addressing the key barriers, enhancing data quality and governance, and adopting a strategic approach, financial institutions can unlock the full potential of AI. This transformation promises increased efficiency, innovation, and growth, positioning firms at the forefront of the digital age in finance.

The future of financial data management is intertwined with AI, and those who navigate this transition wisely will emerge as leaders in the industry. As AI continues to evolve, financial institutions must remain agile and proactive, continuously refining their strategies to harness the transformative power of AI effectively.

In conclusion, while the path to AI adoption in financial data management is fraught with challenges, the potential rewards are immense. By taking a strategic, informed approach, financial institutions can overcome these hurdles and harness the power of AI to drive efficiency, innovation, and growth. The future of financial data management is undeniably intertwined with AI, and those who navigate this transition wisely will emerge as leaders in the new era of finance. ■

Martijn Groot is VP Marketing and Strategy at Alveo

Should AI stay or should AI go?

There is considerable disagreement about the growth potential of AI. Francesco Filippucci, Peter Gal, Cecilia Jona-Lasinio, Alvaro Leandro and Giuseppe Nicoletti argue that this is dependent on domestic and global governance issues

Income and wellbeing gains in advanced economies have been held back by weak productivity performance. The growth rate of labour productivity declined in OECD economies from about 2% annual growth rate between the 1970s and 1990s, to 1% in the 2000s (Goldin *et al* 2024, Andre and Gal 2024). This poses a dramatic challenge for ageing societies and makes it harder to allocate resources for the green transition.

There is widespread enthusiasm about the growth potential of rapidly developing artificial intelligence (AI). Some analysts argue that, under reasonable conditions, AI could lead to large and persistent gains, on the order of adding 1–1.5 percentage points to annual growth rates over the next 10–20 years (Baily *et al* 2023, Artificial Intelligence Commission of France 2024, McKinsey 2023, Briggs and Kodnani 2023).

On the other hand, Acemoglu (2024) contends that the available evidence combined with the economic theory of aggregation supports only moderate total factor productivity and GDP growth impacts, on the order of about 0.1% per year. Recent work from the OECD provides a broad overview of AI's impact on productivity and discusses the conditions under which it is expected to deliver strong benefits, with a focus on the role of policies (Filippucci *et al* 2024).

AI as a new general-purpose technology

Given its transformative potential in a wide range of economic activities, AI can be seen as the latest general-purpose technology (Agrawal *et al* 2019, Varian 2019) – similar to previous digital technologies such as computers and the internet or, going back further, to the steam engine and electricity.

From an economic perspective, AI can be seen as a production technology combining intangible inputs (skills, software, and data) with tangible ones (computing power and other hardware), to produce three broad types of outputs:

- Content, such as texts or images (generative AI)
- Predictions, optimisations, and other advanced analytics, which can be used to assist with or fully automate human decisions (non-generative AI)
- Physical tasks when combined with robotics (including autonomous vehicles).

Further uncertainties surrounding AI include broader societal concerns. More immediate concerns relate to privacy, misinformation, and bias (possibly leading to exclusion in areas such as labour and financial markets), while longer-term concerns include mass unemployment or even existential risks

Additionally, AI has some peculiar features, even compared to previous digital technologies. These include the potential for being autonomous (less dependent on human inputs) and the capacity for self-improvement, by learning from patterns in unstructured data or leveraging feedback data about its own performance.

Altogether, these features imply that AI can boost not only the production of goods and services but also the generation of ideas, speeding up research and innovation (Aghion *et al* 2018).

Initial micro-level evidence shows large productivity and performance gains

According to our overview of the fast-growing literature, initial micro-level evidence covering firms, workers, and researchers is indicative of several positive effects from using AI. First, micro-econometric studies find that the size of the gains from non-generative AI on firms' productivity is comparable to previous digital technologies (up to 10%; see panel a of Figure 1).

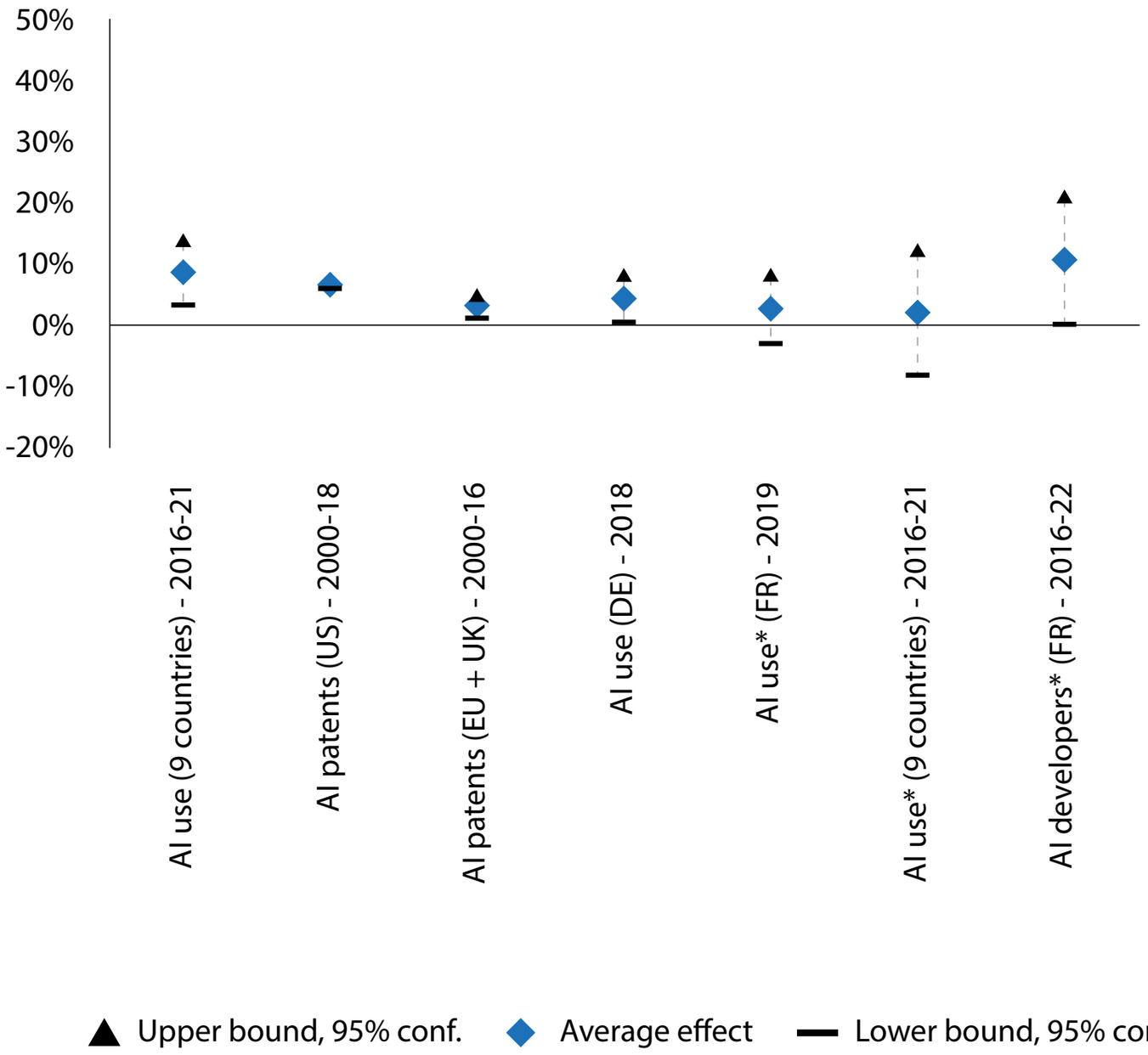
Second, when using more recent generative AI in various tasks – assisting in writing, computer programming, or customer service requests – the estimated performance benefits are substantially larger but vary widely (between 15 and 56%; see panel b of Figure 1) depending on the context.

In particular, Brynjolfsson *et al* (2023) found that AI has a much stronger impact on the performance of workers with less experience in their job. These estimates focus on specific tasks and individual-level gains. Hence, they are narrower in scope than previous firm-level studies but tend to rely more on more causal identification in experimental settings.

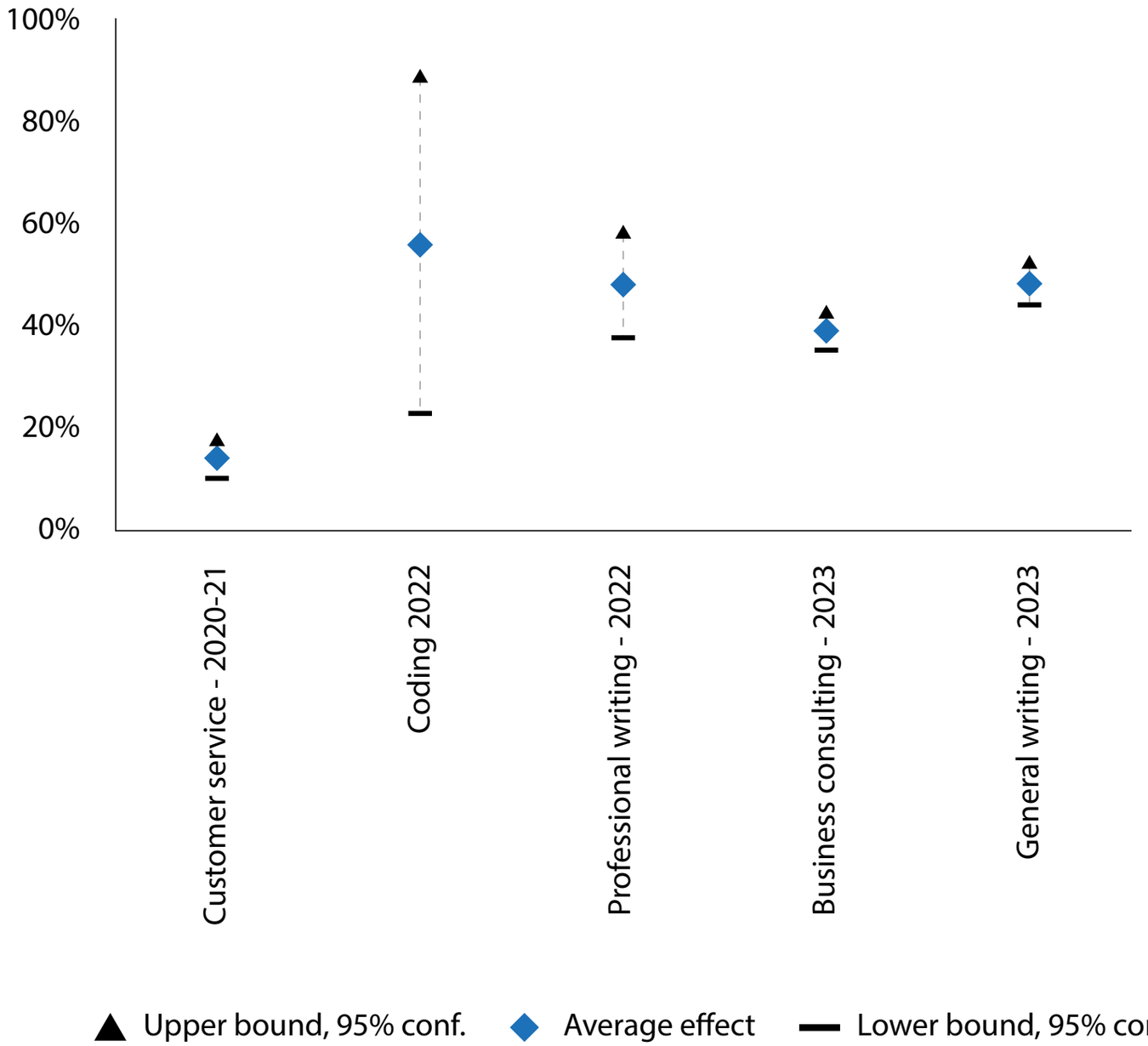
Third, researchers believe that AI allows for faster processing of data – speeding up computations and decreasing the cost of research – and may also make new data sources and methods available, as documented by a recent survey in Nature (Van Noorden and Perkel 2023).

Figure 1. The positive relationship between AI use and productivity or worker performance: selected estimates from the literature

a) Non-generative AI, firm-level studies on labour productivity



b) Generative AI worker-level studies on performance in specific tasks



*Note: In panel a, 'AI use' is a 0-1 dummy obtained by firm surveys, while 'AI patents' refers either to a 0-1 dummy for having at least one patent (US study) or to the number of patents in firms. The sample of countries underlying the studies are shown in parentheses. The year(s) of measurement is also indicated. *Controlling for other ICT technologies. For more details, see Filippucci et al (2024).*

Fourth, AI-related inventions are cited in a broader set of technological domains than non-AI inventions (Calvino *et al* 2023). Finally, there are promising individual cases from specific industries: AI-predicted protein-folding gives new insights in biomedical applications; AI-assisted discoveries of new drugs help with pharmaceutical R&D; and research on designing new materials can be broadly used in manufacturing (OECD 2023).

Long-run aggregate gains are uncertain

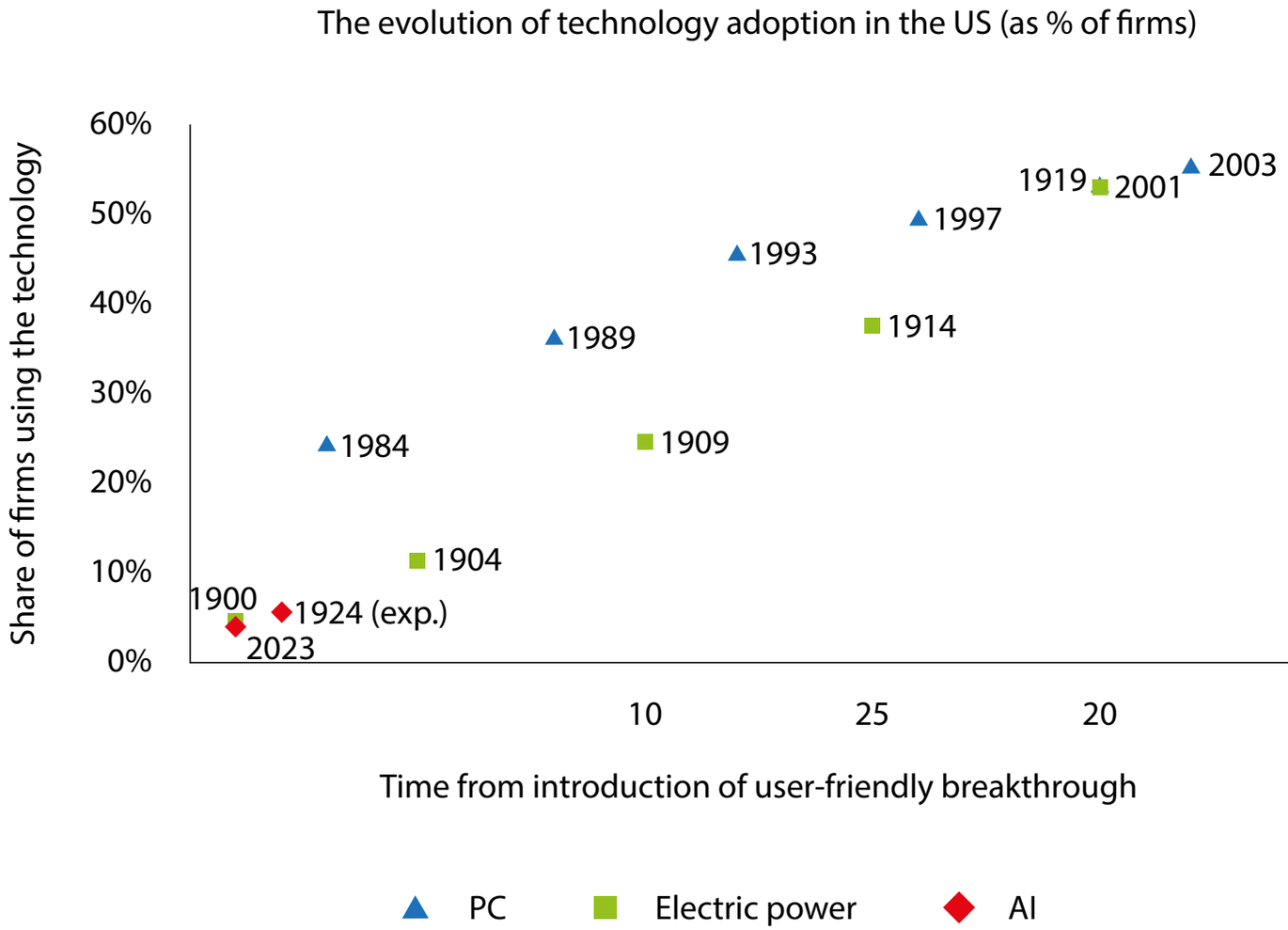
As generative AI's technological advances and its use are very recent, findings at the micro or industry level mainly capture the impacts on early adopters and very specific tasks, and likely indicate short-term effects. The long-run impact of AI on macro-level productivity growth will depend on the extent of its use and successful integration into business processes.

According to official representative data, the adoption of AI is still very low, with less than 5% of firms reporting the use of this technology in the US (Census Bureau 2024; see Figure 2). When put in perspective with the adoption path of previous general-purpose technologies (eg. computers and electricity), AI has a long way to go before reaching the high adoption rates that are necessary to detect macroeconomic gains.

While user-friendly AI may spread faster through the economy, the successful integration of AI systems and exploiting their full potential may still require significant complementary investments (in data, skills, reorganisations) which take time and necessitate managerial talent.

Moreover, future advances in AI development – and its successful integration within business processes – will require specialised technical skills that are often concentrated within a few firms (Borgonovi *et al* 2023).

Figure 2. AI adoption is still limited compared to the spread of previous general-purpose technologies



Note: The 2024 value for AI is the expectation (exp.) as reported by firms in the US Census Bureau survey. For more details, see the sources.
Source: For PC and electricity, Briggs and Kodnani (2023); for AI, US Census Bureau, Business Trends and Outlook Survey, updated 28 March 2024.

It is also an open question whether AI-driven automation will displace (reallocate) workers from heavily impacted sectors to less AI-affected activities or the human-augmenting capabilities of AI will prevail, underpinning labour demand.

Currently, AI exposure varies greatly across sectors: knowledge-intensive, high-productivity activities are generally much more affected (Figure 2), with significant potential for automation in some cases (Cazzaniga *et al* 2024, WEF 2023). Hence, an eventual fall in the employment shares of these sectors would act as a drag on aggregate productivity growth, resembling a new form of 'Baumol disease' (Aghion *et al* 2019).

Historically, the automation of high-productivity activities, combined with saturating demand for their output, has pushed employment from manufacturing to services (Bessen 2018). This structural change also played a role – though a moderate one – in the ongoing slowdown in aggregate productivity growth (Sorbe *et al* 2018).

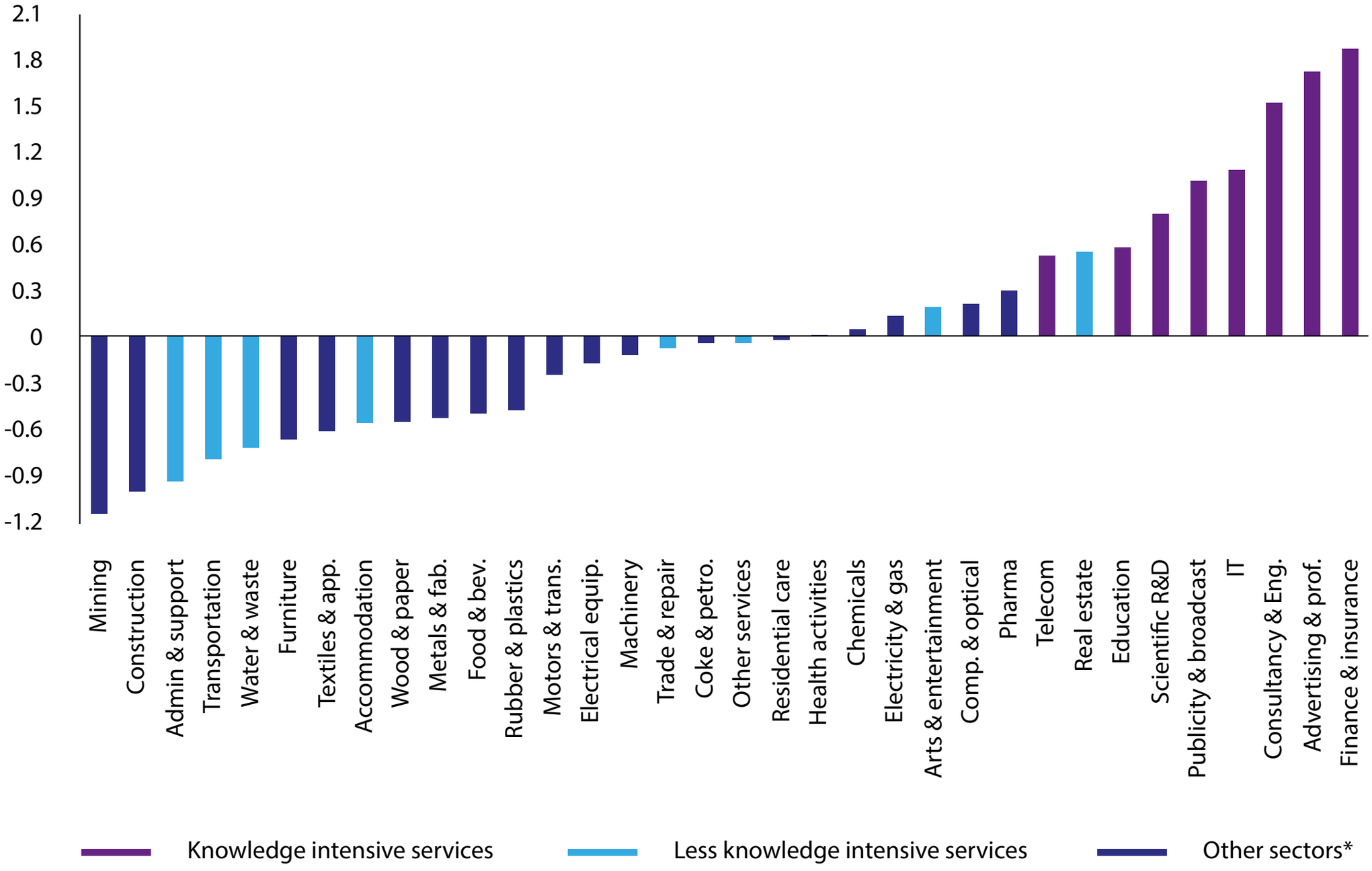
Similarly, if AI enhances productivity only in selected activities, aggregate growth will be limited by the slower productivity growth and higher employment share in sectors that are less exposed to AI (such as labour-intensive personal services like leisure and health care).

This may occur more quickly with AI compared to past technologies given the rapid and wide-ranging advances in its capabilities. However, in the extreme case of AI impacting (nearly) all tasks and boosting productivity in (nearly) all economic activities, this negative effect may be muted (Trammel and Korinek 2023).

AI poses policy challenges related to competition, inequality, and broader societal risks

AI poses significant threats to market competition and inequality that may weigh on its potential benefits, either directly or indirectly, by prompting preventive policy measures to limit its development and adoption.

Figure 3. High-productivity and knowledge-intensive services are most affected by AI



*Note: The index measures the extent to which worker abilities are related to important AI applications. The measure is standardised with mean zero and standard deviation one at the occupation level and then matched to sectors. Figure does not yet include recent Generative AI models. *Including non-market services, manufacturing, utilities, etc.*
Source: Filippucci et al (2024) and OECD (2024) based on Felten et al (2021).

First, the high fixed costs and returns to scale related to data and computing power may lead to excessive concentration of AI *development*. Second, AI *use* in downstream applications may lead to market distortions, especially if it allows first movers to build up a substantial lead in market share and market power.

Moreover, AI-powered pricing algorithms have a tendency to charge supra-competitive prices (Calvano *et al* 2020) and could eventually enhance harmful price discrimination (OECD 2018).

The impact of AI on inequality remains ambiguous. The technology can potentially substitute for high-skilled labour and narrow wage gaps with low-skilled workers, thereby reducing inequalities (Autor 2024) at least within occupations (Georgieff 2024).

Though there are indications that AI can be associated with higher unemployment (OECD 2024), AI could also lead to more inclusion and stronger economic mobility by improving education quality and access, expanding credit availability, and lowering skill barriers (eg. foreign languages).

Further uncertainties surrounding AI include broader societal concerns. More immediate concerns relate to privacy, misinformation, and bias (possibly leading to exclusion in areas such as labour and financial markets), while longer-term concerns include mass unemployment or even existential risks (Nordhaus 2021, Jones 2023).

A comprehensive policy approach is needed to effectively manage these risks and harness AI's full potential. Immediate priorities include promoting market competition and widespread access to AI technologies while preserving innovation incentives (eg. via adapting intellectual property rights protection) and addressing issues of reliability and bias, which require adequate auditing and accountability mechanisms.

Job displacement, reallocation and inequality impacts might emerge over longer periods, but they require preventive policy action through training, education, and redistribution measures to ensure human skills remain complementary to AI. Policymakers should also devise national and international governance mechanisms to cope with rapid and unpredictable developments in AI. ■

Francesco Filippucci is an Economist at the Organisation for Economic Co-Operation and Development (OECD), Peter Gal is Deputy Head of Division and Senior Economist at the OECD, Cecilia Jona-Lasinio is Professor of Applied Economics at Luiss Business School, Alvaro Leandro is an Economist at Caixa Bank and the OECD, and Giuseppe Nicoletti is Senior Fellow, LUISS Lab of European Economics, Luiss University

References
Acemoglu, D (2024), "The Simple Macroeconomics of AI", MIT, 5 April.
Acemoglu, D and T Lensman (2023), "Regulating Transformative Technologies", NBER Working Paper 31461, July.
Aghion, P, B Jones and C Jones (2018), "Artificial Intelligence and Economic Growth", NBER Working Paper 23928, October.
Agrawal, A, J Gans and A Goldfarb (2019), "Economic Policy for Artificial Intelligence", Innovation Policy and the Economy, vol. 19.
Andre, C and P Gal (2024), "Reviving productivity growth: A review of policies", OECD Economic Policy Papers, forthcoming.
Artificial Intelligence Commission of France (2024), "AI, Our Ambition for France" ("IA: Notre Ambition pour la France"), March.
Autor, D (2024), "Applying AI to Rebuild Middle Class Jobs", NBER Woking Paper 32140, February.
Baily, M, E Brynjolfsson and A Korinek (2023), "Machines of mind: The case for an AI-powered productivity boom", The Economics and Regulation of Artificial Intelligence and Emerging Technologies, Brookings, 10 May.
Bessen, J (2018), "AI and Jobs: The Role of Demand", NBER Working Paper 24235, January.
Borgonovi, F et al (2023), "Emerging trends in AI skill demand across 14 OECD countries", OECD Artificial Intelligence Papers, no. 2, Paris: OECD Publishing.
Briggs, J and D Kodnani (2023), "The Potentially Large Effects of Artificial Intelligence on Economic Growth", Global Economics Analyst, New York: Goldman Sachs.
Brynjolfsson, E, D Li and L Raymond (2023), "Generative AI at Work", NBER Working Paper 31161, November.
Calo, R (2013), "Digital market manipulation", The George Washington Law Review 82/995.
Calvano, E et al (2020), "Artificial intelligence, algorithmic pricing, and collusion", American Economic Review 110(10): 3267–97.
Calvino, F and L Fontanelli (2023), "A portrait of AI adopters across countries: Firm characteristics, assets' complementarities and productivity", OECD Science, Technology and Industry Working Papers, no. 2, Paris: OECD Publishing.

Cazzaniga, M et al (2024), "Gen-AI: Artificial Intelligence and the Future of Work", IMF Staff Discussion Notes, International Monetary Fund, 14 January.
Census Bureau (2024), Business Trends and Outlook Survey, 28 March.
Felten, E, M Raj and R Seamans (2021), "Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses", Strategic Management Journal 42(12): 2195–217.
Filippucci, F, P Gal, A Leandro, C Jona-Lasinio and G Nicoletti (2024), "The impact of Artificial Intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges", OECD Artificial Intelligence Papers, no. 15, Paris: OECD Publishing.
Georgieff, A (2024), "Artificial intelligence and wage inequality", OECD Artificial Intelligence Papers, no. 13, Paris: OECD Publishing.
Goldin, I, P Koutroumpis, F Lafond and J Winkler (2024), "Why Is Productivity Slowing Down?", Journal of Economic Literature 62(1): 196–268.
Haslberger, M, J Gingrich and J Bhatia (2023), "No great equalizer: experimental evidence on AI in the UK labor market", SSRN, 2 November.
Jones, C (2023), "The A.I. Dilemma: Growth versus Existential Risk", NBER Working Paper 31837, November.
McKinsey (2023), "The economic potential of generative AI", McKinsey Digital, 14 June.
Nordhaus, W (2021), "Are We Approaching an Economic Singularity? Information Technology and the Future of Economic Growth", American Economic Journal: Macroeconomics 13(1): 299–332.
OECD (2018), "Personalised Pricing in the Digital Era", Background Note by the OECD Secretariat for the joint meeting between the Competition Committee and the Committee on Consumer Policy, 28 November.
OECD (2024), "Labour Market Shortages, Mismatches and Megatrends", OECD Global Forum on Productivity, forthcoming.
OECD (2023), Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research, Paris: OECD Publishing, June.

Sorbe, S, P Gal and V Millot (2018), "Can productivity still grow in service-based economies? Literature overview and preliminary evidence from OECD countries", OECD Economics Department Working Papers, no. 1531, Paris: OECD Publishing.
Trammell, P and A Korinek (2023), "Economic Growth under Transformative AI", NBER Working Paper 31815, October.
Van Noorden, R and J Perkel (2023), "AI and science: what 1,600 researchers think", Nature 621(7980): 672–75.
WEF (2023), "Jobs of Tomorrow: Large Language Models and Jobs", World Economic Forum White Papers, September.

Authors' note: The main paper underlying this column (Filippucci et al. 2024) was developed within the Joint OECD-Italy's Department of Treasury Project for Multilateral Policy Support. This article was originally published on VoxEU.org.

How the financial authorities can take advantage of AI

AI will both help to the financial authorities and bring new challenges. Jon Danielsson and Andreas Uthemann discuss where AI can help, and what to watch out for

Artificial intelligence (AI) will likely be of considerable help to the financial authorities if they proactively engage with it. But if they are conservative, reluctant, and slow, they risk both irrelevance and financial instability. The private sector is rapidly adopting AI, even if many financial institutions signal that they intend to proceed cautiously. Many financial institutions have large AI teams and invest significantly; JP Morgan reports spending over $1 billion per year on AI, and Thomson Reuters has an $8 billion AI war chest.

AI helps them make investments and perform back-office tasks like risk management, compliance, fraud detection, anti-money laundering, and 'know your customer'. It promises considerable cost savings and efficiency improvements, and in a highly competitive financial system, it seems inevitable that AI adoption will grow rapidly.

As the private sector adopts AI, it speeds up its reactions and helps it find loopholes in the regulations. As we noted in Danielsson and Uthemann (2024a), the authorities will have to keep up if they wish to remain relevant.

So far, they have been slow to engage in their approach to AI and will find adopting AI challenging. It requires cultural and staff changes, supervision will have to change, and very significant resources will have to be allocated.

Pros and cons of AI

We see AI as a computer algorithm performing tasks usually done by humans, such as giving recommendations and making decisions, unlike machine learning and traditional statistics, which only provide quantitative analysis. For economic and financial applications, it is particularly helpful to consider AI as a rational maximising agent, one of Norvig and Russell's (2021) definitions of AI.

AI has particular strengths and weaknesses. It is very good at finding patterns in data and reacting quickly, cheaply, and usually reliably.

However, that depends on its having access to relevant data. The financial system generates an enormous amount of data, petabytes daily. But that is not sufficient.

A financial sector AI working for the authorities should also draw knowledge from other domains such as history, ethics, law, politics, and psychology, and to make connections between different domains, it will have to be trained on data that contain such connections.

The authorities need to be aware of AI benefits and threats and incorporate that awareness into the operational execution of the services they provide for society

Even if we can do so, we don't know how AI that has been fed with knowledge from a wide set of domains and high-level objectives will perform. When made to extrapolate, its advice might be judged as entirely wrong or even dangerous by human experts.

Ultimately, this means that when extrapolating from existing knowledge, the quality of its advice should be checked by humans.

How the authorities can implement AI

The financial authorities hold a lot of public and private information that can be used to train AI, as discussed in Danielsson and Uthemann (2024b), including:

1. Observations on past compliance and supervisory decisions

2. Prices, trading volumes, and securities holdings in fixed-income, repo, derivatives, and equity markets

3. Assets and liabilities of commercial banks

4. Network connections, like cross-institution exposures, including crossborder

5. Textual data
    - The rulebook
    - Central bank speeches, policy decisions, staff analysis
    - Records of past crisis resolution

6. Internal economic models
    - Interest rate term structure models
    - Market liquidity models
    - Inflation, GDP, and labour market forecasting models
    - Equilibrium macro model for policy analysis

Data are not sufficient; it also requires considerable human resources and compute. Bloomberg reports that the median salary for specialists in data, analytics, and artificial intelligence in US banks was $901,000 in 2022 and $676,000 in Europe, costs outside the reach of the financial authorities.

This is similar to what the highest-paid central bank governors earn. Technical staff earn much less (see for example Borgonovi *et al* 2023 for a discussion on the AI skill market).

However, it is easy to overstate these problems. The largest expense is training AI on large publicly available text databases. The primary AI vendors already meet that cost, and the authorities can use transfer learning to augment the resulting general-purpose engines with specialised knowledge at a manageable cost.

Taking advantage of AI

There are many areas where AI could be very useful to financial authorities. It can help micro authorities by designing rules and regulations and enforcing compliance with these rules. While human supervisors would initially make enforcement decisions, reinforcement learning with human feedback will help the supervisory AI become increasingly performant and, hence, autonomous.

Adversarial architectures such as generative adversarial networks might be particularly beneficial in understanding complex areas of authority-private sector interactions, such as fraud detection.

AI will also be helpful to the macro authorities, such as in advising on how to best cope with stress and crises. They can run simulation scenarios on alternative responses to stress, advise on and implement interventions, and analyse drivers of extreme stress. The authorities could use generative model models as artificial labs to experiment on policies and evaluate private sector algorithms.

AI will also be useful in ordinary economic analysis and forecasting, achievable with general-purpose foundation models augmented via transfer learning using public and private data, established economic theory, and previous policy analysis. Reinforcement learning with feedback from human experts is useful in improving the engine. Such AI would be very beneficial to those conducting economic forecasting, policy analysis and macroprudential stress tests, to mention a few.

Risks arising from AI

AI also brings with it new types of risk, particularly in macro (eg. Acemoglu 2021). A key challenge in many applications is that the outcome needs to cover behaviour that we rarely observe, if at all, in available data, such as complicated interrelations between market participants in times of stress.

When AI does not have the necessary information in its training dataset, its advice will be constrained by what happened in the past while not adequately reflecting new circumstances. This is why it is very important that AI reports measures of statistical confidence for its advice.

Faced with all those risks, the authorities might conclude that AI should only be used for low-level advice, not decisions, and take care to keep humans in the loop to avoid undesirable outcomes. However, that might not be as big a distinction as one might think. Humans might not understand AI's internal representation of the financial system. The engine might also act so as to eliminate the risk of human operators making inferior choices, in effect becoming a shadow decision-maker.

While an authority might not wish to get to that point, its use of AI might end up there regardless. As we come to trust AI analysis and decisions and appreciate how cheaply and well it performs in increasingly complex and essential tasks, it may end up in charge of key functions. Its very success creates trust. And that trust is earned on relatively simple and safe repetitive tasks.

As trust builds up, the critical risk is that we become so dependent on AI that the authorities cannot exercise control without it. Turning AI off may be impossible or very unsafe, especially since AI could optimise to become irreplaceable. Eventually, we risk becoming dependent on a system for critical analysis and decisions we don't entirely, or even partially, understand.

Six criteria for AI use in financial policy

These issues take us to six criteria for evaluating AI use in financial policy.

1. Data. Does an AI engine have enough data for learning, or are other factors materially impacting AI advice and decisions that might not be available in a training dataset?

2. Mutability. Is there a fixed set of immutable rules the AI must obey, or does the regulator update the rules in response to events?

3. Objectives. Can AI be given clear objectives and its actions monitored in light of those objectives, or are they unclear?

4. Authority. Would a human functionary have the authority to make decisions, does it require committee approval, or is a fully distributed decision-making process brought to bear on a problem?

5. Responsibility. Does private AI make it more difficult for the authorities to monitor misbehaviour and assign responsibility in cases of abuse? In particular, can responsibility for damages be clearly assigned to humans?

6. Consequences. Are the consequences of mistakes small, large but manageable, or catastrophic?

We can then apply these criteria to particular policy actions, as shown in the following table.

Conclusion

AI will be of considerable help to the financial authorities, but there is also significant risk of authorities losing control due to AI. The financial authorities will have to change how they operate if they wish to remain effective overseers of the financial system.

Many authorities will find that challenging. AI will require new ways of regulating, with different methodologies, human capital, and technology. The very high cost of AI and the oligopolistic nature of AI vendors present particular challenges. If then the authorities are reluctant and slow to engage with AI, they risk irrelevance.

However, when the authorities embrace AI, it should be of considerable benefit to their mission. The design and execution of micro-prudential regulations benefit because the large volume of data, relatively immutable rules, and clarity of objectives all contribute to AI's strength.

It is more challenging for macro. AI will help scan the system for vulnerabilities, evaluate the best responses to stress, and find optimal crisis interventions. However, it also carries with it the threats of AI hallucination and, hence, inappropriate policy responses.

Table 1. Particular regulatory tasks and AI consequences

Task	Data	Mutability	Objectives	Authority	Responsibility	Consequences
Fraud/Compliance Consumer protection	Ample	Very low	Clear	Single	Mostly clear	Small
Microprudential risk management Routine forecasting	Ample	Very low	Mostly clear	Single	Clear	Moderate
Criminality Terrorism	Limited	Very low	Mostly clear	Multiple	Moderate	Moderate
Nation state attacks	Limited	Full	Complex	Multiple & international	Moderate	Very severe
Resolution of small bank failure	Limited	Partial	Clear	Mostly single	Mostly clear	Moderate
Resolution of large bank failure Severe market turmoil	Rare	Full	Complex	Multiple	Often unclear	Severe
Globlal systemic crises	Very rare or not available	Full	Complex & conflicting	Multiple & international	Unclear even ex-post	Very severe

It will be essential to measure the accuracy of AI advice. It will be helpful if the authorities overcome their frequent reluctance to adopt consistent quantitative frameworks for measuring and reporting on the statistical accuracy of their data-based inputs and outputs.

The authorities need to be aware of AI benefits and threats and incorporate that awareness into the operational execution of the services they provide for society. ■

Jon Danielsson is Director of the Systemic Risk Centre at the London School of Economics and Political Science, and Andreas Uthemann is Principal Researcher at the Bank of Canada and Research Associate at the Systemic Risk Centre at the LSE

References
Acemoglu, A (2021), "Dangers of unregulated artificial intelligence", VoxEU.org, 23 November.
Borgonovi, F, F Calvino, C Criscuolo, J Nania, J Nitschke, L O'Kane, L Samek and H Seitz (2023), "Tracking the skills behind the AI boom: Evidence from 14 OECD countries", VoxEU.org, 20 October.
Danielsson, J and A Uthemann (2024a), "How AI can undermine financial stability", VoxEU.org, 22 January.
Danielsson, J and A Uthemann (2024b), "On the use of artificial intelligence in financial regulations and the impact on financial stability".
Norvig, P and S Russell (2021), Artificial Intelligence: A Modern Approach, Pearson.

Authors' note: Any opinions and conclusions expressed here are those of the authors and do not necessarily represent the views of the Bank of Canada. This article was originally published on VoxEU.org.

Bridging the innovation gap

Andrés Rodríguez-Pose and Zhuoying You show how artificial intelligence and robotics present a potential solution to the innovation gap problem for cities in China

In recent years, the potential of artificial intelligence and robotics to revolutionise production processes and stimulate innovation has attracted considerable attention (eg. Presidente and Calì 2022). Despite this, their capacity to address disparities in innovation remains underexplored. In our recent paper (Rodríguez-Pose and You 2024) we delve into the capacity of AI and robotics to not only spearhead technological innovation, but also to trigger innovation in the less developed and less innovative Chinese cities – those places where traditional innovation-boosting policies have often failed.

The geographical concentration of innovation

Science and technology (S&T) spending has traditionally been at the heart of innovation policies (Audretsch and Feldman 1996, Pavitt 1982). However, policies based on S&T have mostly favoured regions already at the forefront of technological progress and largely failed in less innovative areas, contributing to a massively uneven distribution of innovation.

Innovation has therefore become increasingly geographically concentrated in a few hubs endowed with substantial human and financial resources (Loumeau and Egger 2019). This concentration results in large disparities in innovation between more and less developed regions (Audretsch and Feldman 1996).

Less developed cities and regions, often positioned at or below the technological frontier, struggle to match the innovation outputs of their more advanced counterparts due to inadequate human capital and financial resources (Aghion *et al* 2019).

Policies aimed at promoting innovation, such as investments in S&T, have therefore contributed to a distribution of innovation across the world that is more geographically concentrated and uneven than that of virtually all other economic indicators, such as employment, income, investment, or productivity.

The potential of AI and robotics in China

AI and robotics – characterised by the implementation of technologies that enable machines to learn and make decisions without human intervention and by the use of programmable machines to perform tasks (Liu *et al* 2020), respectively – offer new avenues for stimulating innovation. It has been argued that investing in these technologies can significantly enhance productivity, economic growth, and innovation (eg. Acemoglu and Restrepo 2020). But are they also capable of bridging the innovation gap?

Less developed cities and regions, often positioned at or below the technological frontier, struggle to match the innovation outputs of their more advanced counterparts due to inadequate human capital and financial resources

China, a country determined to become a global AI leader by 2025 (Ciocca and Biancotti 2018), is actively resorting to AI and robotics as a means to spur economic activity. It pioneered a national campaign to attract AI talent (Zeng 2021) that strengthened its position in the field.

Until 2017, Europe had more than twice the number of AI researchers compared to China. However, from 2015 to 2020, China's AI skill penetration became 1.4 times the global average, trailing only behind India and the US (Lundvall and Rikap 2022). The campaign also aimed to promote AI for innovation in less innovative areas of the country.

Despite significant investments in science and technology, China's technological innovation was (and remains) highly concentrated around the big Eastern hubs of Shanghai, Guangzhou, and Beijing (Li 2009). A shift in focus to AI and robotics was considered a way to address this polarisation in innovation.

The approach to promoting AI adopted in China has been decentralised. Local governments have been allowed to tailor AI development strategies to local conditions (Zeng 2021). Individual cities and regions have developed their own AI strategies, leading to a geography of AI that is more variegated than that of innovation.

AI is concentrated in traditional innovation hubs, such as the Pearl River Delta and the Yangtze River Delta, but its use has also become widespread in many inland provinces (Figure 1).

Similarly, the adoption of industrial robots has grown significantly, particularly in the manufacturing sector. Since 2013, there has been a consistent increase in industrial robots.

Once again, the geography of robotics in China expands beyond the traditional industrial hubs, with robotics making substantial inroads in provinces such as Henan, Shandong, Fujian, Hebei, and Anhui, outside the main Chinese innovation centres (Figure 2).

AI and robotics and the innovation gap in China

Has this drive to promote AI and robotics delivered? And has it contributed to reduce the acute geographical innovation gap in China? We argue that, in the case of China, AI and robotics have acted as catalysts for technological innovation across cities and regions, particularly enhancing innovation in traditionally less innovative places.

After examining the impact of investment in AI and robotics on technological innovation across 270 Chinese cities from 2009 to 2019, we posit that these novel technologies not only have driven innovation directly but have also enhanced the returns on S&T investments. This dual effect is particularly pronounced in regions at or below the technological frontier, thus offering a promising strategy for reducing regional innovation disparities.

We show that investments in AI in Chinese cities correlate with an increase in innovation overall. Similarly, the density of industrial robots is positively associated with technological innovation, suggesting that cities with more robot installations tend to have higher patent intensities.

Furthermore, we also find that the impact of AI and robotics varies across the innovation spectrum. While these technologies enhance innovation in all cities, their effects tend to be more substantial in traditionally less innovative regions. This finding suggests that AI and robotics can help bridge the innovation gap by providing greater relative benefits to cities below the technological frontier.

Figure 1. AI development in mainland China, 2008-2018

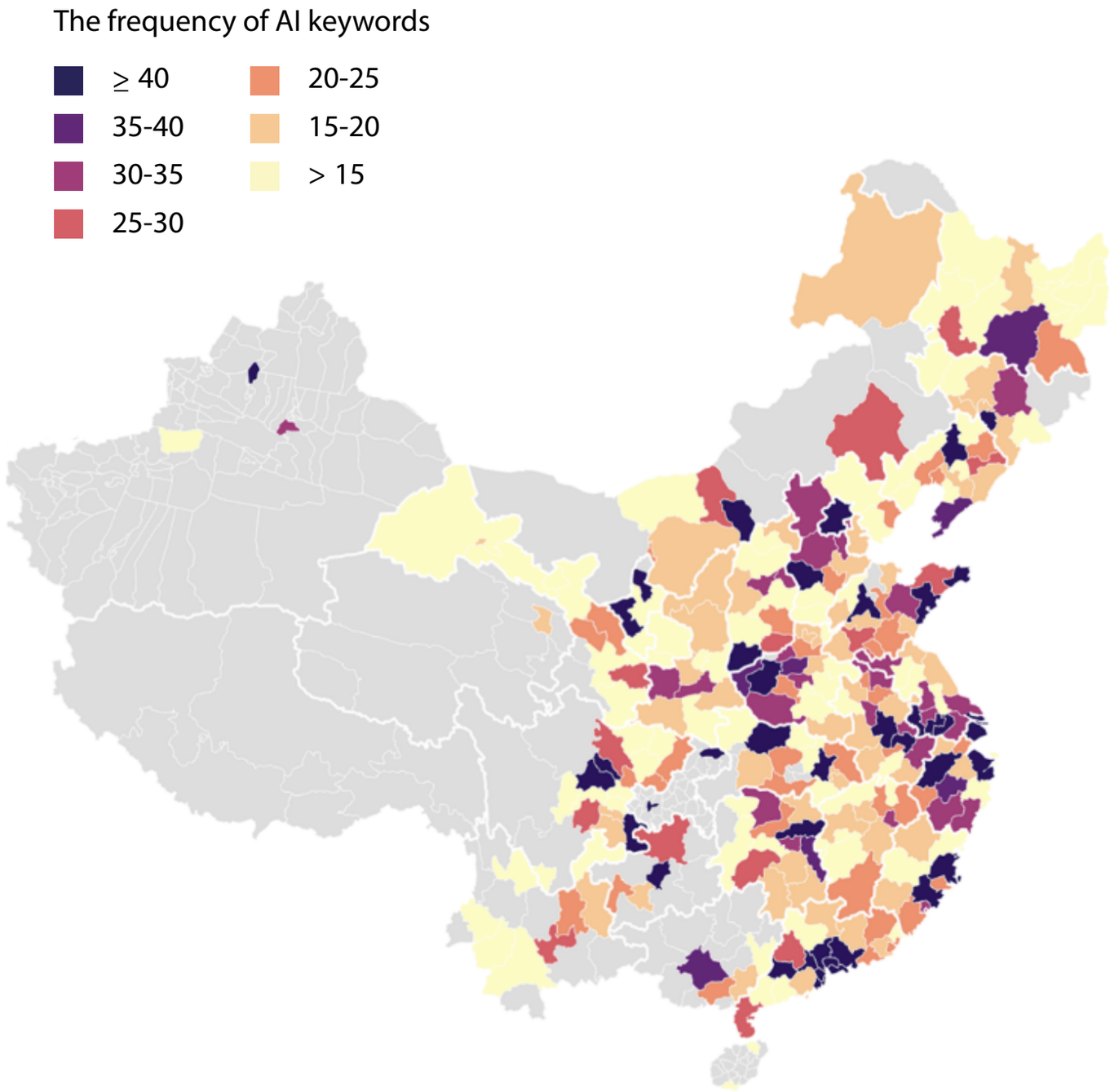
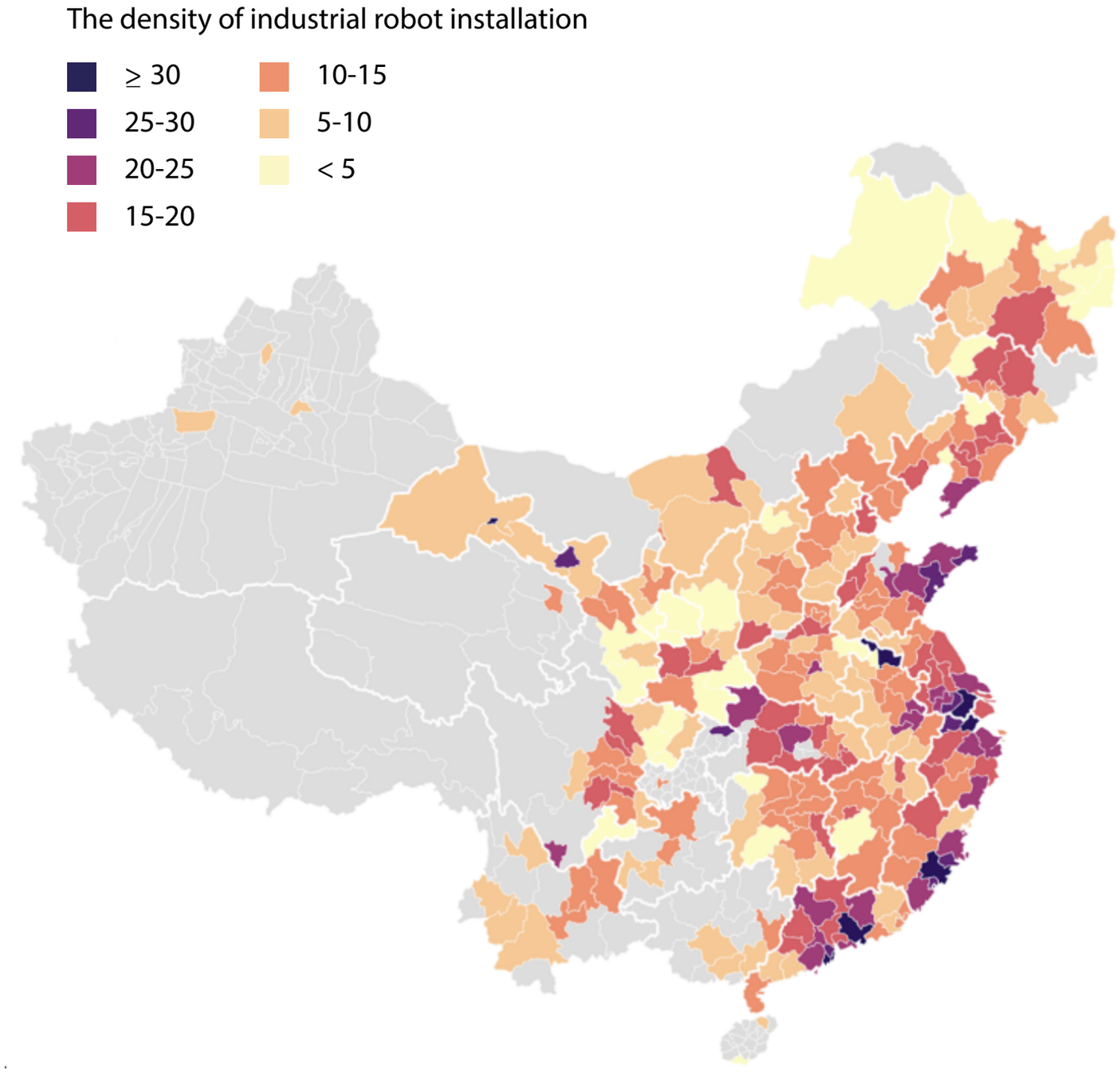


Figure 2. Geographical distribution of the average density of industrial robots in mainland China, 2008-2018



Our study also highlights the moderating roles of AI and robotics on the relationship between S&T expenditure and technological innovation. AI and robotics amplify the positive effects of S&T investments on innovation. This moderating effect is particularly strong in less innovative cities, further underscoring the potential of these technologies to reduce regional innovation disparities.

Policy implications

The results of our study point to the need to include AI and robotics as integral components of innovation strategies aimed at fostering technological progress across all regions. Subnational governments, particularly in less developed areas, need to prioritise the development and adoption of these technologies as a more effective way than old-style innovation strategies to enhance innovation across the board.

The findings underscore the limitations of traditional S&T investments in driving innovation in less developed regions. While S&T expenditure remains important as a tool to promote innovation, its returns are greatly enhanced by initiatives that promote the deployment of AI and robotics.

This integrated approach can boost the returns on S&T investments and ensure that the benefits of technological progress are more inclusive and more evenly distributed across people and territories.

By focusing on AI and robotics, governments can develop targeted strategies that not only increase overall innovation but also specifically support less innovative cities. This targeted support can help mitigate the risks associated with regional disparities in innovation, such as economic stagnation, social inequality, and political discontent.

Overall, AI and robotics have demonstrated themselves in China to be powerful drivers of technological innovation and effective mechanisms for reducing regional disparities in innovation performance. Integrating these technologies into innovation strategies can foster more inclusive and sustainable economic growth, ensuring, as Baldwin (2019) has stressed, that changes towards AI and robotics *"will give more 'head' to people with big hearts, but no extra heart to people with big heads."*

Hence, innovation based on AI and robotics will encourage the use of new types of skills that are more widely distributed across all territories, meaning that the benefits of technological progress, as shown in the case of China, are shared more widely, ultimately leading to greater economic and social wellbeing. ■

Andrés Rodríguez-Pose is a Professor of Economic Geography at the London School of Economics, and Zhuoying You is a Research Fellow at Bocconi University and a PhD student at Ku Leuven

References

Acemoglu, D and P Restrepo (2020), "Robots and jobs: Evidence from US labor markets", Journal of Political Economy 128(6): 2188-2244.

Aghion, P, BF Jones and CI Jones (2019), "Artificial Intelligence and Economic Growth", in The Economics of Artificial Intelligence: An Agenda, NBER.

Audretsch, DB and MP Feldman (1996), "R&D spillovers and the geography of innovation and production", The American Economic Review 86(3): 630-640.

Baldwin, R (2019), "Globalisation, automation and the history of work: Looking back to understand the future", VoxEU.org, 31 January.

Ciocca, P and C Biancotti (2018), "The economic geography of innovation", VoxEU.org, 23 October.

Liu J, H Chang, JYL Forrest and B Yang (2020), "Influence of artificial intelligence on technological innovation: Evidence from the panel data of China's manufacturing sectors", Technological Forecasting and Social Change 158: 120142.

Loumeau, N and P Egger (2019), "Data superpowers in the age of AI: A research agenda", VoxEU.org, 16 January.

Lundvall, BÅ and C Rikap (2022), "China's catching-up in artificial intelligence seen as a co-evolution of corporate and national innovation systems", Research Policy 51(1).

Pavitt, K (1982), "R&D, patenting and innovative activities A statistical exploration", Research Policy 11: 33–51.

Presidente, G and M Calì 2022, "Robots for economic development", VoxEU.org, 13 March.

Rodríguez-Pose, A and Z You (2024), "Bridging the Innovation Gap. AI and Robotics as Drivers of China's Urban Innovation", CEPR Discussion Paper DP19010.

Zeng, J (2021), "China's Artificial Intelligence Innovation: A Top-Down National Command Approach?" Global Policy 12(3): 399–409.

This article was originally published on VoxEU.org.

Broader border taxes in the EU

The EU needs new resources to fund its budget. Pascal Saint-Amans considers the problems of tax leakage and discusses how the EU can access new funds

Executive summary

There is widespread agreement on the need for new resources to fund the European Union's budget in order to meet increasing spending demands, not least repayment of debt incurred as part of the EU's post-pandemic economic recovery.

In particular it is seen as desirable that the EU should have 'own' resources, or reliable ongoing revenue streams. But there is little agreement on what new own resources could consist of.

Limited reform so far has led to the introduction of a levy paid by EU members depending on plastic packaging waste generated in their territory and not recycled. Meanwhile, the European Commission has proposed resources for the EU budget from emissions trading revenues, and from levies collected under the EU carbon border adjustment mechanism (CBAM). There proposals are pragmatic and move in the right direction, but do not go far enough.

The debate about own resources should focus on whether the EU will be able to build genuine own resources based on common tax policies. The EU suffers from 'tax leakage' in which profits are shifted from high-tax to low-tax EU countries, and from there onto no or low-tax non-EU jurisdictions, often without the application of withholding taxes.

It may not be too much of a stretch to compare this situation of tax leakage with the situation addressed by CBAM – a quasi-tax at the border. So far, an opportunity for what could be seen as a tax at the border of the internal market, aiming to protect the market from harmful competition, may have been missed.

Such a tax could reflect the undertaxed profit rule agreed as part of the international deal on the corporate minimum tax. Focusing on protecting the revenues of EU members by common tax borders could offer scope for new own resources.

In proposing new own resources, it was wise for the Commission not to go back to the idea of a European digital services tax

1 Introduction

While the budgets of its member countries are funded primarily by taxes approved by their parliaments, the funding of the European Union budget is much more complex, reflecting in part the ambiguous nature of the EU. The Treaty on the Functioning of the European Union provides that *"without prejudice to other revenue, the budget shall be financed wholly from own resources"* (Article 311).

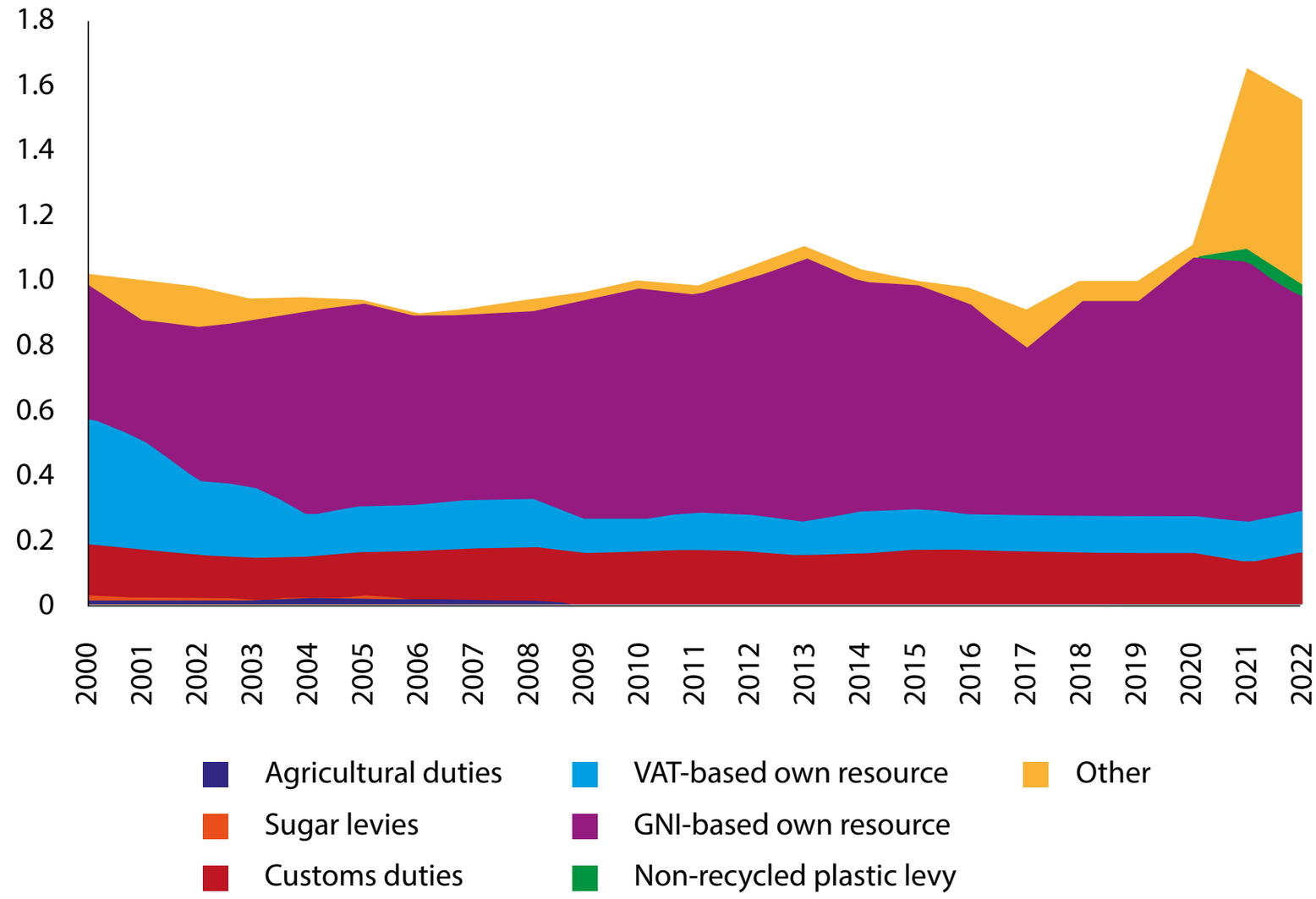
'Own resources' are ongoing streams of revenue mainly collected by member countries and passed onto the EU, with EU governments responsible for deciding by unanimity what these resources should be.

Consequently, revenues for the EU budget comprise a mixture of so-called 'genuine' own resources (levies that belong to the EU, such as custom duties) and other contributions from member countries, usually based on statistical aggregates, like value-added tax and gross national income (GNI)[1]. The latter have significantly increased over time and represented almost three-quarters (72 percent) of the EU budget in 2020 (Figure 1; for pre-2000, see European Commission, 2021).

'Genuine' own resources now account for a small portion of the overall budget, which is a good illustration of why the EU needs a new approach to funding its activities. The adoption in December 2020 of the EU's post-pandemic recovery programme NextGenerationEU (NGEU), allowing debt financing for the first time in European Union history, was an opportunity to reopen the debate on own resources.

Also in December 2020, an EU Inter-Institutional Agreement[2] on the EU budget provided that *"The repayment of the principal of such funds to be used for expenditure under the European Union Recovery Instrument and the related interest due will have to be financed by the general budget of the Union, including by sufficient proceeds from new own resources introduced after 2021."*

Figure 1. Sources of financing for the EU budget (% of EU GNI)



Note: Borrowing to finance NextGenerationEU is included for 2021 and 2022, reflected in the large increase in the 'Other' category. This category also includes smaller sources of revenue such as fines, surplus from the previous year and revenue from EU policies.
Source: Bruegel based on European Commission.

On 14 December 2020, the Council adopted a new own resource – on non-recycled plastic waste – for the first time in years, as if a new era was beginning. The European Commission then proposed, in December 2021, three new own resources:

1. Contributions from the EU emissions trading system (ETS);

2. Contributions from the carbon border adjustment mechanism (CBAM), which is designed to equalise the carbon cost of certain goods, whether produced inside the EU or imported;

3. A share of the revenue expected from the application of an Organisation for Economic Co-operation and Development agreement on the taxation of the residual profits of large multinational companies.

Though endorsed by the European Parliament, this proposal failed to trigger much discussion among EU countries. In June 2023[3], the Commission tabled a revised proposal for 'An adjusted package for the next generation of own resources' (European Commission, 2023a).

As well as setting out new ideas for revenues, the proposal called on EU countries *"to accelerate the negotiations"*, with the objective of getting a unanimous decision by 1 July 2025 for the introduction of the new own resources in January 2026.

The question is what such new own resources should be, and particularly, whether it is possible to identify additional 'genuine' own resources with a European character – as opposed to statistically-based contributions such as VAT and GNI shares, which encourage thinking about the EU budget in terms of net balances received or contributed by member states (Fuest and Pisani-Ferry, 2020).

Before assessing the Commission's proposal, one can only note the lack of appetite among EU countries to move this debate forward. In February 2024, an agreement among EU countries on a midterm review of the EU's seven-year budget (the Multiannual Financial Framework, MFF) gave only a cursory mention to new own resources, with no update on the position of member states on the Commission's proposed package[4].

The purpose of this paper is two-fold: first, to review the Commission's revised proposal in the context of the new financing challenges resulting from NGEU, and second to contribute some new ideas for 'genuine' own resources.

We find that the Commission has taken a pragmatic approach aimed at speeding up the negotiations, rather than revisiting the nature of the own resources. As to new 'own resources', we offer some recommendations that draw on recent progress on international taxation.

2 Background: the impact of NGEU on the EU budget and its financing

In adopting NGEU, EU countries called for a revision and expansion of the EU's own resources, to finance the borrowing costs for the approximately €421 billion in NGEU grants and to reduce reliance on the GNI-based own resource (Council, 2020).

They also agreed to raise the maximum potential amount of their annual contributions to the EU budget by an additional 0.6 percent of GNI, expressly for the purpose of servicing NGEU interest and debt.

For the first time in decades, a new resource based on non-recycled plastic waste, was adopted and entered into force in 2021[5]. However, this new resource is relatively small money compared to the debt service required for NGEU, contributing only about 3 percent of total EU revenues (European Commission, 2023b). Moreover, it is not an EU levy, but is based on contributions from members, reflecting their levels of non-recycled plastic packaging waste.

With the first repayments of NGEU borrowing due in 2028, a timeline was agreed to revisit this issue and find new resources. The December 2020 Inter-Institutional Agreement provided that *"the expenditure from the Union budget related to the repayment of the European Union Recovery Instrument should not lead to an undue reduction in programme expenditure or investment instruments […] It is also desirable to mitigate the increases in the GNI-based own resources for the member states."*

In absence of an agreement on additional own resources, the burden of financing this debt will fall directly on EU countries through the increased ceiling, leading to an even greater reliance on the GNI-based contribution to the EU budget. It could also translate into cuts in current programmes to make room for debt service.

The Commission has already had to propose changes to the current MFF to respond to the much higher-than-expected interest rates on EU borrowing costs (Figure 2)[6]. Hence, the debate on increasing own resources is critical to EU-funded policies.

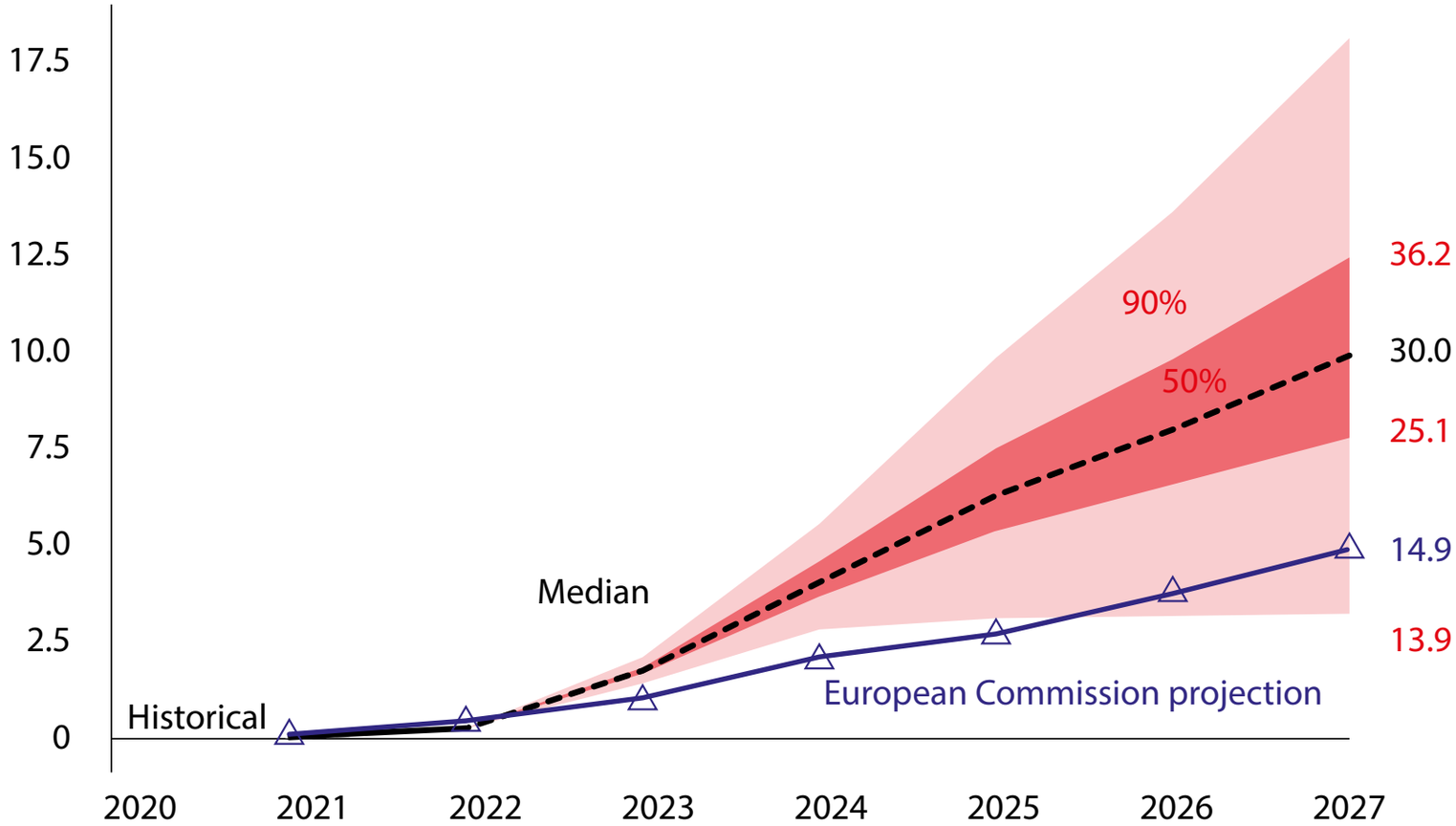
3 The European Commission proposals on own resources

3.1 The initial Commission proposal of December 2021

A first package of three new own resources was proposed by the European Commission in December 2021 (European Commission, 2021). This package introduced what could be considered 'genuine' own resources, with 25 percent of the emissions trading system (ETS) revenues and 75 percent of carbon border adjustment mechanism (CBAM) revenues going to the EU budget.

CBAM, which would levy a border toll (the Commission refrains from using the word 'tax') on certain carbon-intensive imports, entered into force in October 2023 and will become a definitive system in 2026. It will not be a significant revenue raiser.

Figure 2. Projected annual and total interest costs borne by the EU (in € billions)



Source: Claeys et al (2023a).

However, CBAM shares some similarities with custom duties as it is an EU levy at the point of entry of goods into Union territory. The ETS, meanwhile is clearly an EU-wide policy and, even though revenue can be tracked nationally, with most of the revenue allocated to each member state, the over- all approach remains an EU approach.

The Commission's proposal was therefore to rebalance own resources away from contributions from member states, whether based on VAT or GNI, and towards EU policy-based resources.

The third element of the December 2021 proposal related to the potential revenue generated by the agreement reached at the OECD on the reallocation of taxing rights among more than 140 countries to some of the profits of the world's largest and most profitable companies.

This was the culmination of an issue debated for more than a decade in the context of the idea that market jurisdictions were not receiving fair shares of revenues from the world's biggest digital companies. While the OECD negotiations on a global approach progressed slowly, some EU members, led by France, pushed for the introduction of an EU digital services tax (DST), which failed to obtain unanimity in 2019.

Some members introduced domestic DSTs from 2018 to 2021, while the OECD was still negotiating a multilateral solution within its Inclusive Framework on Base Erosion and Profit Shifting (bringing together 140+ countries[7]). In July 2020, EU countries agreed that, in the case of a failure of the OECD negotiations, a tax on digital companies would be agreed and would be an own resource.

In 2021, the Biden Administration rebooted the negotiations, which resulted in a two-pillar agreement at the OECD. Pillar 1 provides that a quarter of the rent (defined as the profit above a 10 percent profit margin on sales) earned

by the largest and most profitable multinationals (above €20 billion in revenues and 10 percent profitability) would be allocated to market countries (countries where the goods or services are sold) using a formula based on sales, whether or not the company is physically present in the country.

This is a significant departure from traditional transfer pricing rules and a move towards what economists call 'destination'. Interestingly, Pillar 1 would not be limited to tech companies, as was initially asked for by most European countries.

The Commission's December 2021 proposal proposed that 15 percent of the revenue accruing to EU countries from the Pillar 1 taxing rights reallocation would become an own resource.

The rationale behind that reallocation seemed to be more reflective of a political mood ('taxing the digital economy', or *"taxing the GAFA"* as the French finance minister repeated, in Council throughout 2018[8], even though the scope of the OECD agreement had already broadened) than about building a genuine own resource, as could have been the case with the initially planned DST.

The proposed rate of 15 percent was hard to explain (at the global level, the reallocation of profit for taxing has been projected to be in the range of €150 billion).

Pillar 1, however, is subject to the development of a multilateral convention, which would require ratification by all signatories, including the United States, with a two-thirds Senate majority. The development of the multilateral convention is running late, with a new deadline in June 2024 for signing, and very uncertain prospects for ratification.

Interestingly, the Commission did not propose anything on own resources in relation to Pillar 2 of the OECD agreement, which provides for the establishment of a global minimum tax of 15 percent on the profits of multinationals with revenue above €750 million.

The expected additional tax revenues globally from this pillar are higher (in the range of €200 billion), with a complex three-tier mechanism that might be interesting from an EU own-resource perspective. We return to this issue in section 4.

3.2 The revised European Commission package

The Commission's June 2023 *"adjusted package for the next generation of own resources"* (European Commission, 2023a) added to the December 2021 plan in three ways: an increased slice of ETS revenues for the EU budget, a change to the date when some supplementary ETS revenues would start to flow into the budget, and a proposed new own resource related to corporate profits.

The EU budget share of ETS proceeds would increase from 25 percent to 30 percent, with no change related to CBAM. As the carbon price has increased, this would still leave more revenue to member states (€46 billion per year from 2028) while securing an annual €19 billion for the EU budget. CBAM, meanwhile, would be expected to generate €1.5 billion as of 2028 for the EU budget.

The June 2023 proposal formally maintains the 15 percent contribution deriving from the OECD deal, despite that deal's uncertain prospects of implementation.

In addition, the Commission proposed a new statistical-based resource on company profits. This was described as a *"national contribution calculated on the basis of statistics from national accounts under the European system of accounts"*, a proxy for company profits.

It would be less of a genuine own resource than CBAM and ETS contributions. It is also a pragmatic reflection of the fact that an EU harmonised tax on company profits is still a distant prospect.

The Commission estimates the base of corporate profits could reach €3 trillion and trigger revenues from €3 billion to €16 billion per year, with a call rate of 0.1 percent to 0.5 percent. The proposed resource has merit in that it would increase the absolute contribution of the largest and most advanced EU members (Germany, France), while having the largest effects in terms of GNI on smaller members that have benefitted from decades of corporate profit shifting (predominantly Ireland and Luxembourg; Figure 3).

While the ETS own resource could disproportionately penalise some Eastern European countries (because of their shares of electricity generation from fossil fuels), this new resource would balance the contribution back to the 'west'.

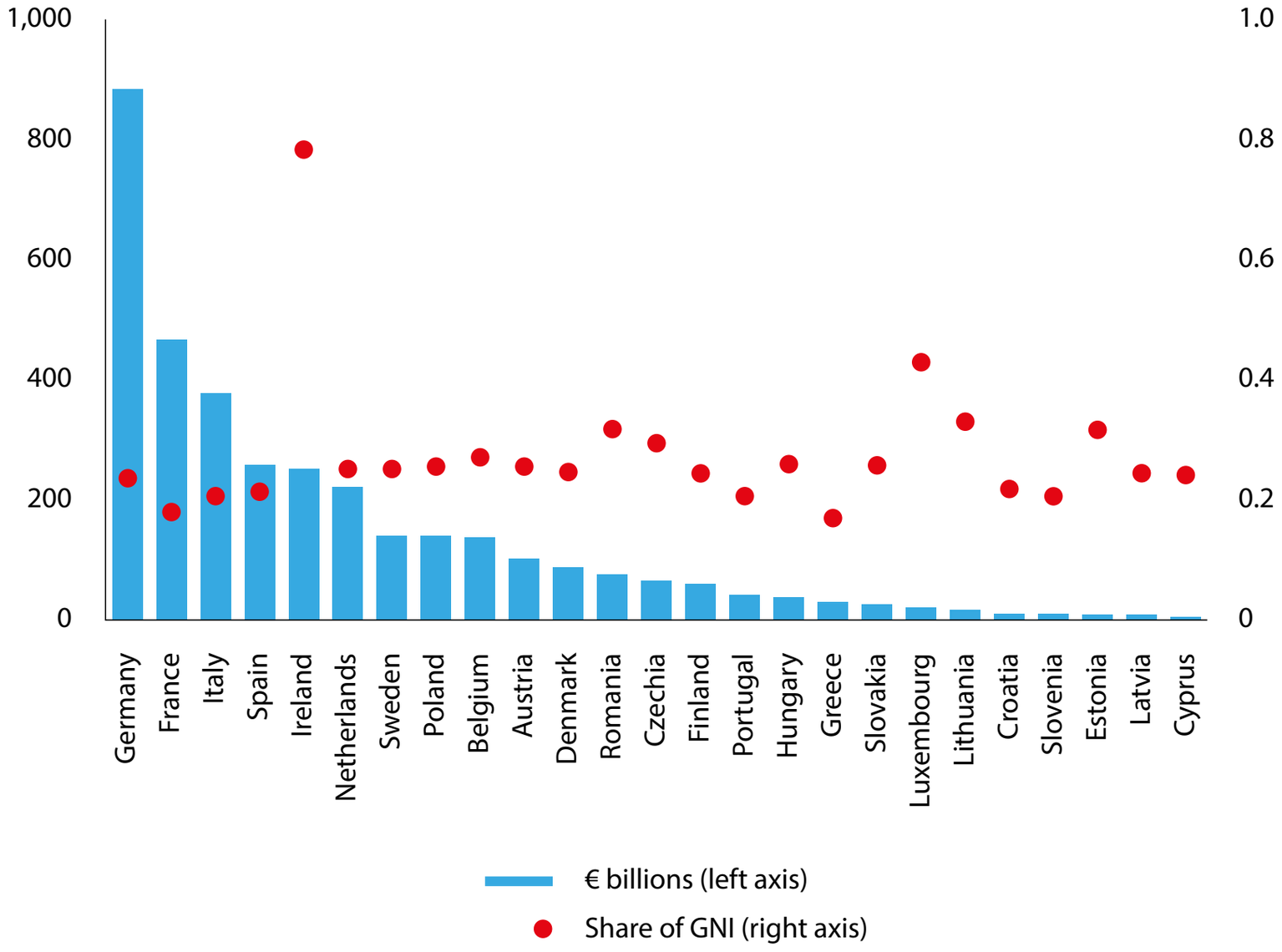
Expected amounts from the resource would be broadly comparable with €19 billion from ETS/CBAM. Finally, this resource is presented as temporary: it would be superseded by a share of taxes on corporate profits based on a common European tax base for corporations, which the European Commission is pushing for[9].

The new statistical based resource on company profits is one of eight potential new own resources, most of them previously mentioned by the European Parliament[10] and scanned by the Commission. This 'scanning' exercise against three selection criteria (revenue potential, simplicity in terms of compliance and administrative burden and fast mobilisation) was brief and could look like lip service to seven potential new own resources[11].

3.3 An evaluation of the Commission proposal

The Commission's revised proposal is pragmatic and moves in the right direction, but does not go far enough.

Figure 3. Gross operating surplus for corporations by country in € billions and as a share of GNI, 2021



Note: Gross operating surplus data is unavailable for Bulgaria and Malta.
Source: Bruegel based on AMECO.

First, the Commission rightfully puts the emphasis on ETS revenue, increasing the allocation to the EU budget from 25 percent to 30 percent. As demonstrated by Fuest and Pisani-Ferry (2020), ETS revenues fit best the criteria for EU budget resources: *"carbon emissions do not primarily cause damage where they occur […] additional emissions in a particular member state should be regarded as negative externalities in other member states […] emission reduction objectives are set at the EU level."* As the authors concluded, *"ETS allowances are not that different from custom duties"*, making them genuine own resources.

In addition, potentially big revenue can be derived from the ETS. Since the Commission's December 2021 proposal, the EU carbon price has increased significantly. The price per tonne of CO_2 was until recently above €90 and likely to rise further in the mid-term, much higher than the price assumption of €55 for the period 2026-2030, as presented in the Commission's legislative proposals related to the 2030 emissions reduction target of minus 55 percent compared to 1990.

Not only has the price increased, but the scope is also broadening, with a second ETS (ETS 2), which covers buildings and road transport, becoming operational in 2027. This dynamic allows the Commission to increase the EU share of ETS revenue from 25 percent to 30 percent, while leaving net increased revenue to member states at €46 billion per year from 2028.

The only downside to this approach may be of a political nature. Carbon pricing remains unpopular when directly borne by households, which is why the allocation of ETS 2 revenue was politically committed to social compensation and redistribution, rather than financing EU own resources.

Allocating part of ETS 2 to policies not directly related to greening the economy represents some political risk that the Commission and other EU institutions must be mindful of.

It will be important for the Commission to assure people that money collected from ETS 2 will be spent on climate objectives to alleviate the burden on households of the energy transition. It may make political sense to allocate the revenue to funding green policies and not divert it to other actions.

Second, the Commission is right to maintain its proposal to allocate 75 percent of CBAM revenues to own resources. Logically all resources arising from CBAM should in fact be allocated to the EU budget, but it was politically impossible for the Commission to not leave 25 percent of CBAM revenue to member states.

The Commission estimates that €1.5 billion from CBAM would accrue annually to the EU budget from 2028. This however would depend on the reaction from the main impacted EU partners. Some may adopt pricing policies, which would reduce CBAM revenues.

This would remain good news as the overall policy objective is about emission reductions and not revenue. In any case, CBAM is not where the bulk of revenue is. In the long run, when decarbonisation occurs, the fundamental question of finding a more stable revenue base will arise.

Third, the new statistical based resource on a proxy for corporate profits can be considered a smart move. It does not rely on the fast adoption of BEFIT, the latest Commission proposal on harmonising corporate income tax in the EU. Whatever the merits of BEFIT, the prospect of adoption is extremely low given how difficult the corporate income tax debate has been in Europe for decades.

More importantly, it is wise for the Commission not to go back to the idea of a European digital services tax (DST) as a substitute for the OECD deal, notwithstanding that, in July 2020, EU governments recommended the adoption of a European DST in case of OECD negotiation failure[12].

DSTs are taxes on transactions, which would be a proxy for EU countries to tax the profits of tech companies that leave very little profits on their territories because of aggressive tax planning that takes advantage of the inadequacy of existing international tax rules.

DSTs may seem like a good idea but to the extent that they are taxes on gross income, they would create double taxation, would be borne by consumers more than companies, and would likely generate trade tensions with the US.

For all these reasons, they are divisive and an EU DST is unlikely to garner the necessary unanimity to be adopted. By not mentioning this option, the Commission risks of being criticised by DST advocates (France, Italy, Spain), but spares itself a difficult and unpromising negotiation within the EU and tensions with the US.

While it adds little to the December 2021 proposal, the adjusted Commission proposal can be defended as a pragmatic move to facilitate a discussion of own resources with member states within a constrained calendar.

European elections are approaching, and a Multilateral Financial Framework proposal will have to be tabled by 2025, while the strategic agenda will have to be approved in 2024. With the first repayments of NGEU debt in 2028, EU institutions are running out of time.

4 EU taxation ideas worth of exploration

Beyond the urgent need to agree on a package to pay back NGEU, the debate about own resources should focus on whether the EU will be able to build genuine own resources based on common tax policies.

This is a more fundamental debate, raising the question of the nature of the European Union, and the debate between those seeing it as a confederation of sovereign states and those believing in its federal destiny.

For the time being, the Treaties reflect the situation in which tax remains at the core of national sovereignty and consent to tax, one of the fundamental human rights, a pure national exercise. The EU's unanimity rule on tax-related decisions is the basic translation of this stubborn reality, in which it is unlikely that the European Parliament would be considered as sufficiently legitimate to consent to tax.

The decision-making difficulty resulting from unanimity is increased by the interests of member countries not being aligned. Large EU countries and other high-tax countries have an interest in establishing a common tax base, which would limit tax leakage, for both individuals and companies, even at the cost of limiting their sovereignty.

On the contrary, to attract investment, most of the small members have to compensate for the sizes of their economies, or their peripherical geography, with lower taxes, in particular on mobile factors, including corporate profits or high-income earners. Diverging interests and unanimity are why the EU is in a stalemate situation.

It could be observed that previous EU enlargement to low-tax countries, such as Malta and Cyprus, without changing the decision-making rules, or asking these countries to change their laws before joining the Union, has just made the issue more intractable.

Overall, this means that the prospect of genuine own resources deriving from harmonised taxes remains remote, as unanimity is unlikely to be reached any time soon.

Moving from unanimity to qualified majority voting in tax decisions, which would require Treaty changes, can only reflect agreement on the nature of the institutions. This does not seem feasible, especially at a time of rising populism when national sovereignty is increasingly emphasised.

The current situation, within the EU, reflects a tax anomaly. To avoid leakage, national tax systems provide for tax borders: residents are taxed on their worldwide incomes and countries tax non-residents via withholding taxes on the incomes they derive from those countries.

In short, to avoid leakage, outbound payments (including dividends, interest, royalties and salaries) are subject to withholding taxes, while anti-abuse rules ensure that residents don't shift profits abroad. With globalisation, the robustness of these rules has been tested.

International efforts driven by the G20 and the OECD since 2008, to introduce a tax regulation of globalisation have aimed to restore these instruments in a coordinated manner, rather than a situation of pure protectionist unilateral tax measures.

The EU offers however a unique environment in which countries have lost their ability to apply taxes at the internal borders (within the internal market) following a set of EU Court of Justice decisions starting in the 1990s, which have found anti-abuse rules to be discriminatory.

As a result, high-tax countries lost their ability to limit the risk of profit shifting within the EU, where there are low-tax countries. Low-tax countries, as part of their 'tax offer' to foreign investors, removed their own external borders, when they had such measures.

For instance, they used to offer hybrid instruments and entities allowing companies to book profits generated in Europe in no-tax jurisdictions like Bermuda or Cayman Islands. They also usually offer no withholding taxes and no controlled foreign company regimes, providing tax planners with easy opportunities to shift profits outside the EU at a very low tax cost.

In short, the EU offers the possibility to do business in a high-tax country, shift the profits to an EU low-tax country, without any toll, and then shift the profits to a low- or no-tax country outside the EU, still without any toll.

In parallel with OECD progress on fighting base erosion and profit shifting (BEPS), the European Union has adopted an unprecedented number of tax directives, with various directives on administrative cooperation (which deal with exchange of information between tax authorities[13]) and two directives on anti-abuse rules.

These EU instruments implement rules adopted at the OECD by the Inclusive Framework. These directives bring more coherence to the system by increasing cooperation between tax authorities, and also by helping members to protect their tax base.

The most recent example is the directive translating into EU law the OECD Pillar 2 agreement establishing a global minimum tax, which EU countries should have implemented by the end of 2023 for entry into force in 2024 (Directive (EU) 2022/2523).

Preceded by global agreements, facilitating a worldwide level-playing field, these EU instruments show that EU members can overcome the constraints of unanimity. The EU has even been able to go beyond OECD efforts with a directive mandating publication of the country-by-country reports of multinationals (the issue was considered as non-tax and therefore was ruled with qualified majority).

Some of this recent progress could facilitate a move towards genuine EU own resources. For instance, the 15 percent global minimum tax could have offered an opportunity to mutualise some resources at the EU level as a genuine own resource.

The minimum tax rules provide for a complex three-tier mechanism to ensure that profits of multinationals, where initially taxed below an effective 15 percent in a jurisdiction, will finally be taxed at 15 percent.

First, the country of residence of the multinational will include any such low-taxed income in its tax base and will collect the additional tax (the Income Inclusion Rule, IIR).

If a country does not exercise that taxing right, countries where the company sells its goods or services will have a right to collect the additional tax (the difference between the effective tax rate in any jurisdiction where the company operates and the 15 percent effective rate), through what is known as an undertaxed profit rule (UTPR).

In addition to the IIR and the UTPR, countries where profits are taxed below 15 percent (either because it is a no-tax country, or because it offers a tax holiday, as can be the case in developing countries) can decide to take the difference themselves through a domestic minimum top-up tax (DMTT).

While the nature of the IIR and the DMTT seems quite national (a country will tax the profit of its own companies abroad), the nature of the UTPR is less domestic. Concretely, if a US or Chinese company (these two countries have not so far moved to implementing the minimum tax rules) operates on the European market with under-taxed profit in a low-tax jurisdiction (say the Cayman Islands or Bermuda where there is presently no corporate income tax for the time being), European countries will be entitled to collect the tax.

Though the collection of the tax will be national, the right to tax, which will depend on allocation rules, seems logically to belong to the internal market and the EU as a whole. It may not be too much of a stretch to compare this with the CBAM, a quasi-tax at the border.

In that sense, it is surprising that the European Commission did not examine this option, and favoured, in its initial proposal, a share of the allocation of taxing rights resulting from the other OECD Pillar (Pillar 1). It is true that the distribution of the global additional annual €150 billion to €190 billion of revenue remains unclear and that, in the long term, this revenue may dry up with tax competition being neutralised.

Still, an opportunity to push for what could be seen as a tax at the border of the internal market, aiming to protect the market from harmful competition, may have been missed.

In theory, one could argue that the DMTT is a way for low-tax countries to put an end to their aggressive tax offers, which allowed excess profit to be allocated to their territory, in a way that is not commensurate to activity deployed there.

The OECD estimates that a significant part of the additional revenue will be captured, at least in the short run, through DMTTs (Hugger *et al* 2024). This additional revenue could in theory be mutualised, even though, focussing on UTPR, as an external tax border, seems like a more realistic and practical way. It is also consistent with the fundamental structure of tax systems.

More broadly, exploring how other external tax borders of the EU could be restored could be a way to move towards genuine new own resources. For instance, in the area of personal income tax, establishing a common exit tax on EU countries' residents moving abroad to avoid paying capital gain taxes could serve the purpose of protecting EU countries' tax bases and developing a new own resource.

This could also be considered in the field of wealth taxation or inheritance duties, even though it must be recognised that the lack of harmonised approaches to these taxes by EU countries does not help define a common external policy.

Fundamentally, however, the idea of establishing external tax borders, to limit the risk of the delocalisation of the tax base (through exit taxes on unrealised capital gains for instance), could be further explored and may be a way to move forward the tax conversation in Europe.

Rather than harmonising taxes, which proves difficult, focusing on protecting the revenues of EU members by common borders may unleash some potential.

5 Conclusion

The European Commission's June 2023 adjusted proposal for own resources was motivated by the need to ensure a swift move towards adopting additional own resources to fund NGEU. The agreement to start debt financing the EU included an agreement to adopt new own resources.

Failure to move forward would jeopardise the ability of the EU to keep funding its existing projects, especially at a time when interest rate increases will make the repayment of both capital and interest heavier.

Time is running out, and the Commission proposed an adjusted mechanism that is pragmatic and rebalances the burden to make it more acceptable to Eastern European countries. It is a good move, even though no conversation has yet seriously taken place in the Council.

More importantly, the real debate on how to establish genuine own resources still needs to take place. A move to ETS and CBAM revenue to be mutualised is good and would give more weight to real own resources, aligned with EU policy objectives.

More needs to be done and recent international tax progress are a unique opportunity for the EU to explore how it could bring more consistency to tax systems in the EU while developing own resources. ■

Pascal Saint-Amans is a Non-Resident Fellow at Bruegel

Endnotes

1. The VAT and the GNI resources, based on statistical aggregates, are paid by members, which consider them to be national contributions, rather than resources owned by the EU.

2. In December 2020, the European Parliament, the Council and the Commission adopted an agreement on budgetary discipline, cooperation on budgetary matters, sound financial management and new own resources; see https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.LI.2020.433.01.0028.01.ENG&toc=OJ%3AL%3A2020%3A433I%3ATOC.

3. This was brought forward: a decision on a second basket of own resources was initially envisaged for June 2024.

4. See European Council notice of 1 February 2024, 'Special European Council, 1 February 2024'.

5. This own resource is proportional to the quantity of plastic packaging waste that is not recycled. EU countries contribute €0.80 per kilogramme of plastic packaging waste that is generated in their territory and not recycled.

6. See European Commission press release of 20 June 2023, 'EU budget: Commission proposes to reinforce long-term EU budget to face most urgent challenges'.

7. See https://www.oecd.org/tax/beps/about/.

8. See for example Reuters, '"Enough excuses!" France's Le Maire grows impatient over GAFA tax', 18 October 2018. GAFA refers to Google, Amazon, Facebook and Apple.

9. The Business in Europe: Framework for Income Taxation (BEFIT) proposal, which aims to reboot negotiations on a common EU approach to taxation of corporate profits. See European Commission press release of 12 September 2023, 'Taxation: new proposals to simplify tax rules and reduce compliance costs for cross-border businesses'.

10. See the European Parliament resolution of 10 May 2023, 'Own resources: A new start for EU finances. A new start for Europe'.

11. The examined additional seven own resources were: (i) corporate tax BEFIT (no fast mobilisation planned), (ii) a financial transaction tax (same), (iii) an EU fair border mechanism aimed at fighting social dumping (modestly meeting the criteria), (iv) a tax on crypto-currencies (same), (v) a statistical resource based on gender pay gap (no fast mobilisation), (vi) a statistical resource on food waste, and (vii) a statistical resource based on e-waste, the latter two with

a good prospect of fast mobilisation but only adequate simplicity and revenue potential.
12. See https://www.consilium.europa.eu/media/45109/210720-euco-final-conclusions-en.pdf.
13. See https://taxation-customs.ec.europa.eu/taxation-1/tax-co-operation-and-control/administrative-co-operation-and-mutual-assistance/enhanced-administrative-cooperation-field-direct-taxation_en.

References

Claeys, G, C McCaffrey and L Welslau (2023a) 'The rising cost of European Union borrowing and what to do about it', Policy Brief 12/2023, Bruegel.
Claeys, G, C McCaffrey and L Welslau (2023b) 'An estimate of the European Union's long-term borrowing cost bill' Briefing for the European Parliament BUDG Committee.
Council (2020) 'Council Decision 2020/2053 of 14 December 2020 on the system of own resources of the European Union', Council of the EU.
European Commission (2021) 'The evolving nature of the EU budget', EU Budget Policy Brief 1.
European Commission (2023a) 'An adjusted package for the next generation of own resources', COM(2023) 330 final.
European Commission (2023b) 'Staff Working Document accompanying the amended proposal for a Council Decision on the system of own resources of the European Union', SWD(2023) 331 final.
Fuest, C and J Pisani-Ferry (2020) 'Financing the European Union: new context, new responses', Policy Contribution 16/2020, Bruegel.
Hugger, F, AC González Cabral, M Bucci, M Gesualdo and P O'Reilly (2024) 'The Global Minimum Tax and the taxation of MNE profit', OECD Taxation Working Papers 68, Organisation for Economic Co-operation and Development.

This article is based on Bruegel Policy Brief Issue n°06/24 | March 2024.

Finance21 is pleased to announce that Euro Exim Bank has been awarded the Best CSR in Finance 2024.

The Finance21 awards celebrate achievement, innovation and excellence across several fields of endeavour. Our award programs are tailored to provide a comprehensive analysis of the very best in each market.

The Finance21 awards are recognised as the principal indications of professional conduct and excellence. The selection panel took into account product innovation, on-going customer support and best practice criteria as well as a continuing commitment to deploying the best possible solutions for the benefit of their clients.

Bulgaria in the eurozone: when?

Bulgaria meets all the nominal convergence criteria, except the one of inflation. Dimitar Radev examines the benefits in joining the eurozone as quickly as possible

The short answer to this question is: as soon as possible. Why? Because delaying our full integration into the core of the EU, such as the eurozone, has its price, and it is constantly increasing. More generally, this cost is expressed in continued marginalization in the political and economic periphery of Europe with the inherent themes of this periphery such as poverty, corruption and external dependencies. These topics are shifting Bulgaria's real agenda, which should focus on its modernisation, outpacing economic growth and people's wellbeing.

More specifically, the price of delay is expressed in harder conditions for business, trade and investment; less favourable financing conditions; higher non-productive costs for businesses and households. The scale of losses is measured not in millions, but in billions. Therefore, the short answer to the question raised is: Bulgaria, as soon as possible, must finalise the process of European integration by joining the eurozone.

I will also try to give a bit more detailed answer. For this purpose I will touch on three issues: the background of the accession process; the current situation; and some necessary steps to successfully complete this process.

Bulgaria has a long history in the European integration process with many lessons learned, but unfortunately also with a few lessons not learned. Due to time constraints, I will not go back that far, but only comment on the period after 2018. We should remember that Bulgaria was to a very large extent the initiator and leader of the last stage of the enlargement of the eurozone, which, as a rule, has begun with the accession of the national currency to the European Exchange Rate Mechanism.

The formal beginning was made in June 2018 with a letter of intent signed by the Minister of Finance and the BNB Governor and addressed to the Eurogroup, the EC, the ECB and the then 19 eurozone member states. We proposed an approach that was different from the one applied up till then. The main difference was that, in addition to the

inclusion of the Bulgarian lev in the European Exchange Rate Mechanism, Bulgaria was to join the Banking Union by establishing close cooperation between the BNB and the ECB.

Our partners have supported this approach, and the Eurogroup has announced that this is the approach that will be used for all new member states. Pursuant to this decision and exactly one year later this approach was also applied to Croatia, and thereafter the two countries moved as a package in the accession process.

Bulgaria, as soon as possible, must finalise the process of European integration by joining the eurozone

The first historically significant result was achieved in July 2020 when a European decision was made that the two currencies - the lev and the kuna – would join the Exchange Rate Mechanism and the two countries - Bulgaria and Croatia – would join the Banking Union.

Unfortunately, from the day of that decision, the two countries parted ways. Croatia clearly made the entry into the eurozone its top priority and mobilised all its political and expert capacity to achieve it. At the same time, the day of the decision marked the onset of the political crisis in Bulgaria which is still going on.

The results are known to all. Croatia has been a member of the eurozone since 1 January 2023, and Bulgaria continues to wander in the labyrinth of the political crisis.

What is the situation at the moment apart from the political context?

Now Bulgaria meets all the nominal convergence criteria, except the one of inflation, and the country meets them by a significant positive margin. The negative margin for the price stability criterion is narrowing, including in the last month, but it is expected to remain above the requirements for this criterion at the date of the forthcoming assessment to be made at the end of May.

Institutionally, the BNB and the banking sector have already become somewhat part of the eurozone by means of the close cooperation established between the BNB and the ECB in 2020. Now the BNB is the only central bank outside the eurozone that is operating in such a mode.

This ensures us a full and effective participation in the process of making and implementing the decisions on supervision and resolution of the banks in the eurozone, plus the banks in Bulgaria. The banking sector's results in the last more than three years testify to the success of this participation.

The Bulgarian lev is one of the two currencies, other than the euro, that participate in the European Exchange Rate Mechanism, which is one of the key conditions for joining the eurozone.

The adoption of the new Law on the BNB, which has received the support of the ECB and the EC, has virtually completed the process of legal convergence - one of the unalterable conditions for accession. The draft Law on the introduction of the euro, which establishes the administrative process of introducing the new currency, without being a formal requirement for legal convergence, is actually ready to be presented to, and adopted by, the National Assembly.

The logistical and technical preparations, which are largely within the BNB's competences, have reached a very advanced stage. The payment, information, accounting and statistical systems, for which the BNB is responsible, are practically ready to operate under the conditions of the eurozone and only need a final fine calibration.

We have ended the process of preparing the minting of the Bulgarian euro coins, which included coordination with the Commission and all member states, and we are moving on to the test minting of 8 million coins, 1 million of each denomination. The coin blanks, both for the test series and for the regular production, have been contracted and their delivery is about to begin.

We have provided the necessary areas for the exchange process, including in Sofia, Plovdiv and Varna, as well as the necessary machines and technical equipment for this process. By the end of May we expect to finalise the decision for Burgas, and by the end of this summer to put into operation the newly built cash centre in Pleven. The construction of this centre is part of our large-scale program for development and modernisation of the system of cash centres, which are practically industrial enterprises.

The Mint is fully prepared in terms of equipment and expertise for the minting of Bulgarian euro coins. It is yet to receive the necessary license for the minting of euro coins, which will take place after the upcoming certification of the test series of euro coins.

The joint venture of our Printing Works with the French company Oberthur Fiduciaire already has a license and prints euro banknotes for the needs of the eurozone central banks.

We are in the process of completely renewing our fleet of armoured and security vehicles, as well as approving the new transport schemes and security systems. To give you a general idea of the scale of the operation, I will point out that the total amount of banknotes and coins in the process of exchange alone amounts to about 12,300 tons, or from the point of view of transportation – the capacity of about 620 TIR trucks.

Over the past year, we have been actively working to create the capacity to conduct an active monetary policy, something that the BNB has not done in the past quarter of a century. The necessary organisational structure will be operational by the end of June.

We have created the necessary organisation to guide and control the entire preparation process. Twice a month, the Governing Council examines four reports of the three deputy governors and the general secretary, respectively, on the progress achieved, potential problems and measures to overcome them.

We are closely monitoring the commercial banks' preparations, which are also progressing according to plan, and reacting as necessary. With this incomplete list, I want to emphasise that the BNB and the banking sector will be fully ready within the current year for the introduction of the euro in our country.

Is this enough in purely technical, logistical terms? The answer is no. There are a number of, above all, information and accounting systems under the control of the executive and the municipalities, which must be adapted to work in the conditions of the eurozone. These have been identified, but considerable work is still required to be fully completed.

In addition, there are four sets of conditions in the powers of the executive that must be met before joining the eurozone, including in relation to the non-banking financial sector, insolvency, state-owned enterprises and anti-money laundering measures. I highly appreciate the caretaker government's intention to continue work on these topics.

What do we need to do to successfully finalise the joining process?

Above all else, we need a clearly established, sustainable, pro-European political structure. This is something that has been missing since we joined the European Exchange Rate Mechanism and the Banking Union.

In the earlier stage of the political crisis, the executive showed a hesitant position regarding the eurozone. For example, it took nearly a year for the government to adopt the plan drawn up and approved by the Coordination Council for the introduction of the euro, with six ministers voting against and one abstaining, including ministers key to the process.

In contrast to the earlier period, the last government stood on a clearly pro-European platform, but as the development of events has shown, it turned out to be extremely unsustainable. I also leave without comment the fact that since our admission to the Banking Union and the Exchange Rate Mechanism, we have had five different Ministers of Finance.

In establishing a sustainable pro-European political structure, we are expected to achieve several goals. First, re-establishing the political contacts on the subject at the highest level, as we must not forget that in the end it is a political process and a political decision. Unfortunately, in recent years the contacts on this topic have been protocol rather than substantive.

Second, returning to the path of fiscal consolidation, which is important not only for the accession process, but is also the basis of our most important comparative advantages in economic and financial terms. Such a development is also of great importance for the BNB, due to the need to harmonise the monetary and fiscal conditions in our country.

Third, adopting as quickly as possible the Law on the introduction of the euro, which will give clear indications, but also legal guarantees for businesses and households, as to what lies ahead.

Fourth, accelerating the work on the remaining conditions and technical preparation. Here I mean above all the issues of the government's competence, since, as I indicated, the BNB and the banking sector are working according to plan and will be fully ready before the end of this year.

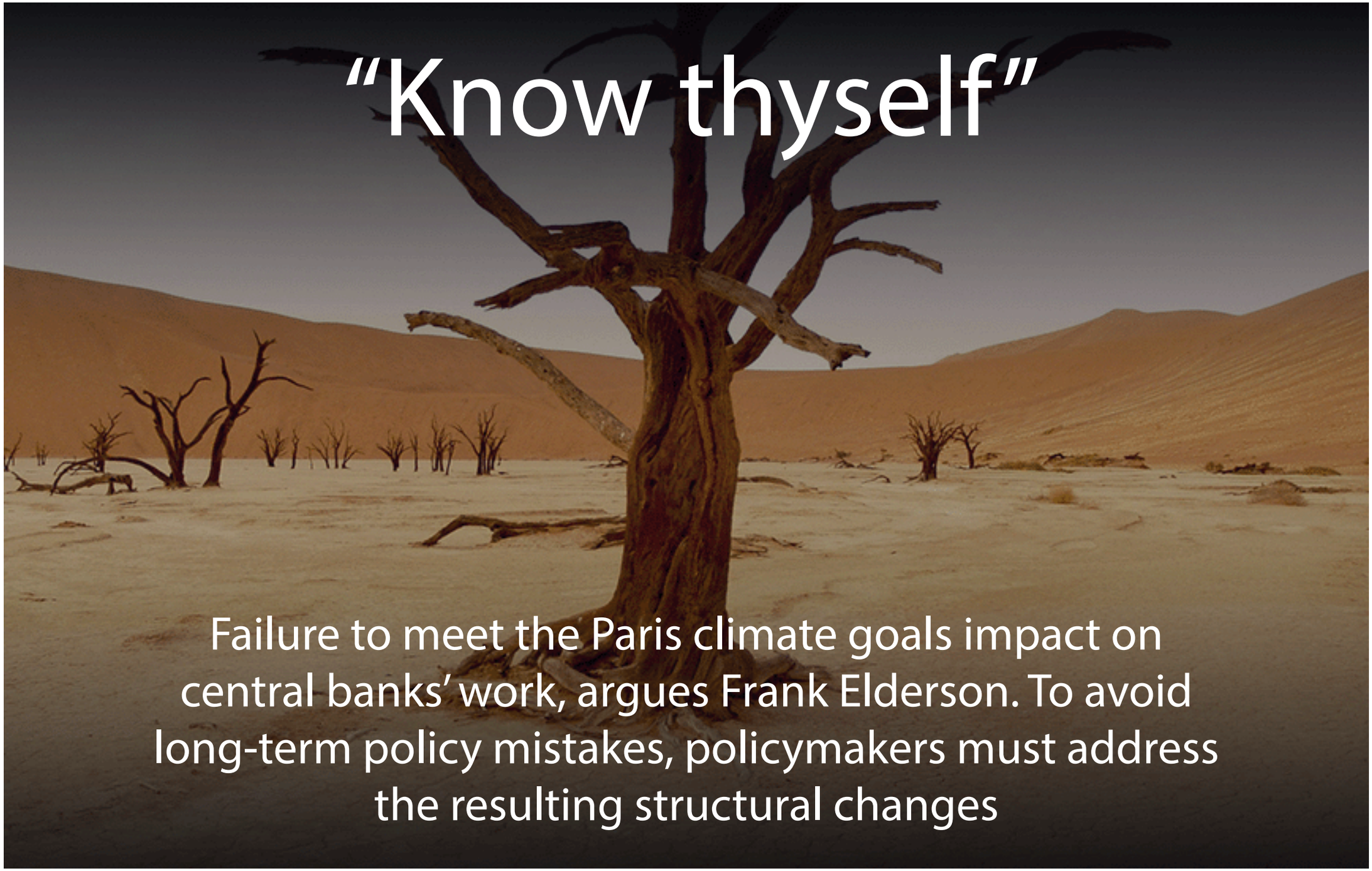
If this scenario materialises as soon as possible after the upcoming elections, the chances of joining the eurozone in 2025 remain strong and entirely realistic.

In conclusion, let me sum up what I said. As a result of the political crisis of recent years, we have lost both the initiative and the leadership in the process of joining the eurozone. Nevertheless, our readiness in terms of

accession conditions, legislative and technical framework remains high. To successfully finalise the process in 2025, we also need a sustainable, pro-European political structure. ■

Dimitar Radev is Governor of the Bulgarian National Bank

This article is based on a speech delivered at the opening of the conference 'Bulgaria in the eurozone, when?', Sofia, 14 May 2024. This article was originally published on BIS.



“Know thyself”

Failure to meet the Paris climate goals impact on central banks’ work, argues Frank Elderson. To avoid long-term policy mistakes, policymakers must address the resulting structural changes

For centuries the oracle of Delphi guided those seeking advice on what the future might hold[1]. Perhaps the most famous prophecy originating here from the Temple of Apollo is the one delivered during the Ancient Greek era to Croesus, the King of Lydia. When he consulted the oracle about going to war with Persia, he was told that if he were to attack, *"a great empire would fall."*

Emboldened by this apparent foresight, King Croesus went to war. And an empire did indeed fall. But it was the Lydians, not the Persians, who were defeated. The oracle was right. Yet King Croesus had overlooked the considerable room for interpretation that the prophecy allowed, with significant implications for his assessment of the outlook and the consequences of his decisions.

Today, policymakers count not on prophecies and oracles but on facts and science when assessing the outlook so they can make informed decisions. But while facts and science leave far less room for interpretation and uncertainty than ancient prophecies, they cannot eliminate it entirely.

The scientific method requires established knowledge to be scrutinised and reviewed, especially – though certainly not exclusively – knowledge that pushes the boundaries of modern science. So science-based models that are used to describe what happens in the real world need to be updated regularly, in terms of both their structure and their parameters.

And we have to acknowledge that these models are subject to uncertainty, including statistical, measurement and policy uncertainty. These caveats are relevant whenever we use these models to describe what has happened in the past, and they are especially relevant when assessing how present day knowledge is used to project an outlook for the future.

At the same time, policy must remain robust in the face of this uncertainty and build on what is scientifically established. Policymakers need to identify and spell out those questions that, if resolved, would reduce uncertainty and increase the level of confidence with which decisions are taken.

Analysis by the ECB and other central banks and supervisors repeatedly shows that, from an economic perspective, an orderly transition is by far preferable to alternative scenarios of doing nothing or doing too little too late

I will discuss how the prevailing evidence from climate and nature science can inform the actions of public authorities, even those that are not responsible for climate and nature policy, such as central banks and supervisors[2]. These public authorities – just like companies and individuals – are increasingly taking decisions whose outcomes will be subject to the tangible consequences of the ongoing climate and nature crises.

In fact, in a ground-breaking ruling earlier this week, the European Court of Human Rights, explicitly referring to *"the compelling scientific advice provided, in particular, by the Intergovernmental Panel on Climate Change"*, established that States *"need to put in place the necessary regulations and measures aimed at preventing an increase in greenhouse gas concentrations in the Earth's atmosphere and a rise in global average temperature beyond levels capable of producing serious and irreversible adverse effects on human rights."*[3]

So how can we ensure that decisions taken today reflect what we know about climate science while remaining robust in the face of uncertainty?

Fundamental challenges of failing to meet the goals of the Paris Agreement

Currently, the best assessment by climate scientists tells us that the world is not on a path to limit the increase in the average global temperature to 1.5 degrees Celsius above pre-industrial levels – the overarching goal of the Paris Agreement. We are not even on course to limit the increase to 2 degrees.

In fact, last November the UN Emissions Gap Report concluded that the world is on track for an average increase of 2.9 degrees, and even that will only be achieved if all government commitments to mitigation measures are implemented[4].

In other words, without a full and prompt implementation of these commitments, we will see an increase of even more than 2.9 degrees. In any case – acknowledging the uncertainty – the world is currently heading for a temperature rise far above the Paris Agreement goals.

This raises a number of critical challenges for maintaining wellbeing as we know it. These go far beyond the economic challenges that may emerge and will be particularly relevant for central banks and supervisors.

In a recent report, the European Environment Agency sent a dire message about climate risks, pointing out that *"several climate risks have already reached critical levels"* and observing that *"[i]f decisive action is not taken now, most climate risks could reach critical or catastrophic levels by the end of this century."*[5]

Global heating will have an impact on food, water and energy security and the health of the general population, and these effects will be aggravated by ecosystem degradation, which is itself worsened by global heating. Moreover, increasing climate and natural hazards can disrupt critical infrastructure, putting people's livelihoods and even their basic needs at risk.

There may also be second-round effects that compound the direct impact of an increase in climate and natural hazards. One example of this would be changes in migration flows, which like other such second-round effects are generally not yet accounted for in models of the impact of climate change and nature degradation. But the more severe the climate scenario, the more likely it is that these flows will increase, and the greater the impact these increasing flows will have[6].

In addition, the Intergovernmental Panel on Climate Change (IPCC) has been increasingly emphasising the risks of various tipping points. These are critical thresholds that, when breached, will lead to large, accelerating and irreversible changes to our climate system.

According to the most recent IPCC assessment report from 2021, the risk of reaching these tipping points is already assessed as being high if the average global temperature increase amounts to between 1.5 and 2.5 degrees. And it is assessed as very high if global temperatures increase by 2.5 to 4 degrees[7].

Climate science can provide indications of potential tipping points and what their consequences might be, like the melting of the Greenland ice sheet and the impact it would have on global sea levels. There is, however, no scientific consensus yet on the systemic changes that might occur after these tipping points are reached.

Further research is therefore urgently required here, especially in light of the current trajectory for global heating[8]. Over the last 12 months, the global average temperature was already 1.5 degrees above pre-industrial levels.

Structural economic challenges

Let me now turn to the implications for the global economy if temperatures increase by significantly more than 2 degrees. The structural economic consequences will be profound, with impacts on both the supply and demand sides of the economy.

First, resources will have to be dedicated to protecting citizens and society from increased climate and natural hazards like wildfires, droughts and floods.

Second, to the extent that the increase in hazards can no longer be avoided, the economy will need to cater for the critical needs that the European Environment Agency identifies as being at risk. Specifically, maintaining adequate food production, water availability and health care will require substantially more resources than those sectors currently receive.

Third, beyond catering for these critical needs, the economy will undergo further structural transformation as both preferences and production possibilities change. Tourism is a case in point, with destinations that are currently popular no longer being similarly in demand or even accessible in the future.

Another example is international trade, which may be forced to redevelop as existing routes and ports become unavailable and others open up. And there will also be a reallocation between sectors, with some losing out while others benefit, much like we have seen following the pandemic and the energy crisis.

Fourth, the economy needs to be made resilient to the increase in climate and natural hazards. The existing capital stock – including people's homes – will need to be upgraded and adapted, with all the increases in structural costs this entails. Achieving such resilience may even require physically relocating part of the capital stock to avoid proximity to areas that will be heavily exposed to hazards.

Any capital stock that is not made resilient to hazards will most likely see its economic lifespan shorten significantly. This will take the form of higher depreciation rates, which imply greater financial risks for anyone with exposures to the capital stock.

It is particularly noteworthy that investments that are currently being made to green the capital stock may not be immune to this effect. For example, a hydroelectric power plant may become obsolete prematurely if a river runs dry or changes course.

Resilience to the more disastrous climate and nature outcomes that are the consequence of failing to meet the Paris Agreement goals should, therefore, feature prominently in any decisions related to mitigation investment that are being taken today.

A key challenge for economic policymakers will be to ensure that the economy is suitably prepared to undergo these structural transformations. If it is not, there is a significant risk that economic and financial factors will actually exacerbate the critical challenges we will face in a world that overshoots the goals of the Paris Agreement.

Against this backdrop, it will be crucial for economic policymakers to identify potential barriers to effective and efficient adaptation. First, a failure to coordinate may lead to investment being misallocated. Some investments may not materialise at all if the private sector fails to consider the benefits for society.

And others may materialise but only inefficiently, for example if investment in cooling homes and offices takes place at the level of individual households and firms.

Second, structural adjustment in an economy requires the right combination of flexibility, education and social safety nets to navigate an inclusive and effective adaptation process.

Third, financial bottlenecks may emerge. Increased uncertainty due to potential climate and natural hazards may lead to an increase in risk premia, which in turn could hold back investment. And this situation could be exacerbated if it is no longer possible to obtain insurance against certain risks – or if it is only possible at a prohibitive cost[9].

Besides the greater frequency and impact of hazards, uninsurable risks occur when hazards become systemic – in other words, when a hazard would affect the entire population at once if it were to materialise. And when such risks are uninsurable, individuals and firms – as well as the financial institutions that finance them – need greater loss-absorbing capacity themselves. This self-insurance will mean that – all other things being equal – the aggregate propensity to invest decreases further.

Bottlenecks in the flow of finance that reduce investment or that lead to misallocation can be mitigated with a sound banking system and well-developed capital markets that bolster transparency and ensures climate- and nature-related risks are properly priced.

Against this backdrop, there is an urgent need to complete the banking union and the capital markets union – as the ECB has previously called for – irrespective of the climate and nature scenario that ultimately materialises.

In areas where private investment bottlenecks cannot be resolved, governments may need to step in with increased public investment and safety nets. This would give rise to significant government contingent liabilities that are not yet appropriately reflected in credit ratings or in institutional economic governance frameworks.

The relevance for central banks and supervisors

Many of the challenges I have mentioned – both the critical and the structural economic challenges – fall to policymakers in other areas, rather than central banks and supervisors. But the challenges presented and the policy choices that are taken in response will have a bearing on the environment in which central banks and supervisors pursue their mandates to maintain price stability and ensure the safety and soundness of banks.

First, our objectives are even more important in a world that is facing increased climate and natural hazards. Price stability and sound banks provide an anchor that makes an economy – and therefore a society – more resilient to shocks. The more frequent and intensive the shocks, the more important it becomes that the anchor doesn't break.

Second, while our tasks become more important when the world around us becomes more daunting, maintaining price stability and a sound banking sector becomes more complicated. And this is not just because shocks are more frequent and more intense. It also becomes more complicated to assess the type of shock that is hitting the

economy, yet this is crucial to gauging the potential risk to price stability or to the soundness of banks, as well as the appropriate policy response.

It could raise questions about whether climate and natural hazards can be fully captured in the traditional categorisation of demand, supply and financial shocks that are inherent in most macroeconomic models. For example, my fellow ECB Executive Board member Isabel Schnabel has suggested thinking about the impact of climate change on inflation using concepts that she has referred to as 'climateflation', 'fossilflation' and 'greenflation'[10].

The Basel Committee on Banking Supervision, meanwhile, has already established that climate-related risks translate into the traditional types of risk that banks consider[11]. This covers credit risk, liquidity risk, market risk and operational risk, including litigation risk[12].

However, the exact mechanisms of mapping actual hazards to risks still need to be analysed further to fully capture climate-related factors in quantifiable regulatory and supervisory requirements.

Third, climate and natural hazards limit the productive capacity of the economy. Some of the consequences may eventually fade – although they may well persist for quite some time – for example if supply chains are disrupted as a result of hazards materialising. Others may be permanent, for example if nature providing critical services – including land use and fisheries – becomes degraded.

In both cases, the risk of the economy running into capacity constraints would be greater. Therefore, to properly assess the state of the economy and identify risks, central banks and supervisors need to further deepen their understanding of the supply side of the economy, just as we had to do after the pandemic and the energy crisis.

This also means that we need to extend the horizon of our analyses well beyond the typical horizon considered today. Climate science gives us a window into the rest of this century. What we can see through this window should be taken seriously, including by central banks and supervisors as we identify and assess risks in the pursuit of our mandates. The time to think seriously about the long term is now.

Fourth, the combination of heightened uncertainty and a greater need for self-insurance could lead to an increase in the propensity to save in the private sector. This could create space for the investment that is so urgently needed and – in the absence of increased savings – would lead to an increase in the equilibrium real interest rate[13].

At the same time, if owing to coordination failures the increased savings are not channelled towards providing the investment needed, the equilibrium real rate of interest would instead be depressed. As this equilibrium rate is the interest rate that prevails when all shocks to the economy have dissipated and monetary policy is neither accommodative nor restrictive, it is an important yardstick for central banks. Thus, for monetary policy, understanding which of these effects ultimately dominates will be key.

Fifth, increasing financial risks arising from the climate and nature crises can impair the soundness of financial institutions and the stability of the financial system as a whole. Should these risks materialise – despite all our efforts to mitigate them – the transmission of our monetary policy could be affected.

Monetary policy decisions would be transmitted through the financial system and the economy in a less orderly and less predictable manner, potentially making it more difficult for us to achieve our price stability objective.

More generally, the effectiveness and efficiency of our policies benefit from well-functioning markets. This holds true in terms of both our ability to maintain price stability and the need to avoid the risk of our monetary policy impulses unduly contributing to a misallocation of resources.

Concluding remarks

The Temple of Apollo in Delphi famously bore the inscription *"Know thyself"* – a maxim that is often understood to mean *"know your limits."*

Know what you know and know what you don't know – this is what I have sought to convey to you.

And act upon that knowledge in a way that is robust in the face of known and unknown uncertainties, to avoid making avoidable mistakes like that of King Croesus after he consulted the oracle of Delphi. This includes identifying and seeking answers to questions that reduce uncertainty and increase the scope of 'no-regret' policy actions.

This will require policymakers to engage with stakeholders beyond their own fields of expertise – just like the Bank of Greece is doing through the interdisciplinary Climate Change Impacts Study Committee, which recently announced the preliminary results of analytical work on the economic, social and environmental impacts of climate change in Greece[14].

Experts from all disciplines – including climate and nature scientists, biologists, economists, legal experts and sociologists, to name just a few – will need to work closely together in responding to the multifaceted challenges ahead. If ever there was an urgent need to pool knowledge and draw on different fields of expertise, it is now.

Let me be clear: my remarks are by no means a signal that we should throw in the towel on mitigation. Quite the opposite. I hope that I have been able to show you why, in light of the prevailing climate science, no effort should be spared in working towards the goals of the Paris Agreement.

The European Climate Law requires it, and the European Court of Human Rights has ruled that governments that fail to meet their climate commitments are violating human rights. Analysis by the ECB and other central banks and supervisors repeatedly shows that, from an economic perspective, an orderly transition is by far preferable to alternative scenarios of doing nothing or doing too little too late[15].

That said, even though climate and nature policymakers are under a legal obligation to deliver on the goals of the Paris Agreement and even if they have committed to achieving these objectives, they still have a duty to prepare for risks that lie ahead as the entire world needs to live up to its obligations – and it is not a given that it will – and critical thresholds may have already been surpassed.

The duty to prepare for these risks also holds for central banks and supervisors in the pursuit of their mandates. We must both unwaveringly strive for the best and diligently prepare for what climate science tells us lies in store.

It is not a Delphic prophecy that is calling for action. It is facts and science. ■

Frank Elderson is a Member of the Executive Board and Vice-Chair of the Supervisory Board of the European Central Bank

Endnotes

1. To my knowledge, there are at least three instances of central banks and supervisory authorities paying tribute to the ancient oracle. The semi-structural macroeconomic model of the Dutch economy that De Nederlandsche Bank uses for its projections is named DELFI. In ECB Banking Supervision we have developed a tool named Delphi that integrates market indicators and information from the media to better understand risk developments affecting banks in real time. And central banks have been described as giving "Delphic" forward guidance when communicating about how they intend to adjust policy in relation to incoming data.

2. I have emphasised in other speeches that central banks are not climate and nature policymakers, but climate and nature policy takers. See, for example, Elderson, F (2023), "Policymakers as policy takers – accounting for climate-related and environmental factors in banking supervision and monetary policy", speech at the Peterson Institute for International Economics, 21 April.

3. European Court of Human Rights (2024), "Judgment Verein KlimaSeniorinnen Schweiz and Others v. Switzerland – Violations of the European Convention for failing to implement sufficient measures to combat climate change", press release, 9 April.

4. United Nations Environment Programme (2023), Emissions Gap Report 2023: Broken Record – Temperatures hit new highs, yet world fails to cut emissions (again).

5. European Environment Agency (2024), European climate risk assessment.

6. According to the World Bank, climate change could contribute to the movement of 216 million people within their own countries by 2050, unless concrete climate and inclusive development actions are taken. See Clement, V et al (2021), Groundswell Part 2: Acting on Internal Climate Migration, World Bank Group, Washington, D.C.

7. Intergovernmental Panel on Climate Change (2021), Climate Change 2021 – The Physical Science Basis.

8. The Central Banks and Supervisors Network for Greening the Financial System has previously developed scenarios to assess how economies might look on different climate policy paths. In future work it will prioritise the inclusion of non-linear elements – like climate tipping points – in its models (see Aerts, S, Spaggiari, M and Stracca, L (2023), "Climate

scenarios: procrastination comes at high cost", The ECB Blog, 4 December). For it to achieve this, climate and nature science will be crucial in advancing its understanding of tipping points.

9. Together with EIOPA the ECB has issued a discussion paper that outlines policy options to promote climate catastrophe insurance that could mitigate the effect of reduced insurability, see ECB and EIOPA (2023), "Policy options to reduce the climate insurance protection gap", Discussion Paper, April.

10. Schnabel, I (2022), "A new age of energy inflation: climateflation, fossilflation and greenflation", speech at a panel on "Monetary Policy and Climate Change" at The ECB and its Watchers XXII Conference, 17 March.

11. Basel Committee on Banking Supervision (2021), Climate-related risk drivers and their transmission channels, April.

12. On litigation risk, see Elderson, F (2023), ""Come hell or high water": addressing the risks of climate and environment-related litigation for the banking sector", speech at the ECB Legal Conference, 4 September.

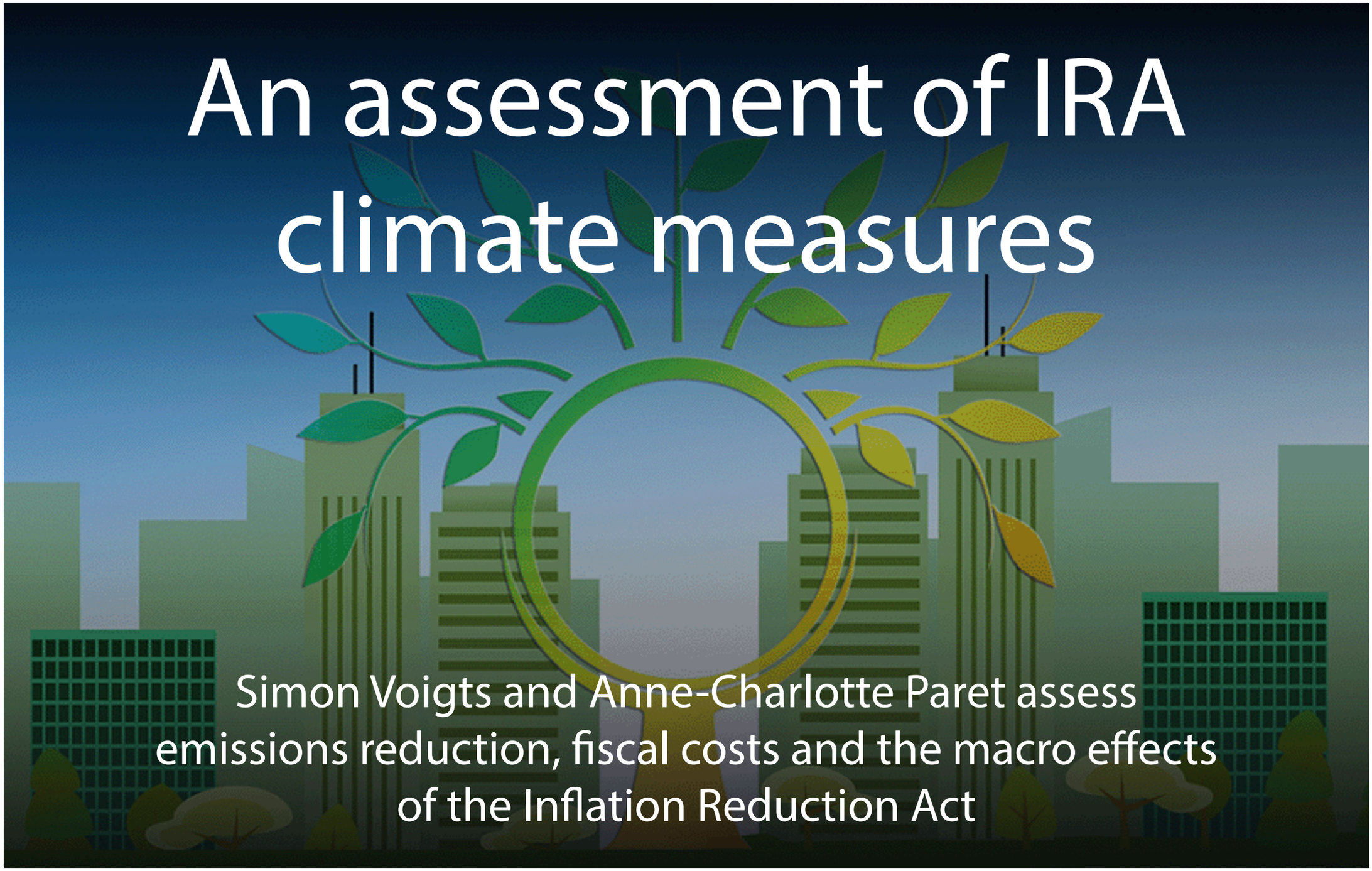
13. See, for example, Schnabel, I (2024), "R(ising) star?", speech at The ECB and its Watchers XXIV Conference session on Geopolitics and Structural Change: Implications for Real Activity, Inflation and Monetary Policy, 20 March.

14. Bank of Greece (2023), "Preliminary results of the studies on the vulnerability assessment and the impact of climate change in Greece", 15 December.

15. Emambakhsh, T et al (2023), "The Road to Paris: stress testing the transition towards a net-zero economy", Occasional Paper Series, No 328, ECB

This article is based on a keynote speech delivered at the Delphi Economic Forum IX, Delphi, 12 April 2024.

An assessment of IRA climate measures



Simon Voigts and Anne-Charlotte Paret assess emissions reduction, fiscal costs and the macro effects of the Inflation Reduction Act

There is a wide gap between most countries' greenhouse gas (GHG) mitigation pledges and actual policy implementation, putting the global economy widely off track to honour the 2015 Paris Agreement (UN Environment Program 2023). The Inflation Reduction Act (IRA), signed into law by President Biden on 16 August 2022, aims to significantly narrow that implementation gap in the US, in addition to pursuing other objectives.

In a recent paper (Paret and Voigts 2024) we apply the IMF's new Global Macroeconomic Model for the Energy Transition model (GMMET, see Carton *et al* 2023) to assess the impact of those IRA measures that are related to climate and energy security, focusing on both GHG emissions and the macroeconomy up to 2030.

While the IRA has been discussed extensively (including on Vox, see for example Fajeau *et al* 2023 and Attinasi *et al* 2023), we contribute to the literature by employing a model that captures key measures in a granular fashion, by assessing complementary policies to bridge the remaining gap to the US' medium-term climate pledge, and by shedding light on the dynamic implications of a permitting reform.

GMMET builds on the IMF's Global Integrated Monetary and Fiscal model (GIMF), which is a large-scale, non-linear, structural, multi-country New Keynesian dynamic general equilibrium model for quantitative monetary and fiscal policy analysis. GMMET adds a granular, sector-specific modelling of key GHG-emitting sectors that allow to capture sectoral idiosyncrasies playing a crucial role for emission mitigation.

These sectors include: (i) an electricity generation sector with different technologies (renewables, coal, gas, nuclear) and explicit treatment of intermittent generation from renewables; (ii) a transportation sector with conventional cars, electric vehicles (EVs) and a charging station network (giving rise to network effects); and (iii) fossil fuel-specific mining sectors.

Due to GMMET's sectoral granularity, most key measures have a direct representation in the model, so that their uptake is determined endogenously. To proxy for the IRA's tax on profits made by large corporations and the excise tax on stock buybacks and exemptions, the measures are assumed to be funded by corporate income taxes, implemented as a tax on the profit from the ownership of firms.

The social value of the induced emission cuts outweighs their fiscal costs

The following measures are modelled:

- **Electricity sector measures:** The Clean Electricity Production Tax Credit (PTC) is represented by a 33% subsidy on the model's renewable utility's total production cost, while the New Advanced Manufacturing Production Tax Credit is a 40% subsidy on the price of capital good employed by the utility. The Nuclear Power Production Tax Credit raises nuclear power investment such that capacity increases by roughly 15%.

- **Transport sector measures:** The Clean Vehicle Credit is a gradually increasing subsidy on EV purchases reaching 15% by 2030 (proxying for a slow increase in the share of manufacturers fulfilling domestic content requirements). The Alternative Fuel Refueling Property Credit exogenously increases the charging network density by 13.4%, based on an estimate of the charger station supply elasticity in Cole *et al* (2023). Both measures have a direct model representation, as the choice between both car type and the charging station network are explicitly modelled to capture network externalities.

- **Other measures:** The Carbon Capture and Sequestration Tax Credit and measures related to agriculture and waste exogenously reduce the tradable goods emission intensity. The various measures aimed at improving residential energy efficiency improve productivity on the bundle of natural gas and oil that is used by households for home heating. This allows to capture the general equilibrium impact of reduced fuel demand. Spending for all three measures is captured as government spending.

The impact of the IRA

Selected key results emerge from our analysis and are presented in the following. Absent permitting-related investment delays, IRA climate measures deliver large emission reductions at manageable fiscal costs and with an expansionary but very small impact on the overall economy:

- The share of renewables in the electricity mix rises by around 19 percentage points by 2030, at the expense of gas and coal. The Clean Electricity Production Tax Credit and the New Advanced Manufacturing Production Tax Credit lower the renewable utility's overall generation costs and, respectively, the price of capital.

  As shown in Figure 1, this triggers a surge in renewable investment and reduces investment in coal utilities via crowding out. Gas investment ticks up slightly due to its role as a back-up for renewables, and investment in nuclear power rises mildly from the Nuclear Power Production Tax Credit. The subsidies boost the total volume of electricity generation and thereby reduce the electricity price.

- The share of EVs in newly purchased cars increases by slightly less than 5 percentage points upon introduction of the Clean Vehicle Credit in 2022, and thereafter rises gradually to reach 19% by 2030. The charging station density increases when the Alternative Fuel Refueling Property Credit is adopted, and then rises gradually until it more than doubles by 2030. EV adoption and charger deployment reinforce another in a virtuous circle driven by network externalities, where a denser charging network incentivises EV adoption, while a rising EV share incentivises charger deployment.

- The impact on the macroeconomy is expansionary but very modest in size. Aggregate investment rises, reflecting investments in the electricity and manufacturing sectors. The latter is driven by lower electricity prices as energy and capital are assumed to be complements. This complementarity, combined with the non-distortionary source of funding, provide the key explanation for the mild increase in output of close to 0.25% by the end of the decade.

  The remaining aggregates exhibit a milder response, and the impact on inflation and the policy rate are

Figure 1. Electricity mix

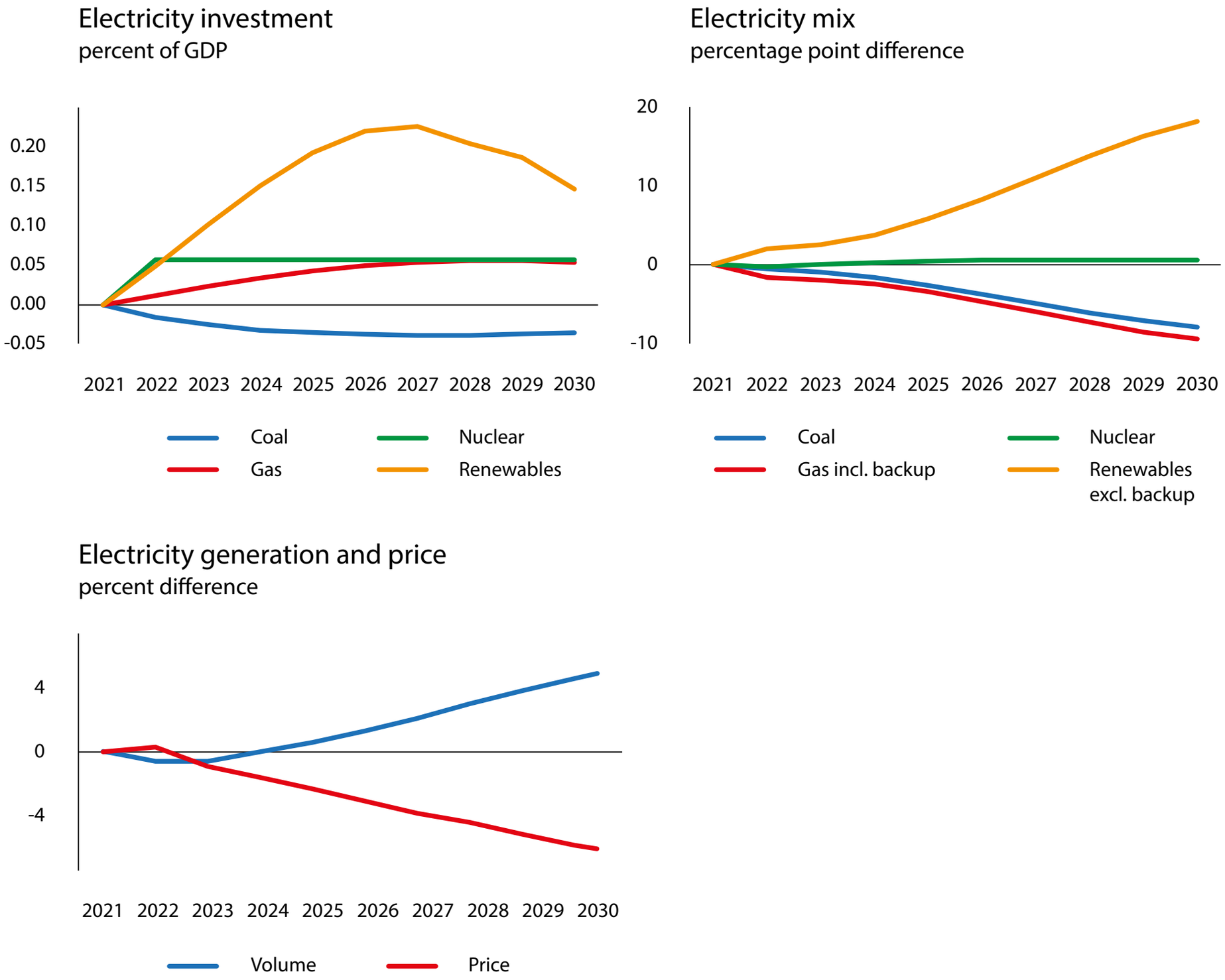


Figure 2. EV share

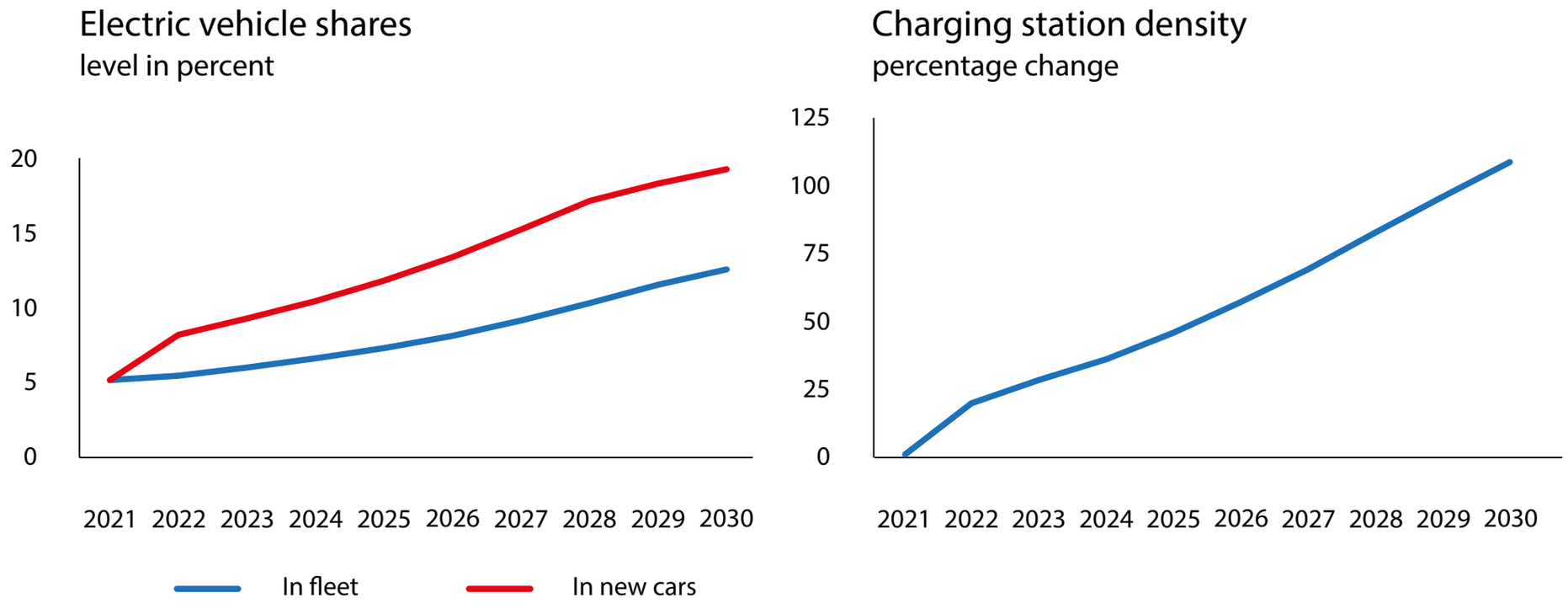
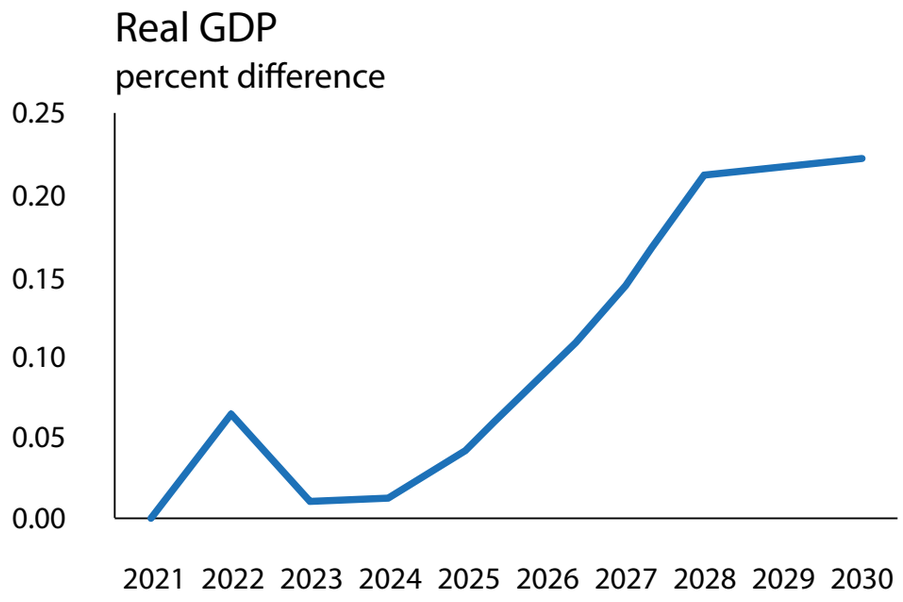
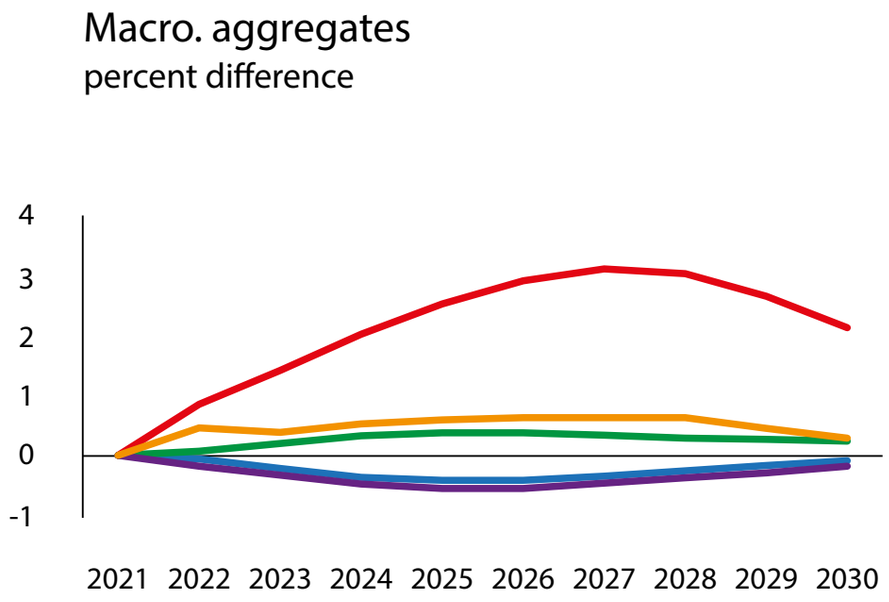
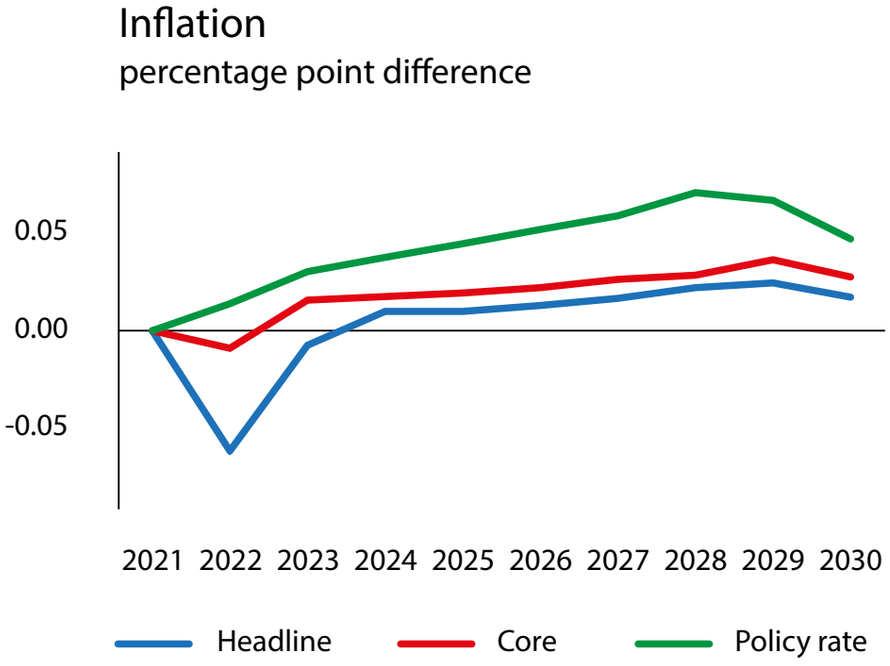


Figure 3. Impact on the macroeconomy







negligible. The main explanation for the muted macroeconomic impact is that the electricity sector, where the IRA has the largest effect, is small. While the depicted impact is obtained from a simulation with lump-sum funding of the measures, the results under corporate income financing are virtually identical; the only noticeable difference is that output rises by roughly half as much (given the small magnitude of the adjustment, it is almost identical in absolute terms), owing to the additional tax-induced distortions.

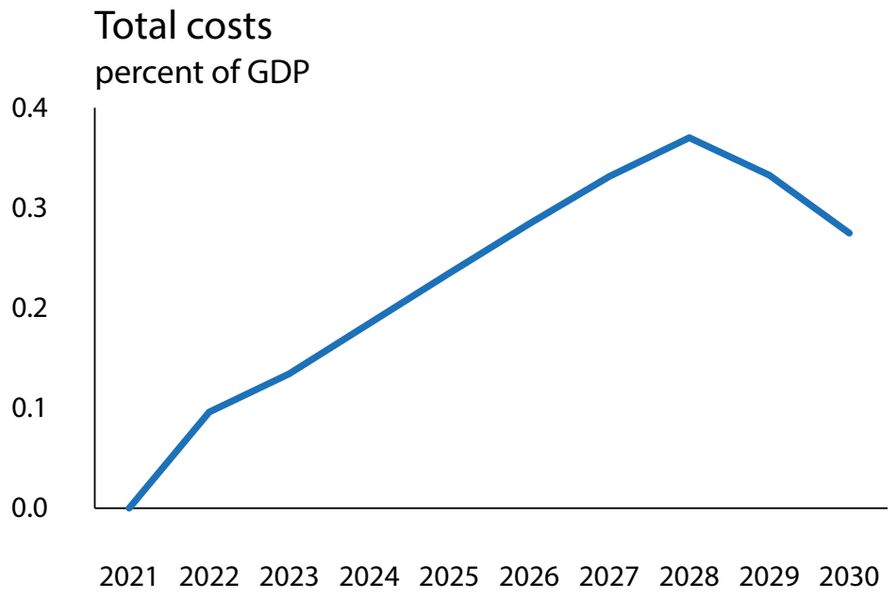
The takeaway is that, even when the funding side is modelled appropriately, climate-related IRA measures have a vanishingly small impact on output and inflation.

- Total fiscal costs are estimated to be in the same order of magnitude as in the 2023 update by the CBO/JCT (CFRB, 2023). They rise over time as subsidy take-up increases (especially for capital goods employed by the renewables utility), and peak at about 0.4% GDP towards the end of the decade, when renewables investment comes down as the capital stock has grown. Applying these cost shares to nominal GDP projections from the October 2023 World Economic Outlook, and cumulating through 2030, yields undiscounted total costs of about $700 billion.

  Together with IRA climate-measures that are not modelled (and whose emission impact is therefore not captured) total costs stand at $820 billion. This is well above the initial CBO/JCT estimate of about $350 billion over this period (CBO, 2022), but only moderately above the $590 billion by 2030 estimated in the 2023 update of JCT scores, and relatively close to some other recent estimates (eg. Fajeau *et al* 2023).

- Total annual emissions decline by about 710 MMT by 2030, mostly driven by electricity generation, followed by the transport sector and agricultural measures. Assuming a no-policy emission reduction of 27% between 2005 and 2030, IRA measures bridges slightly less than half of the way to the mitigation target of

Figure 4. Fiscal costs



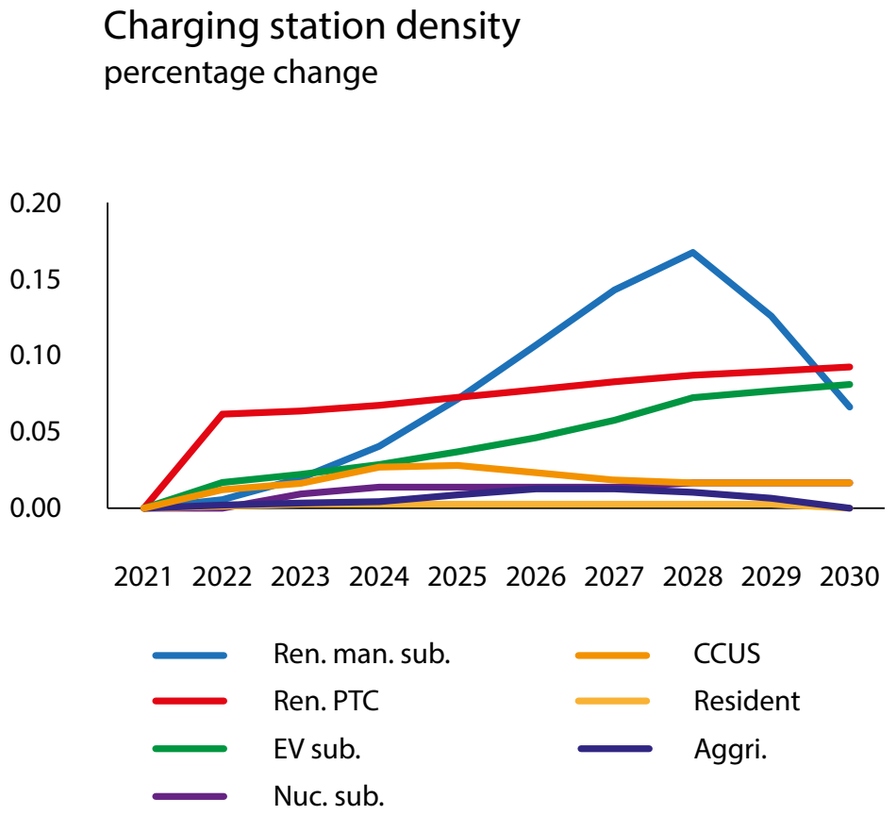
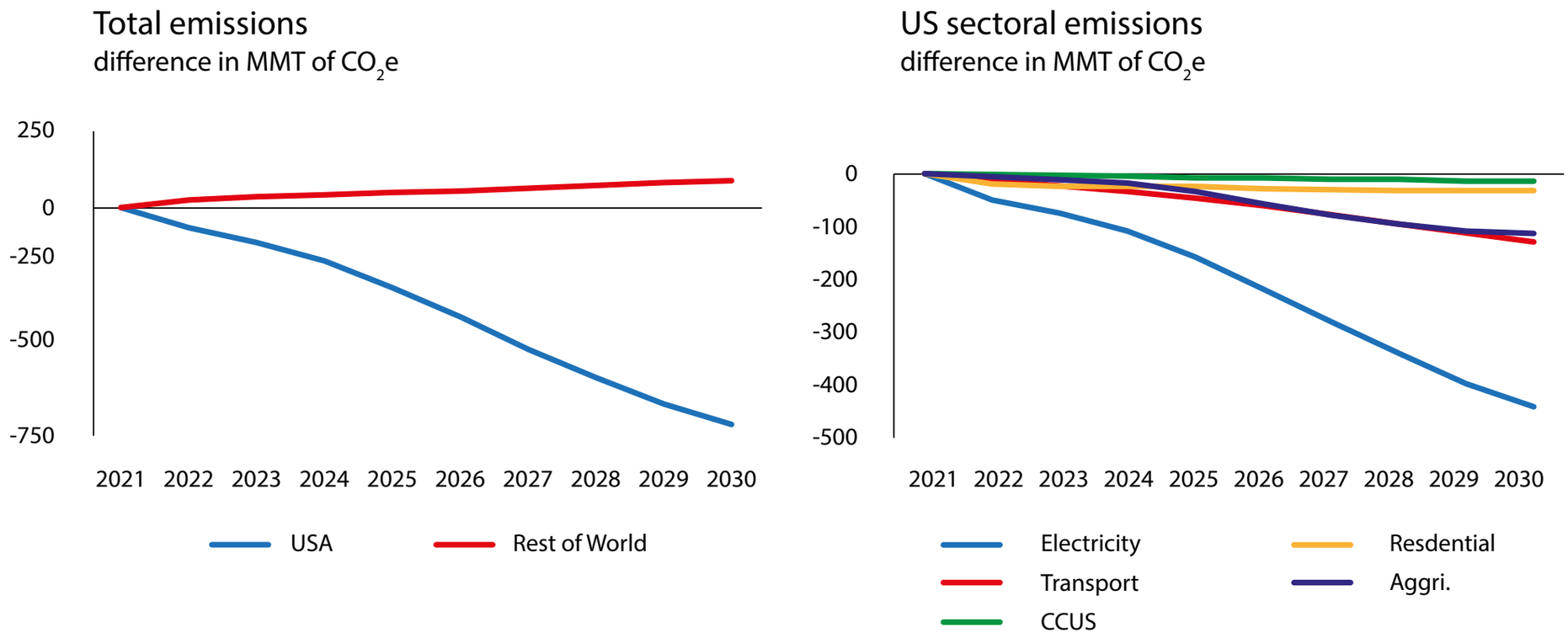


Figure 5. Annual emissions



a 50-52 percent reduction over this period. Emissions in the rest of the world increase by about 100 MMT. With weaker demand from the US, fossil fuel prices decline on global markets, causing an uptick in foreign consumption.

The social value of the induced emission cuts outweighs their fiscal costs.

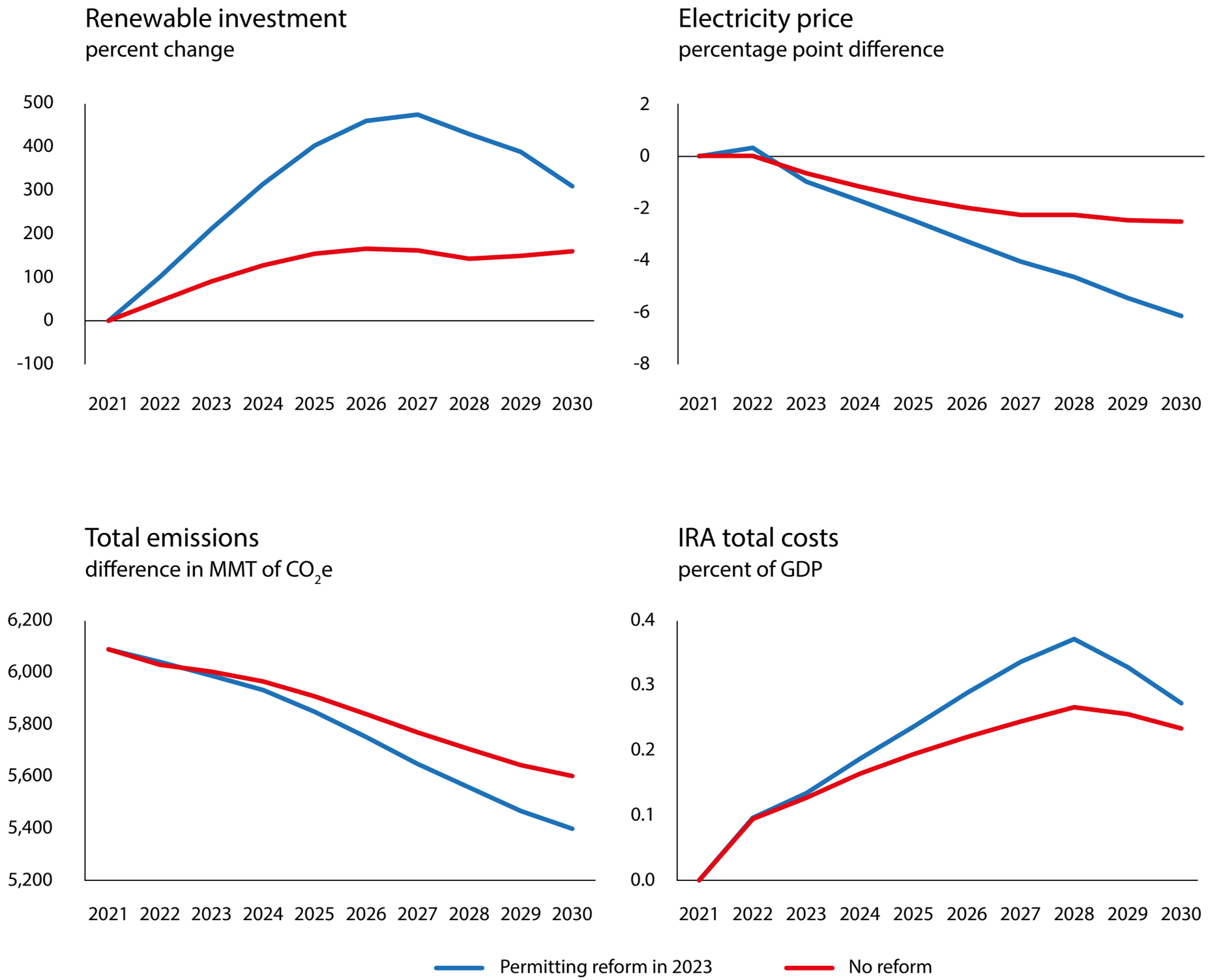
While GMMET does not feature warming damages, and therefore does not allow for a cost-benefit analysis, we still provide an indication that the IRA's climate measures carry a social value greater than their fiscal cost. To approximate the measures' desirability, we compare the fiscal costs per ton of GHG reduction from our simulation with a plausible estimate of the social costs of carbon (SCC), $185 per tonne, taken from Rennert *et al* (2022).

The ratio of cumulative fiscal costs over cumulative emission reductions – a metric for average fiscal abatement costs – stands at about $400/tCO$_2$ in 2022, but then declines swiftly to reach the SCC of $185/tCO$_2$ in 2029 and settles at $50/tCO$_2$ in the long run (the decline results from subsidy-induced investments yielding long-term emission reduction benefits). This suggests that by the end of the decade, the social value of IRA emission reductions greatly outweighs their fiscal cost, making the measures highly desirable from a cost-benefit viewpoint.

Reducing permitting-related delays in energy investment is crucial to unlock the measures' full potential.

The National Environmental Policy Act of 1970 requires federal permit for infrastructure projects, including for energy, and this permitting process takes around 4.5 years on average (eg. American Clean Power Association 2023). This is captured by an adjustment in electricity investment rigidity.

Figure 6. Outcomes



If permitting delays remain in place (ie. if permitting processes are not shortened, in contrast to the previous simulation), the investment surge in renewable generation is attenuated, delaying the addition of new generation capacity and with it the decline in the electricity price.

Emissions drop by only two-thirds of the amount in the absence of permitting delays, and the dampened take-up of renewable subsidies cuts fiscal costs, while the implications for the adjustment of output and inflation are negligible in absolute terms.

Additional policies

In our paper, we also consider two hypothetical measures that could complement the IRA to substantially reduce the mitigation policy implementation gap. The measures target areas of low-cost emission abatement that are not addressed to a significant extent by the IRA:

- A regulatory measure (introduced as a feebate-like tax) that reduces the coal share in the electricity mix by about one percentage point each year. This would tap into room for low-cost abatement resulting from 15% post-IRA coal share in 2030 (lacking targeted measures, IRA only curbs electricity generation from coal via crowding-out).

- A regulatory measure leading oil and gas industries to abate about three-quarters of today's methane emissions. The associated costs are modelled as a decline in the productivity in GMMET's oil and gas mining sectors, calibrated based on abatement costs estimates. Emissions from these industries have recently been estimated to be vast, at nearly 400 MMT of CO_2e, while about 300 MMT could be abated at minimal costs (IEA, 2023).

However, observed emission cuts have been negligible relative to their potential. The effectiveness of IRA provisions aimed at curbing methane emissions is likely to be minor, as legislative details (including reporting thresholds and emission aggregation rules) limit the scope of emissions that are covered (Mahajan *et al* 2022 expect 2030 emissions to decline by only 29 MMT).

When the IRA climate measures are complemented by the two regulatory measures, the drop in coal power plant investment becomes stronger, which amplifies the rise in renewables and gas investment. Regarding electricity generation volume and price, the disinvestment from coal triggered by the regulation works in the opposite direction from IRA subsidies, boosting renewable generation capacity.

Initially, the decline in coal generation dominates, but the IRA-induced surge in capacity more than offsets this from 2026 onwards, leading to a rise in the electricity volume (and a decline in the price) by the end of the decade. In the short term, the regulatory measures slightly reduce output and push up inflation, but the overall picture is virtually unchanged. This is not surprising given that methane abatement comes at minimal cost and that the coal regulation lowers its electricity share very gradually.

However, the complementary regulatory measures greatly reduce emissions, which would drop by a total of about 1,300 MMT by 2030, nearly covering the remaining gap to the emission reduction target. ■

Simon Voigts and Anne-Charlotte Paret are Economists at the International Monetary Fund

References

American Clean Power Association (2023), "U.S. Permitting Delays Hold Back Economy, Cost Jobs".

Attinasi, MG, L Boeckelmann and B Meunier (2023), "Unfriendly friends: Trade and relocation effects of the US Inflation Reduction Act", VoxEU.org, 3 July.

Carton, B, C Evans, DV Muir, and S Voigts (2023), "Getting to Know GMMET: The Global Macroeconomic Model for the Energy Transition", IMF Working Paper No. WP/23/269.

Cole C, M Droste, C Knittel, S Li and JH Stock (2023), "Policies for Electrifying the Light-Duty Vehicle Fleet in the United States", AEA Papers and Proceedings 113.

CRFB – Committee for a Responsible Federal Budget (2023), "IRA Energy Provisions Could Cost Two-Thirds More Than Originally Estimated".

Fajeau, M, N Garnadt, V Grimm et al (2023), "The US Inflation Reduction Act: How the EU is affected and how it should react", VoxEU.org, 17 October.

International Energy Agency (2023), Methane Tracker Database, IEA, Paris.

Mahajan, M, O Ashmoore, J Rissman, R Orvis and A Gopal (2022), "Updated Inflation Reduction Act Modelling Using The Energy Policy Simulator", Energy Innovation, Policy and Technology LLC.

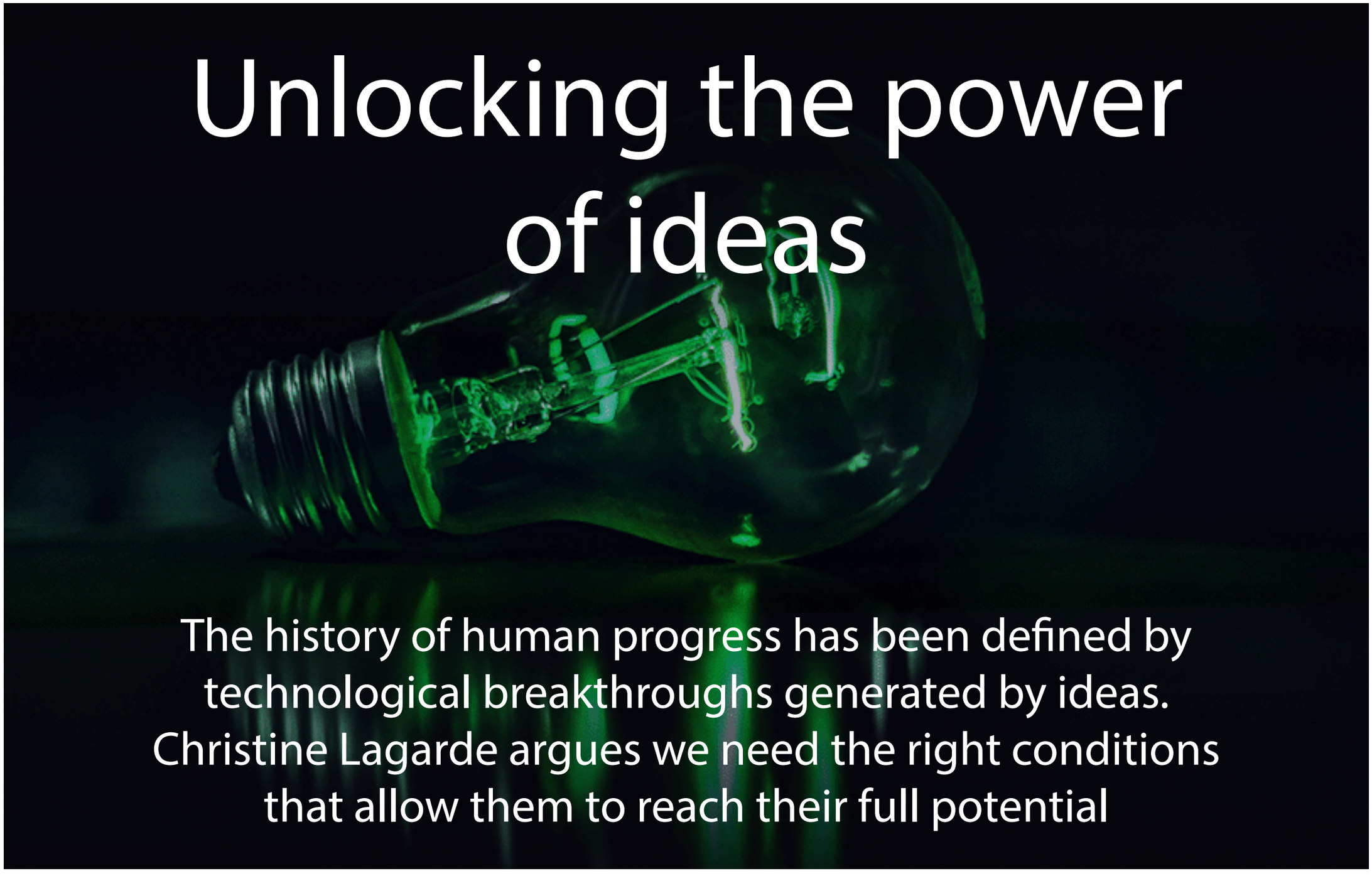
Paret, A-C and S Voigts (2024), "Emissions Reduction, Fiscal Costs, and Macro Effects: A Model-based Assessment of IRA Climate Measures and Complementary Policies", IMF Working Paper No. WP/24/24.

Rennert, K, F Errickson, BC Prest et al (2022) "Comprehensive evidence implies a higher social cost of CO2", Nature 610(7933): 687-692.

United Nations Environment Programme (2023), Emissions Gap Report 2023: The Closing Window.

Authors' note: The views expressed are those of the authors and do not necessarily reflect those of the International Monetary Fund. This article was originally published on VoxEU.org.

Unlocking the power of ideas



The history of human progress has been defined by technological breakthroughs generated by ideas. Christine Lagarde argues we need the right conditions that allow them to reach their full potential

Tucked away in the Sterling Memorial Library in the heart of Yale's campus lie the papers of America's first diplomat, Benjamin Franklin. Franklin was many things – ambassador to France, scientist, inventor, writer and publisher, to name a few – but above all, he was a man of ideas. As a young man, Franklin understood the power of ideas.

"All our ideas are first admitted by the senses and imprinted on the brain, increasing in number by observation and experience," he wrote. *"There they become the subjects of the soul's action."*[1]

By inspiring action, ideas can help us grow. This might be personal growth – a student's learning, say, allowing them to make the right decisions throughout their future career. But it holds at the societal level too: ideas help push our economies forward.

In recent decades, we had few barriers globally to the flow of ideas. Advanced economies shared their technologies with emerging ones, and emerging economies shared their cheaper input costs with us – the process we knew as 'globalisation'. But in recent years, the global economic order as we know it has been changing.

We now see that previously emerging economies are taking leadership in some advanced technologies. And we are seeing globalisation go into reverse, threatening access to the resources on which advanced technologies depend. So, how do we all prosper in this new world? I will argue that the key ingredient for our prosperity remains the same as ever: generating and sharing new ideas.

But history tells us that ideas can only drive growth if we first create the right conditions that allow them to reach their full potential – and if we are committed to breaking the bottlenecks that stand in their way. This is the challenge we all face today to thrive in this new world. And today, I will focus on what this challenge means for our economies and, in particular, for Europe.

The power of ideas across history

The history of human progress has been defined by technological breakthroughs generated by ideas. But ideas do not immediately translate into economic prosperity. Take Johannes Gutenberg's printing press – an ingenious device that combined metal prisms for moulding letters with an oil-based ink and techniques found in wine production[2].

What truly unlocks growth is when these three forces combine: when ideas translate into innovations, innovations diffuse into productivity growth, and our societies have the necessary ambition to remove any barriers that are in the way

By reducing the cost and increasing the speed at which books were produced, the printing press unleashed a communications technology that would revolutionise our world. In fact, an original Gutenberg bible is on display in the beautiful setting of the Beinecke Rare Book and Manuscript Library in Yale. But the printing press arrived at a time when literacy rates were still exceptionally low – around 9% in Gutenberg's native Germany[3].

Its ultimate benefits depended on rising literacy rates in the centuries that followed, with cheaper and more plentiful books also lowering the costs of learning. Countries that were quicker to embrace literacy reaped the gains in higher rates of economic growth and GDP per capita – a correlation that persists to this day[4]. In more recent centuries, we can identify three conditions that need to be in place for ideas to reach their full potential: translation, diffusion and ambition.

Translation means the ability to translate ideas into socially usable projects. And history has shown us that this ability depends on having the right economic ecosystems in key areas like finance and the supply of inputs. For example, until the turn of the 17th century, the ability to fund new ideas was severely limited by underdeveloped financial markets. One factor that helped change the game was the emergence of the modern joint-stock, limited-liability company around this time[5].

Suddenly, large pools of capital could be raised to fund bold proposals, such as expanding global shipping routes from east to west, which facilitated supplies of inputs. Countries that embraced joint-stock companies tended to experience faster growth[6].

If the right economic ecosystem infrastructures can facilitate ideas, the reverse is also true. The pioneering rollout of railroads across the US continent proved revolutionary in spurring the development of the country's capital markets[7].

But for ideas to be truly impactful at the macro level, there also needs to be diffusion. Technologies need to spread through an economy and become widely used. History suggests that a key factor in the diffusion of ideas is scale: that is, operating in a large, integrated market. Scale encourages firms to adopt new technologies, so that by expanding their production they can achieve lower unit costs.

The clearest example of the impact of scale is in the United States. While its constitution brought together thirteen disparate colonies, the country's economic trajectory would ultimately depend on how that constitution – in particular its Commerce Clause - was interpreted.

A pivotal moment occurred in 1824, when the Supreme Court's decision in Gibbons v. Ogden asserted the right of Congress to regulate interstate commerce and, in effect, to override state-granted monopolies that risked fragmenting the US market. This decision helped establish a truly nationwide economy and allowed the ideas of US entrepreneurs to spread and flourish. By several estimates, US GDP per capita at least doubled between 1800-20 and 1820-40[8].

In many of these cases, however, change did not happen by itself. It happened because of the ambition of entrepreneurs, economists, jurists or policymakers, their courage in overcoming hurdles to progress, and their ability to inspire others to follow their vision. But the nature of this ambition always evolved with the times.

In the 1800s, remote states scattered across the United States needed visionary entrepreneurs like Cornelius Vanderbilt, whose railways helped unify the country's economy. But as railroad tycoons established monopolies that undermined the public good, it took the ambition of policymakers like Theodore Roosevelt to break them up and foster competition.

What truly unlocks growth is when these three forces combine: when ideas translate into innovations, innovations diffuse into productivity growth, and our societies have the necessary ambition to remove any barriers that are in the way.

The power of ideas today

This brings me to the present day. As our economies grow, the relative importance of the different forces that drive growth changes[9]. For emerging economies that are far away from the technological frontier, deploying first their labour and later capital can help them to catch up.

But once economies mature and become advanced, productivity increases are mostly what propels us forward. And productivity is above all about ideas.

Most advanced economies, however, have seen productivity decelerate for some time. This slowdown led to a debate in the 2010s between techno-pessimists, who believed that most groundbreaking ideas were behind us, and techno-optimists, who believed that we were on the cusp of a new technological revolution.

Developments in recent years suggest that the case for optimism was stronger. Just as in Gutenberg's time, new revolutionary technologies like artificial intelligence (AI) and robotics are on the verge of transforming our societies. One study finds that generative AI alone has the potential to add up to almost USD 4.5 trillion annually to the global economy, roughly 4% of global GDP[10].

The good news for global productivity growth is that we see these new ideas flourishing across major economies, a direct legacy from the common ties that were crafted during the era of globalisation. And Europe, in contrast to what some may believe, is actually well placed to benefit from these ideas.

The European Union accounts for around one-fifth of the world's most-cited publications, patents, and research – despite making up less than 7% of the global population[11] – and this innovative activity includes key sectors such as AI and machine learning. According to one study, Europe draws in more AI talent than the United States, with over 120,000 active AI roles, and last year, Europe accounted for one-third of total early-stage capital invested in AI and machine learning across the two economies[12].

Our region also has many innovative companies in other high-tech sectors. Europe's manufacturing firms often operate at the global frontier, be it in producing photolithography machines for advanced chips or industrial robotics. In fact, Europe's share of the market for such robots is double that of China and more than thirtyfold that of the United States[13].

And many of Europe's most successful companies are not even listed. Of the 2,700 'hidden champions' worldwide – that is, small and medium-sized enterprises that are global leaders in their niche markets – more than half are found in Germany, Austria and Switzerland[14].

But as globalisation recedes and technological change accelerates, all economies are facing bottlenecks in transforming these ideas into sustained productivity growth. And these bottlenecks are in the same three areas that have been critical to unlocking the potential of ideas throughout history: translation, diffusion and ambition. So, the question we face is: how can we break these bottlenecks?

Breaking the bottlenecks

Translation

Let me start with the first bottleneck, translation. To translate new ideas into marketable projects, we need economic ecosystems that are suited to the specific requirements of today's technologies. We need financial systems that allow us to invest massively in innovative firms.

Sectors like AI, for example, need a lot of cash upfront to build up computing power and server capacity. According to industry leaders, the cost of training AI models is set to jump tenfold in the space of a year, and could soon rise to between USD 5 and 10 billion[15].

And we need secure access to a wide range of natural resources. The International Energy Agency estimates that training a single AI model uses more electricity than 100 US households consume in an entire year[16]. And as we electrify our transport systems and invest in renewable energy technologies, global demand for rare earth elements may increase three to sevenfold between now and 2040[17].

So, all our economies need to be proactive in ensuring that we have these ecosystems in place. But in Europe we face two specific challenges. First, we have a large financial sector, backed by high rates of saving from European households. But intermediation mainly takes place through bank lending rather than capital markets, which issue bonds and equities.

Bank lending works well for established companies that are relatively low risk and have generous collateral, such as our traditional manufacturing leaders. But it works less well for young, high-risk companies that typically drive radical innovation.

Innovative companies need access to ample risk capital, which requires a large venture capital sector that can back them until they go public. But the availability of risk capital is around ten times lower in Europe than in the United States[18], meaning that even firms that find backing at the early stage have less support when they enter the growth stage. The average venture capital-backed company in the EU receives about five times less backing than its US peers over its life cycle[19].

This gap often means that European entrepreneurs have to go overseas to get the financing they need – and sometimes their ideas go with them. And it is a key reason why, last year, Europe invested just USD 1.7 billion in generative AI compared with USD 23 billion of US venture capital and private equity[20].

Second, we are not endowed with significant natural resources in Europe, meaning that we depend heavily on imports[21]. And this dependency leaves us vulnerable in a less globalised world and a changing geopolitical landscape.

The brutal Russian invasion of Ukraine, which led to an almost complete shut-off of gas supplies to Europe, shows what is at stake. Even though we have successfully replaced Russia as a supplier, that process has left our firms at a notable cost disadvantage.

Before the pandemic, electricity costs for European firms were 1.7 times higher than those in the United States and 1.2 times as high as China. Now, that gap is 2.5 and 2.3 times respectively. In both cases, however, Europe is creating solutions in response to these constraints. As the former French President, Valéry Giscard d'Estaing, is reputed to have said, *"We may not have oil, but we have ideas."*

Where we can, we are acting to build the ecosystems we need internally. Europe's leaders have agreed to push forward with developing Europe's capital markets union, with a strong focus on improving the conditions for the financing options of European scale-ups[22]. We are also frontloading investment in renewables, which will ultimately make us more energy independent, although this process will take time and we will need to be realistic.

In the interim, we may need to depend even more on countries that have the necessary resources. For example, 80% of the global supply for rare earth metals currently comes from just three countries[23]. But we are also working

together with our friends and allies who face similar bottlenecks, like the United States, to make our supply more diversified. For example, the EU intends to establish a Critical Raw Materials Club, inviting partners with similar geopolitical and economic security concerns to join in the pooling of investments[24].

Diffusion

But once ideas are commercialised, they need to be diffused. Remember that what drives long-term growth is not only innovation by superstar firms, but also that innovations spread widely to less productive ones. Historically, one of the strongest drivers of technology diffusion has been free trade, especially between our two economies. For example, analysis points to a lag of three to four years between innovations in US industry and those in European industry[25].

But research suggests that diffusion has slowed across advanced economies in recent decades[26] – a trend that may partially reflect the nature of the digital economy itself, which tends to create 'winner-takes-the-most' markets[27]. And in Europe's specific case, slow diffusion also reflects the fact that, unlike the United States, we have not yet fully unlocked our innate scale as a continental-sized economy.

We have developed a business model in Europe that is unusually reliant – for a large economy, at least – on selling to other large economies, including capital goods that enable them to exploit their own scale. More than a third of our manufacturing GDP is absorbed outside the EU, compared with around a quarter for China and just a fifth for the United States[28].

But we have not made full use of our own scale to encourage our companies to adopt more technology. We are home to over 445 million consumers and 23 million firms[29] and yet our internal market remains fragmented, especially for services[30]. Intra-EU trade in services accounts for only about 15% of GDP compared with over 50% for goods[31].

This untapped potential is costing us dearly in terms of foregone growth and productivity gains. Remaining trade frictions in the EU mean that we are leaving around 10% of potential EU GDP on the table, according to one estimate[32].

And it is also affecting our competitiveness. We now see that other major economies are using their combination of technology and scale to push ahead faster in key sectors. China may now be leading in 37 of 44 critical technologies including electric batteries, hypersonics and advanced high-frequency communications such as 5G and 6G[33].

But Europe is also acting on this front to lift its constraints. Europe's leaders welcomed a major new report on the Single Market, calling for removing the remaining barriers in the crossborder provision of services as well as a 'policy shift' to reflect the new geopolitical and competitive environment[34].

And here again, Europe and the United States have shared interests in working together, especially in ensuring a level playing field between countries that play by the rules, while acting robustly in instances where rules are being broken to create an unfair advantage[35].

In other words, we should not become engaged in a subsidy race between our economies, which creates a zero-sum game. We should instead ensure that we use our collective weight in international trade to discourage others from anti-competitive practices, while increasing the free flow of ideas amongst ourselves – a positive-sum game.

Ambition

Will we be able to achieve all this? Ultimately, it is a question of ambition – and that is the final bottleneck we will have to break. In recent years, leadership has often been *reactive* in nature. This has been somewhat understandable in an era of 'permacrisis' – in which one shock, like the pandemic, is quickly followed by another, such as the outbreak of war.

But reactive leadership is no longer enough. Crises are becoming ever more global, requiring unprecedented levels of coordination across several sectors of society. And at the same time, the world is moving in directions that make such cooperation more difficult.

That is why we need *proactive* leadership – where we define the flow of events instead of simply responding to them. And for this we need to be far more ambitious.

The history of Europe gives us many examples of how effective such leadership can be. In the 1950s, an era marked by supply shortages and rationing, Europe started building common supply chains and pooling the production of inputs such as coal and steel.

In the mid-1980s, when Europe had exhausted the potential of what was then its common market, it forged ahead by creating the Single Market and reinvigorating growth. And in the 1990s, when exchange rate volatility threatened the stability of our currencies, we pushed forward with our monetary union to anchor our Single Market.

In doing so, we achieved what many had once thought impossible, and progressively united a continent that had been torn apart by two world wars.

When I look across advanced economies today, I am confident that our leaders understand what is required of them. Both the CHIPS Act and Inflation Reduction Act in the United States are accelerating the take-up of new technology. And I have listed many initiatives in Europe that are in the works, while there are many more that I have not touched upon.

But focusing on Europe in particular, what gives me hope is that, unlike after the great financial crisis, both leaders and citizens are aligned on what needs to be done. We realise that we can no longer afford to see ourselves as a loose club of independent economies.

That perspective is outdated in a world that is fragmenting into geopolitical blocs centred around the largest economies. And we know that we need to start seeing ourselves as a single, large economy with predominantly shared interests. This change in perspective also calls for joining forces in more areas.

We face increasing demands on spending from ageing populations, the climate transition and a changing security environment that we will only be able to meet together. And if we do not, we will face some difficult choices between sustaining our social model, delivering on our climate ambitions and playing a leading role in global affairs.

By acting as a Union to raise our productivity growth, and by pooling our resources in areas where we have a tight convergence of priorities – like defence and the green transition – we can both deliver the outcomes we want and be efficient in our spending so that we do not have to make sacrifices elsewhere.

And while this approach may require breaking some long-established taboos, we say in French that *"nécessité fait loi"* – or necessity knows no law. Our citizens understand this reality, even in a context where populism is on the rise. We see in poll after poll that Europeans believe that acting together is the best route to prosperity and security.

Over two-thirds of EU citizens feel that the EU is a place of stability in a troubled word[36], more than three-quarters are in favour of a common defence and security policy[37], and eight out of ten agree that the EU needs to invest massively in areas like renewable energy[38]. And in the euro area, support for our single currency remains close to record levels[39].

So I am confident that the ambition of our policymakers and the will of our people are aligned, and that we will break the bottlenecks that are preventing us from reaching our potential.

Conclusion

The global economy finds itself at a turning point, with old realities being replaced by new uncertainties. But amid all this change, some things remain resolutely the same. It is by generating new ideas, and creating the conditions in which they can spread and flourish through our economy, that we can drive future growth.

To create those conditions, Europe needs to break key bottlenecks in translation, diffusion and ambition. This will not be easy. But for too long we have simply talked about these problems instead of solving them through concrete actions. As Franklin once put it, *"Well done is better than well said."*[40]

In the end, we have a simple choice to make: either we break these bottlenecks, or we let these bottlenecks break us. Given the sense of urgency, the support for action and the consensus on what Europe needs to do, I know which side I stand on. And I am confident we can succeed. ■

Christine Lagarde is President of the European Central Bank

Endnotes

1. Franklin, B (1725), "A Dissertation on Liberty and Necessity, Pleasure and Pain".

2. Manguel, A (1997), A history of reading.

3. "Literacy rate, 1475 to 2022" on Our world in data.

4. "Literacy rate vs. GDP per capita, 2022", on Our world in data.

5. Ferguson, N (2008), The ascent of money: a financial history of the world.

6. By extension, the embrace of the joint-stock company arguably facilitated what some economic historians have dubbed the "great divergence" between a dynamic Europe and the rest of the world. A short overview of the great divergence – a concept that has attracted debate among economic historians – is offered in Allen, RC (2011), The great divergence, Oxford University Press.

7. As recently discussed in Lagarde, C (2023), "A Kantian shift for the capital markets union", speech at the European Banking Congress, 17 November.

8. Valued in 1840 prices. See Weiss, TJ (1992), "U.S. Labor Force Estimates and Economic Growth, 1800-1860", in RE Gallman and JJ Wallis (eds.), American Economic Growth and Standards of Living before the Civil War, Chicago University Press.

9. As put forward in the Solow Growth Model.

10. McKinsey Digital (2023), "The economic potential of generative AI: The next productivity frontier", 14 June.

11. European Commission (2023), "Long-term competitiveness of the EU: looking beyond 2030", 16 March.

12. Atomico (2023), "State of European Tech 23".

13. European Commission (2023), "AI Watch - Evolution of the EU market share of Robotics", JRC Technical Reports, 14 April.

14. Simon, H (2012), Hidden Champions – Aufbruch nach Globalia: Die Erfolgsstrategien unbekannter Weltmarktführer, Campus Verlag, Frankfurt/New York.

15. The New York Times (2024), "Transcript: Ezra Klein Interviews Dario Amodei", 12 April.

16. IEA (2023), "Why AI and energy are the new power couple", 2 November.
17. IEA (2021), "The Role of Critical Minerals in Clean Energy Transitions", May.
18. Risk capital as a percentage of GDP. See ECB (2023), "The EU's Open Strategic Autonomy from a central banking perspective – challenges to the monetary policy landscape from a changing geopolitical environment", Occasional Paper Series, No 311, Frankfurt am Main, March.
19. European Investment Fund (2023), Scale-up financing gap, 12 September.
20. McKinsey Global Institute (2024), "Accelerating Europe: Competitiveness for a new era", 16 January.
21. For instance, the euro area imports around two-thirds of its energy, compared with just over one-fifth for the United States.
22. European Council conclusions, 17-18 April 2024.
23. China, South Africa and the Democratic Republic of Congo. See European Commission (2020), "Study on the EU's list of critical raw materials - Final report".
24. Council of the European Union, "An EU critical raw materials act for the future of EU supply chains".
25. International Monetary Fund (2015), "World Economic Outlook: Uneven Growth: Short-and Long Term Factors", April.
26. Calvino, F and Criscuolo, C (2022), "Gone digital: Technology diffusion in the digital era", Brookings Institution, 20 January.
27. Schnabel, I (2024), "From laggard to leader? Closing the euro area's technology gap", inaugural lecture of the EMU Lab, 16 February.
28. ECB (2023), op. cit.
29. Council of the European Union, "EU single market".
30. Around 70%.
31. Schnabel, I (2024), "From laggard to leader? Closing the euro area's technology gap", inaugural lecture of the EMU Lab at the European University Institute, 16 February.
32. in 't Veld, J (2019), "Quantifying the Economic Effects of the Single Market in a Structural Macromodel", Discussion Paper Series, No. 94, European Commission, February.

33. Australian Strategic Policy Institute (2023), "ASPI's Critical Technology Tracker - The global race for future power", 22 September.
34. European Council conclusions, op. cit.
35. European Commission (2023), "Commission launches investigation on subsidised electric cars from China", 4 October.
36. Eurobarometer (2023), "Standard Eurobarometer 100 - Autumn 2023".
37. Eurobarometer (2023), "Standard Eurobarometer 99 - Spring 2023".
38. Eurobarometer (2023), "Standard Eurobarometer 100 - Autumn 2023".
39. Eurobarometer (2023), "The euro area".
40. Franklin, B (1737), "Poor Richard's Almanack".

This article is based on a speech by delivered at Yale University in New Haven, USA, 22 April 2024.

Reforming EU innovation policy

Europe lags behind in innovation. Clemens Fuest, Daniel Gros, Philipp-Leo Mengel, Giorgio Presidente and Jean Tirole argue that EU innovation policy should support disruptive innovation to compete

That Europe is lagging in innovation has been diagnosed for a long time. More than a decade ago, the EU launched the Innovation Union, and increasing expenditure on R&D to 3% of GDP has been an official goal since the launch the Lisbon Strategy in 2000. However, gross domestic expenditure on R&D in the EU is still below 2% of GDP, lower than in other major economies such as the US, Japan, and China.

The reason why the EU lags behind other regions is not that governments (national and EU) spend less on R&D than its rivals. In 2020, government-funded R&D amounted to €110 billion in the EU (mostly by national governments) and €150 billion in the US, accounting for a very similar percentage of GDP (around 0.7%) and higher than in many other regions of the world.

The key reason for the overall transatlantic difference is the lower engagement in R&D by the business sector, whose spending amounts to only 1.2% of GDP in the EU, versus 2.3% of GDP in the US. These often-cited OECD figures, however, do not allow for a sectorial breakdown.

To analyse in more detail the sectoral composition of R&D, in our recent paper (Fuest *et al* 2024), we use data from the EU Industrial R&D Scoreboard, which are based on the accounts of the 2,500 largest companies in the world in terms of R&D spending[1].

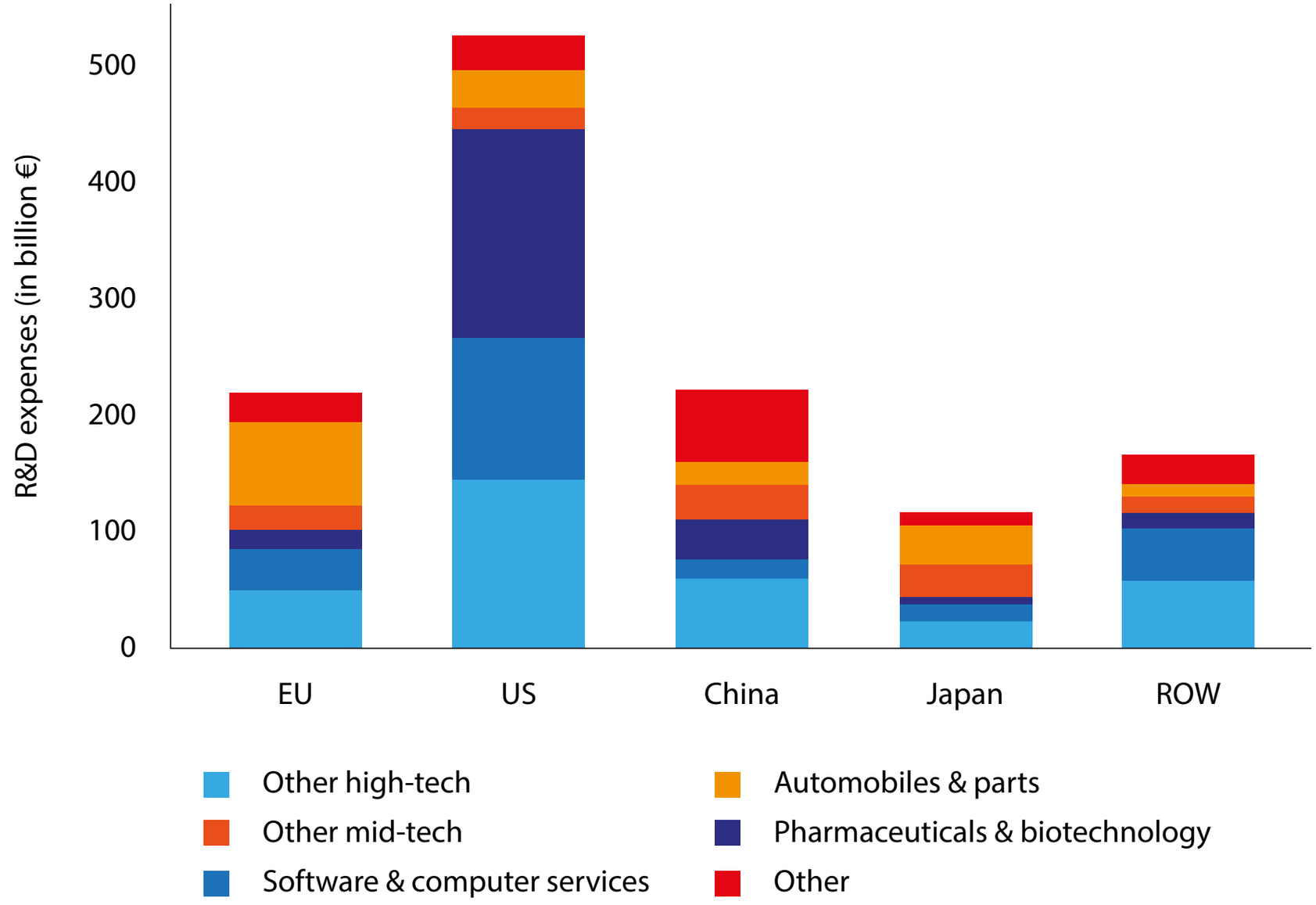
Europe's middle technology specialisation

Figure 1 shows the sectoral composition of business R&D spending (BERD) in nominal terms for businesses headquartered in the four regions plus a residual, the rest of the world (ROW). In the US, high-tech industries – mostly software & computer services and pharmaceuticals & biotechnology – account for 85% of BERD.

In the EU, by contrast, mid-tech industries – especially automobiles & parts – account for roughly 50% of BERD, a much higher share than in the US[2]. The sectoral composition of corporate R&D spending by EU-headquartered firms is more similar to that of Japan and China than the US.

The European pendant to DARPA was supposed to be the European Innovation Council (EIC), created in 2021 with the aim of supporting disruptive innovations

Figure 1. BERD by technology level, top 2,500 companies



Source: Industrial R&D Investment Scoreboard (2023).

Not surprisingly, high-tech industries are much more R&D-intensive than mid-tech industries. Therefore, the larger share of high-tech industries in the US is a key factor explaining why BERD is so much higher than in other economies.

What is more, evidence suggests that public-sector support is more likely to crowd out business R&D in low R&D-intensity industries (eg. Marino *et al* 2016, Szücs 2020), which might explain the low business-sector multiplier in the EU relative to the US[3,4].

Europe's middle technology specialisation is permanent – a trap?

Table 1 shows the top three R&D spenders and their industries over time as a further illustration of the diverging development across the Atlantic. It gives the top three companies in terms of R&D spending and their respective industries over the last 20 years in the US, EU, and Japan[5].

In the US, Microsoft is the only company appearing more than once among the top three R&D spenders. Meanwhile, in the EU and Japan, Volkswagen (VW), Mercedes, and Toyota remain in the top three over the 20 years, while Panasonic, Bosch, and Honda appear at least twice.

Interestingly, in the US two of the three top R&D spenders in 2003 were also in the automotive industry, but this changed over time. The software industry became increasingly important over the years; by 2022, all top-three spenders produced software.

In the EU and Japan, the auto industry tended to dominate throughout the 20-year period. These patterns are consistent with the literature on path dependence in innovation and industrial specialisation (eg. Acemoglu, 2023, Aghion *et al* 2021, Aghion *et al* 2016)[6,7].

Table 1. Top three R&D spenders and their industries compared over time

	2003	2012	2022
US	Ford (auto)	Microsoft (software)	Alphabet (software)
	Pfizer (pharma)	Intel (hardware)	Meta (software)
	GM (auto)	Merck (pharma)	Microsoft (software)
EU	Mercedes-Benz (auto)	VW (auto)	VW (auto)
	Siemens (electronics)	Mercedes-Benz (auto)	Mercedes-Benz (auto)
	VW (auto)	Bosch (auto)	Bosch (auto)
JPN	Toyota (auto)	Toyota (auto)	Toyota (auto)
	Panasonic (electronics)	Honda (auto)	Honda (auto)
	Sony (electronics)	Panasonic (electronics)	NTT (telecom)

Source: Industrial R&D Investment Scoreboard (2004, 2013 and 2023).

EU (and Japanese) industry thus failed to transition to high-tech sectors. One reason might that the incentive to do so was much lower in Europe, where the profit margin of high-tech industries was only about 3 percentage points higher than mid-tech ones, whereas in the US the difference in profit margins between high-tech and mid-tech industries was about 7 percentage points (Redding and Melitz 2021). The incentive to allocate capital to high tech firms was thus much higher in the US than in Europe.

It is possible that the higher profit margins of US high-tech firms at least partially reflect the near-monopoly position of US software giants in their respective markets. But this does not alter the fact that the availability of higher profit margins for US firms presented a strong incentive to invest in these industries.

R&D-intensive industries can be considered natural oligopolies, in which a few market leaders emerge, sustained by the dynamics of large market shares fuelling R&D, which in turn sustain large market shares in a virtuous cycle leading to dominant positions[8]. In these industries, sales and R&D expenditures follow a similar pattern (Sutton 2007).

The evolution of profits in our data reflects these patterns of natural oligopoly formation. The initial advantage of the US in high-tech was magnified over time, whereas EU (and Japanese) industries remained in their specialisation pattern. Breaking this path dependency justifies public-sector intervention to provide the seeds for an alternative model of specialisation.

How to break out: fostering innovation

In the US, the Defense Advanced Research Projects Agency (DARPA) is widely credited as having played a crucial role in fostering the emergence of high tech, including such pivotal innovations as the internet.

ARPA, as it was called initially, was created in response to the 'Sputnik shock' of the late 1950 to support, as the name suggests, advanced research projects that are not of commercial interest because their significance might reveal itself only later. The selection of the projects to be financed is left to the Agency that employs by now close to 100 highly qualified programme managers.

This model of supporting advanced research is not limited to the defence sector, there exist now ARPAs for energy health and artificial intelligence[9].

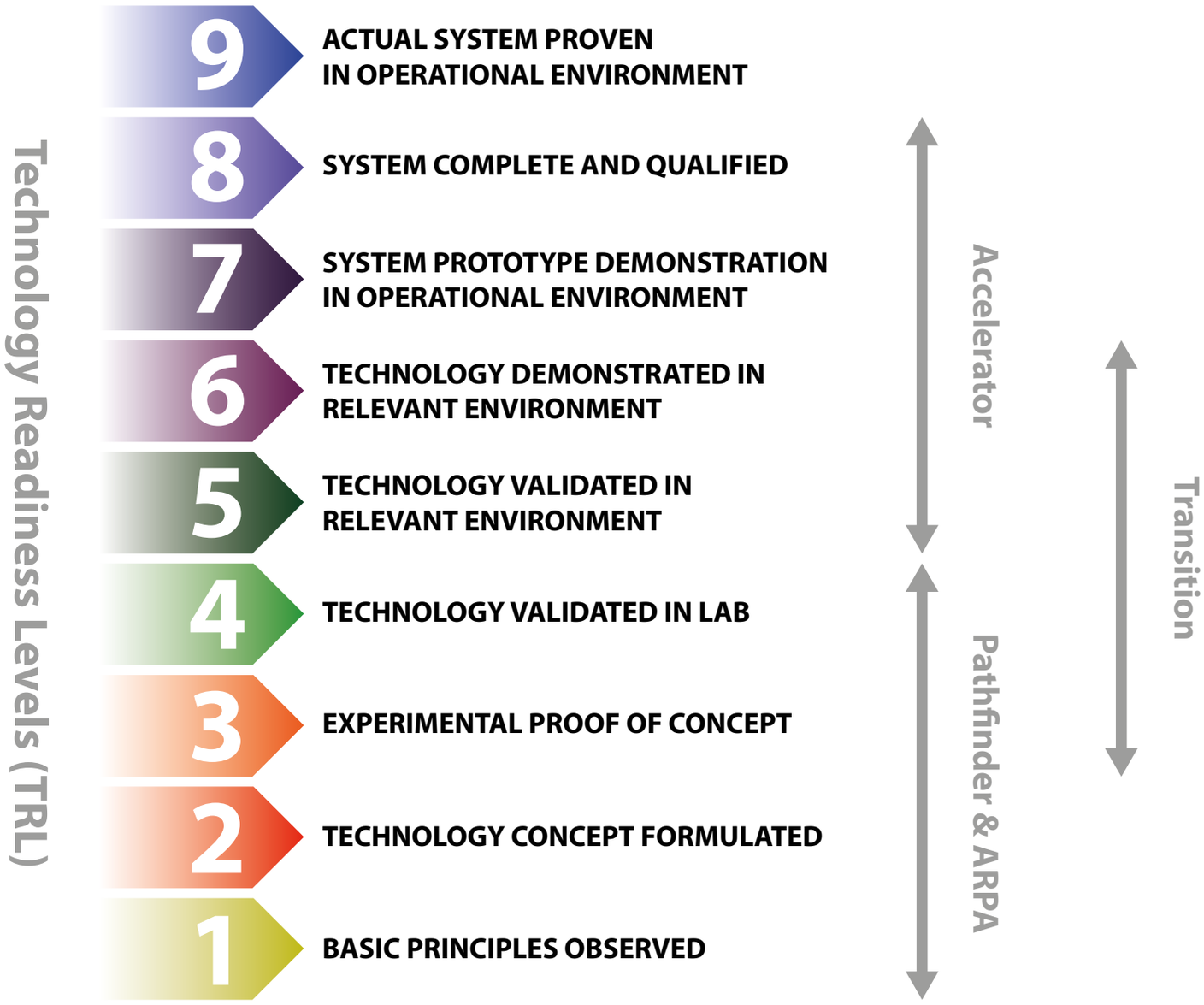
The European pendant to DARPA was supposed to be the European Innovation Council (EIC), created in 2021 with the aim of supporting disruptive innovations. The name 'Council' is actually misleading since the EIC consists essentially of three separate programmes, called Pathfinder, Transition, and Accelerator.

As the headings suggest, Pathfinder finances projects at their early stage, whereas the purpose of Accelerator is to 'accelerate' the commercial application of emerging technologies and support the growth of start-ups. The annual budget of the EIC is about €1 billion (against about $4 billion for DARPA alone).

Similar to the ARPAs, the details of EIC calls are set by programme managers within the overarching objectives of the European Commission. Programme managers also determine the specific goals of the individual projects and group them in thematic portfolios. Still, despite efforts to emulate the salient features of the ARPA model, the EIC falls short in at least two key respects.

A first key aspect, the type of projects financed by the EIC, can best be explained using a technology readiness level (TRL) indicator, described in Figure 2. This indicator goes from 1 (only basic principles observed) to 9 (actual systems in operation).

Figure 2. Technology readiness levels



Source: Authors' representation based on official sources.

ARPAs typically focus on developing 'proof-of-concept' (Azoulay *et al* 2019) or projects up to TRLs 3-4 at most. Once projects reach a sufficient maturity, usually taken to be the demonstration stage (TRL 5 or above), they 'graduate' and leave ARPAs with the expectation that private capital will flow and scale them up.

Azoulay *et al* (2019) position ARPA-funded projects on the initial flat part of the innovation S-curve, relating research effort and technical progress (Foster 1986)[10]. On the initial part of the curve, a high degree of effort results in very limited performance gains, and delayed payoffs limit incentives to pursue the project. This is where public-sector support is most needed because it addresses a clear market failure.

Instead, about two thirds (€700 million) of the annual budget of the EIC goes to the Accelerator progamme that finances projects with TRLs above 5. This is expressed in the official task of the EIC to *"support disruptive innovations throughout the lifecycle from early stage research, through to the financing and scale up of start-ups and SMEs."*

The EIC thus has a dual mission not only to support disruptive innovation, but also to finance scale-up and SMEs. It is thus not surprising that the management of the EIC programmes is housed in the former EU executive agency for SMEs, renamed the European Innovation Council and SMEs Executive Agency (EISMEA). Given the large share devoted to high TRL products, the EIC thus has only about €300 million for ARPA type projects, one tenth of DARPA alone.

A second key difference concerns the selection of projects and their management. First, the selection process is still politically controlled, which is in conflict with the best international standards (ARPA structures here, and also NSF, NIH and the EU's own ERC for fundamental research).

Second, the EIC has only a very small number of project managers, each of which oversees dozens of projects. This means that each project manager has to deal with projects that are outside her areas of expertise. Moreover, only

about one half of the Board of the EIC is composed of scientists and engineers that might have the qualification to find the most promising projects.

Based on this analysis we propose two approaches to improve EU support for innovation that do not require an increase in the EU budget:

1. Better management. Reform the governance of the EIC, hiring a larger number of independent and highly qualified programme managers and giving them greater discretion over project selection and management.

2. Better use of budgetary resources toward disruptive research, which currently accounts for a paltry 5% of total funding. Scale down existing under-performing programmes like the European Institute of Innovation and Technology (EIT) and the European Innovation Ecosystem (EIE).

Replace the financing of equity stakes from the Accelerator budget with other sources whose mission is investment, rather than innovation. For example, the EIC could be merged with the European Investment Fund (EIF) or the proposed Sovereignty Fund. This would free up €0.41 billion per annum.

With this combination of management reforms and redirection of existing resources, Europe could create a much stronger structure to prioritise and boost game-changing innovations through a budget neutral restructuring, thus taking into account limitations for the overall EU budget. ■

ABOUT THE AUTHORS

Clemens Fuest is Professor of Economics and Public Finance at Bibliothek Wirtscharftswiseneschaften – University of Munich; Director of the Center for Economic Studies; Executive Director of Cesifo; Speaker at the European Network for

Economic and Fiscal Policy Research; and President of the ifo Institute; Daniel Gros is Director of the Centre for European Policy Studies, Brussels; Philipp-Leo Mengel is a PhD Researcher at Bocconi University; Giorgio Presidente is a Postdoctoral Researcher at the Oxford Martin School, University of Oxford; and Jean Tirole is Director of the Toulouse School of Economics.

Endnotes

1. The data are taken from the EU Industrial R&D.

2. For the purpose of this exercise, we have used three broad categories: high- and mid-tech plus the remainder, 'other', mostly including services and utilities. Our classification is similar to that adopted by Eurostat and the OECD.

3. One reason might be that R&D-intensive industries need resources far exceeding the typical amounts of a grant.

4. In the EU, €1 of public-sector spending is associated with €2 spent by the private sector. In the US, the private-sector multiplier is equal to 3.

5. We do not include China because some companies there have changed their reference industry over the years.

6. Typically, in these models increasing returns to scale resulted in past advances (in a given sector or technology) facilitating further advances in the same sector.

7. These patterns are also consistent with evidence of declining business dynamism around the world (eg. Akcigit 2024, Biondi et al 2024, Decker et al 2020), but analysing that aspect goes beyond the purpose of this study.

8. We wish to thank Michele Polo for pointing this out.

9. They are called ARPA E-ARPA H and ARPA-I.

10. The metric of technological progress depends on the technology considered, such as kilowatt-hour of electricity generated, or computational speed.

References

Acemoglu, D (2023), "Distorted innovation: does the market get the direction of technology right?", AEA Papers and Proceedings 113.

Akcigit, U (2024), "Economic Renaissance: Unleashing Business Dynamism for Economic Growth".

Aghion, P, J Van Reenen, R Martin, D Hémous and A Dechezleprêtre (2012), "Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry", CEPR Discussion Paper 9267.

Azoulay, P, E Fuchs, AP Goldstein and M Kearney (2019), "Funding breakthrough research: Promises and challenges of the 'ARPA Model'", Innovation Policy and the Economy 19(1): 69-96.
Biondi, F, S Inferrera, M Mertens and J Miranda (2024), "Declining business dynamism in Europe: The role of shocks, market power, and responsiveness", VoxEU.org, 21 March
Decker, RA, JC Haltiwanger, RS Jarmin and J Miranda (2020), "Changing business dynamism and productivity: Shocks versus responsiveness", American Economic Review 110(12): 3952-3990.
Foster, RN (1986), "Working the S-curve: assessing technological threats", Research Management 29(4): 17-20.
Fuest, C, D Gros, P-L Mengel, G Presidente and J Tirole (2024), "EU Innovation Policy: How to Escape the Middle Technology Trap", Report by the European Policy Analysis Group, Institute for European Policymaking at Bocconi University.
Marino, M, S Lhuillery, P Parrotta and D Sala (2016), "Additionality or crowding-out? An overall evaluation of public R&D subsidy on private R&D expenditure", Research Policy 45(9): 1715-1730.
Redding, S and M J Melitz (2021), "Trade and innovation", VoxEU.org, 28 July.
Szücs, F (2020), "Do research subsidies crowd out private R&D of large firms? Evidence from European Framework Programmes", Research Policy 49(3): 103923.
Sutton, J (2007), "Market structure: theory and evidence", Handbook of Industrial Organization 3: 2301-2368.

This article was originally published on VoxEU.org.

Not a ‘side dish’

Industrial policy is undergoing a major resurgence. Cristina Caffarra and Nathaniel Lane argue that getting Europe to improve performance will require a massive, concerted effort at national and EU levels

Speaking recently to an audience in Brussels (at an event that has since become known as the 'Anti-Davos')[1] the European Commission's top competition enforcer provocatively remarked that, *"[w]hen it comes to the big issues of our times, I am afraid competition policy is neither the problem, nor the solution – it's a side dish."*

The statement triggered a cascade of reactions and responses. Weeks away from a European election that may change Europe's political landscape, in this column we discuss why competition policy has a broader role than the comment credits it for. Far from being a 'side dish', antitrust will be important to the transformative role that resurgent industrial policy will need to play, particularly in Europe.

Beyond digital regulation

Digital enforcement ('taming Big Tech') has assumed totemic value in Europe: it is the test of our credibility, resolve, and effectiveness in confronting surveillance business models and the entrenched market power of Big Tech. As antitrust enforcement in digital has failed, regulation is now the beach on which we fight.

If Europe can achieve results here, it shows the world what can be done. But digital enforcement is also getting too much attention in relation to the scale of Europe's real structural problems. While we pride ourselves of our efforts to 'tame Big Tech', policymakers must urgently confront Europe's underwhelming economic performance across sectors.

Europe is falling behind on multiple fronts: productivity, competitiveness, R&D, IT investments, and more. This view is not controversial: many have been sounding the alarm, from the ECB (Schnabel 2024) to the Bruegel think tank with their recent report for the Commission's EGov Directorate (the 'Bruegel EGOV report'; Pinkus *et al* 2024)[2].

Digital enforcement will do its thing, and it is underway. Yet on the eve of a European election, deep structural problems are urgently on the table for the incoming 2024-29 Commission. The European Commission has tasked two former Italian prime ministers, Mario Draghi and Enrico Letta, to report on competitiveness and progress towards the Single Market, respectively.

Europe is falling behind on multiple fronts: productivity, competitiveness, R&D, IT investments, and more. This view is not controversial: many have been sounding the alarm

Their reports will inevitably discuss the causes of Europe's fragmentation – its multitude of languages, cultures, rules, financial markets, capital markets, and economic trajectories – and what can be done to reduce those barriers and address our 'competitiveness crisis' (Schnabel 2024).

The reports are also expected to highlight the need for massive ramp-up of investment in multiple strategic sectors, to promote green transition and digitalisation, to increase Europe's resilience in the global economy, and curb de-industrialisation. The future demands and scope of these investments will be unprecedented.

While Europe has experience with large-scale government spending, both historically and with the pandemic and its aftermath (eg. the EU Recovery and Resilience Facility), calls for state-led investment are now much more expansive. The confluence of post-pandemic emergencies and the waning of the Washington Consensus have created a further sense of urgency in two dimensions.

First, a global resurgence of economic interest in the design of appropriate industrial policies (eg. Juhász *et al* 2022); and second, a significant pivot in the US away from the traditional aversion to the state playing a role in markets (Armstrong and Wu 2024). What does this all mean for Europe, and the role of competition policy?

The 'New Industrial Policy'

Industrial policy has returned as a major object of interest, with a proliferation of new thinking over the last five years by academics and practitioners (Rodrik *et al* 2023). Questions around industrial policy have turned from 'whether' (ie. 'should governments carry out industrial policy?') to 'how' ('how should industrial policy be carried out?').

A recent wave of research, the 'New Economics of Industrial Policy', has generated more nuanced views, and nascent work is tempering historical concerns that industrial policies are necessarily harmful: because 'losers pick governments', and state support necessarily produces 'zombie companies' – inefficient national champions.

While these risks are real, recent empirical work has also established key episodes where industrial strategies likely *"shifted resources in the desired direction, often producing large long-term effects in the structure of economic activity"* (Juhász *et al* 2023).

A diverse community of industrial policy thinkers is coalescing around critical themes. There are undoubtedly differences in the 'how' (from Mazzucato 2020, urgent advocacy for mission-oriented 'moonshots', to more mainstream economic theories of intervention in Juhász *et al* 2023), but there are also key commonalities: the importance of focusing on strategic sectors, the need to go beyond blunt instruments of post-war policy, and a focus on collaborative and deliberative policymaking, with input from the private sector.

Scholars are also emphasising the importance of averting government failures through the design of guard rails and conditionalities (Lane 2021, Mazzucato and Rodrik 2023).

What makes the rethink all the more salient is the big 'pivot' of the current US administration towards 'industrial strategy', with large public funds being allocated and disbursed to support a variety of goals: green transition, rebuilding domestic capacity offshored in the neoliberal era, supporting deindustrialised areas, reducing dependency on concentrated and brittle supply chains, and 'crowding in' complementary private investments.

Foroohar (2024) argues this does not yet amount to a fully coherent industrial policy, but we would be inclined to be indulgent given the magnitude of the pivot.

‘Antimonopoly’ thinking as foundational

The focus on averting the misadventures of past industrial policy (in particular, support for ‘national champions’) is an important reason why antitrust thinking has a major role to play in the new landscape.

In the US, the worlds of industrial policy and antitrust have recently been colliding. With the major shift in antitrust thinking in Washington over the past few years has come recognition that antimonopoly values (fairness, equality, citizenry) must pervade and motivate other economic policy tools – including trade an industrial policy.

As the Chair of the Federal Trade Commission, Lina Khan, recently articulated, *“we are hearing arguments that America must protect its domestic monopolies to ensure that we stay ahead on the global stage. (…) we should be extraordinarily sceptical of these arguments, and instead recognize that monopoly power is a major threat to America’s national interest.”*

Further, *“the choice to bring antitrust lawsuits against AT&T and IBM ended up fostering waves of innovation”* (Khan 2024). And yet further: *“competition policy will be a key complement to achieve industrial policy goals. As we’re handing out subsidies, are there going to be strings attached, that create trajectories on an open and competitive path, rather than a closed and monopolistic path? If the industrial policy vision is one of government as a more active participant in ‘market making’ and ‘market shaping’, we need to make sure that our values and our vision around competition policy are wholly a part of that decision making.”*[3]

Tim Wu, a key architect of Biden-era thinking on antitrust, also describes antitrust (and, in particular, past lawsuits breaking up monopolies) as ‘industrial policy’.

Indeed, where antimonopoly promotes intentional economic change, it is, by definition, *"an industrial policy"* (Juhász *et al* 2022). Making antimonopoly thinking part of the industrial policy toolbox can help break with the past: there is clear recognition among industrial policy scholars that where strategic investments are made, markets must remain 'oxygenated' – not favour dominant players; and that the more successful industrial policies are those which have supported competition (Aghion *et al* 2015, Nahm 2021).

Just like the 'efficiency paradigm' of the neoliberal era has been superseded in antitrust, efficiency goals may not sufficiently capture the broader aims of an industrial policy – for suppliers, regional economies, communities, citizens, and more.

'Antimonopoly', the fight against market power and its pathologies, is a fundamental value that must underpin also the direction of investments to lift entire sectors and communities. What may not be 'efficient' may have other social benefits.

The European predicament

Europe has responded to the pandemic and the 'polycrisis' also with a large increase in public spending initiatives, but we continue to face a large and widening gap in economic performance with the US and other blocks.

This reflects in part a deep structural problem of persistent fragmentation along national borders, which has been Europe's 'Sisyphean rock' for decades notwithstanding major past efforts (Pinkus *et al* 2024).

Confronting our declining economic performance will require a major increase in public spending in selected strategic sectors, which is hard for a collection of sovereign countries, with limited federal-level resources and persistent fears that common public spending could benefit some countries more than others.

The Bruegel EGOV Study suggests as a possible way forward what they describe as *"coordination for competitiveness"* – the European Commission performing a central coordination role to *"cooperate in areas that offer the greatest gains on a sector-by-sector basis, supported by some EU-level funding. Energy policy coordination and an EU Advanced Research Projects Agency (ARPA) are two examples."*[4]

What should competition policy do?

What should be the role of competition policy in designing these policies? Pushing back against 'national champions' is certainly not new in Europe, where European Commission state aid policy has been traditionally tasked to control excess spending by national governments, and their incentives to prop up their own 'zombie firms' with state funds.

State aid is a large part of Directorate-General (DG) for Competition, systematically vetting national schemes dreamt up by member states to support local interests, with the objective of avoiding distortions to 'competition in the Internal Market'. The traditional requirement for state support not to fall foul of state aid rules is that it 'addresses a market failure' in the 'most efficient way possible'.

It is thus unsurprising that in a recent contribution to the debate on the need for more industrial policy, senior DG Competition officials drawing from their state aids experience recommended that each industrial policy intervention be justified by a specific market failure, and adopted measures be 'efficiency-enhancing' (Piechucka *et al* 2023).

The paper mentions 'efficiency' over 100 times, 'efficiency-enhancing' over 30 times, and 'market failure' 80 times. While of course we want to avoid wasteful effort, this focus seems out of line with the evolution of current thinking both in antitrust and industrial policy.

For instance, efficiency criteria are at odds with the rationales driving policy discussions on place-based policies and 'good jobs', which are aimed to produce larger social benefits. Efficiency criteria are also empirically incomplete: how do we prove something is 'efficiency enhancing', particularly in the case of policies with long gestation periods and whose benefits are borne in the future (Lane 2020).

But more fundamentally, the usefulness of 'efficiency' as a principled goal has come deeply into question. As put by Deaton (2024), *"we valorize (efficiency) over other ends. Many subscribe to (the vision) that says economists should focus on efficiency and leave equity to others, politicians or administrators. But the others regularly fail to materialise, so when efficiency comes with upward redistribution – frequently though not inevitably – our recommendations become little more than a license for plunder."*

Simply put, extending traditional 'state aid' thinking to industrial policy is undesirable at a time when thinking and practice around interventions are evolving. We don't need to reassure ourselves we are being 'orthodox' by casting everything as a market failure (which is easy to do, in any event, if one tries – but so what?).

Nor is efficiency the 'north star' we need to be pursuing. We will need major increases in spending from the centre, and coordination of spending at the national level to ensure that collective objectives are not undermined. But if the aim is to build capacity and improve performance in Europe, industrial policy intervention that benefits citizens (not merely as consumers) cannot be held to a 'market failure/efficiency-enhancing' paradigm.

Competition insights and capabilities can and should be involved in industrial policy design to provide not just an assessment of state plans along traditional state aids lines, but also affirmative values of antimonopoly, de-concentration, fairness, and distribution.

Traditionalists will say 'but what is the limiting principle?' – this the generic objection to everything by those who want no change. We cannot be stuck with 'efficiency' when we need to accomplish so much more, and neoliberal 'trickle down' has been shown to be a chimera.

Urgent proof of concept: a digital industrial policy

Major focus needs to be placed on powering Europe's digital infrastructure. Europe has set huge store by its ability to 'tame Big Tech' via multiple laws: the Digital Markets Act (DMA), Digital Services Act (DSA), Data Act, and AI Act. Whether this effort will truly enable European challengers to acquire more than a marginal role remains to be seen.

But we need more than extracting from Big Tech concessions to provide better deals to app developers, search rivals, ad-tech companies and e-commerce sellers, and create access regimes to platforms that are now critical infrastructure. We also need to invest locally in an independent, federated, decentralised digital infrastructure on which Europeans can build.

Europe has lower advanced technology adoption than the US, and the productivity divergence between high- and low-productivity firms has widened more in digital-intensive sectors (Criuscolo 2021). On the positive side, the number of EU-based start-ups is high, and there is vitality in terms of researchers, digital skills, and emerging technologies (Meyer 2024). Yet Europe's fragmentation and its dependencies on US giants make it challenging to implement, commercialise, and scale hi-tech activity.

We thus need a robust digital industrial policy alongside the existing diet of EU regulation and innovation policies. This means coordinating national and EU-level efforts to create autonomous infrastructures and reduce the dependency on Big Tech. These goals also align with narratives on 'sovereignty'.

As summarised by Bria (2023), *"to strengthen our economic and political sovereignty in a complex geopolitical environment, Europe needs a combination of regulatory frameworks and active digital industrial policies. This objective goes beyond merely crafting regulations. It's about building new markets and industries, creating innovative institutions, and fostering ecosystems that truly serve the public interest."*

Investing in a 'Europe stack' tech ecosystem should be an attractive candidate for EU-level funding because the crossborder externalities are high. Bria (2023) suggests a €10 billion EU Digital Sovereignty Fund, which would *"blend grants and equity investments, fostering pan-European collaboration among our national innovation agencies (…) to establish robust digital public infrastructures and digital commons, offering viable alternatives to current monopolistic digital platform models, supporting the deployment of open AI models and decentralized applications, sovereign data spaces, open knowledge tools and content, privacy-preserving digital IDs, and digital payment systems."*

The Bruegel EGOV Study (Pinkus *et al* 2024) suggests an 'EU ARPA' involving the creation of an independent agency with a €5 billion budget to pool investment projects and coordinate spending at national level, to be topped up with additional funds from EU programmes. Objectives would be set by the EU Council and the European Parliament, but the agency would be autonomous in policy implementation.

While prior initiatives have proven insufficient for multiple reasons (for example, the Juncker Plan of 2015 and the European Fund for Strategic Investments), and significant obstacles remain – not least risk-taking appetite and competencies – we need to double down now that we have more scholarship, experience, and expertise.

Critically, the experience and expertise of DG Competition in digital markets will be critical here: successful investment in federated decentralised infrastructures requires understanding of the regulatory environment,

and of competitive dynamics which can facilitate private complementary investment and innovation on these infrastructures.

Overall, getting Europe to improve performance will require a massive, concerted effort at national and EU levels to identify strategic sectors and disburse funds in a targeted way that will crowd in private investment.

Competition thinking has a key role to play, not to enforce narrow and nebulous efficiency goals, but to ensure initiatives are consistent with antimonopoly values, fairness, and opportunities. Not a 'side dish'. ■

Cristina Caffarra is Honorary Professor at University College London, and Nathaniel Lane is Associate Professor and Post-doctoral researcher, Department of Economics, at Massachusetts Institute of Technology; and Associate Professor at the University of Oxford

Endnotes

1. https://www.politico.com/news/2024/02/16/how-some-of-the-worlds-most-powerful-regulators-are-trying-to-upend-the-economic-system-00141802.

2. The analysis of the causes of the gap in the Bruegel EGOV report includes: low labour and total factor productivity growth relative to the US especially since 2020, with large intra-EU differences; especially dramatic difference in productivity with the US in information and communication technology (ICT); slower accumulation of IT capital and better technology adoption in the US; much lower intensity of R&D spending especially in three key sectors – pharma/biotech, software and IT; significantly higher industrial electricity retail prices than the US; higher hourly labour costs; much higher cost of equity finance and lower volume of venture capital funding and therefore much greater restrictions in accessing risk capital; and finally, much lower trade across national borders than one would expect to see given past effort at market integration.

3. https://cepr.org/events/competition-policy-rpn-reinvigorating-antitrust-citizens-not-just-consumers.

4. The reference is to the US ARPA, which has been instrumental in mobilising resources and investing them in high-risk, high-reward projects.

References

Aghion, P, J Cai, M Dewatripont, L Du, A Harrison and P Legros (2015), "Industrial Policy and Competition", American Economic Journal: Macroeconomics 7(4): 1-32.

Armstrong, R and E Wu (2024), "Dani Rodrik: doing industrial policy right", Financial Times, 9 February.

Bria, F (2023), "Open, Sovereign, Independent AI: Europe's greatest challenge?", Speech at the European Parliament, 10 December.

Caffarra, C (2021), "What are we regulating for?", VoxEU.org, 3 September.

Caffarra, C (2023), "Europe's Tech Regulation Is Not an Economic Policy", Project Syndicate, 11 October.

Caffarra, C (2024a), "Antitrust and the political economy: Part 1", VoxEU.org, 5 January.
Caffarra, C (2024b), "Antitrust and the political economy: Part 2", VoxEU.org, 6 January.
Caffarra, C and B Kilic (2024), "Re-joining trade with antitrust", VoxEU.org, 7 March.
Criscuolo, C (2021), "Productivity and Business Dynamics through the lens of COVID-19: the shock, risks and opportunities", European Central Bank Forum on Central Banking, September.
Deaton, A (2024), "Rethinking My Economics", IMF Finance & Development, March.
Foroohar, R (2024), "America (still) has no industrial policy", Financial Times, 18 March.
Juhász, R, N Lane, E Oehlsen and VC Pérez (2022), "The Who, What, When, and How of Industrial Policy: A Text-based Approach", SocArXiv, 25 August.
Juhász, R, NJ Lane and D Rodrik (2023), "The New Economics of Industrial Policy", NBER Working Paper 31538.
Khan, L (2024), "America Has a Resilience Problem", Foreign Policy, 20 March.
Lane, N (2020), "The New Empirics of Industrial Policy", Journal of Industry, Competition and Trade 1(2): 1–26.
Lane, N (2021), "A Flight Plan That Fails", Boston Review, 15 September.
Mazzucato, M (2020), Mission Economy: A Moonshot Guide to Changing Capitalism, Allen Lane.
Mazzucato, M and D Rodrik (2023), "Industrial Policy with Conditionalities: A Taxonomy and Sample Cases", UCL Institute for Innovation and Public Purpose Working Paper 2023/07.
Meyer, Z (2024), "Helping Europe's Digital Economy Take Off: An Agenda for the Next Commission", Centre for European Reform, 20 February.
Nahm, J (2021), Collaborative Advantage: Forging Green Industries in the New Global Economy, Oxford University Press.
Piechucka, J, L Saurí-Romero and B Smulders (2023), "Industrial Policies, Competition and Efficiency: The Need for State Aid Control", Journal of Competition Law & Economics 19(4): 503-526.
Pinkus, D, J Pisani-Ferry, S Tagliapietra, R Veugelers, G Zachmann and J Zettelmeyer (2024), "Coordination for EU competitiveness", Bruegel EGOV Report, March.
Rodrik, D (2014), "Green Industrial Policy", Oxford Review of Economic Policy 30(3): 469-491.

Rodrik, D, R Juhasz and N Lane (2023), "Economists Reconsider Industrial Policy", Project Syndicate, 4 August.
Schnabel, I (2024), "From laggard to leader? Closing the euro area's technology gap", Inaugural Speech: European and Monetary Union Lab, 16 February.

This article was originally published on VoxEU.org.

How to de-risk

How should the EU ‘de-risk’ its external economic relationships without foregoing the benefits of trade? Jean Pisani-Ferry, Beatrice Weder di Mauro and Jeromin Zettelmeyer discuss

Executive summary

Pandemic-related supply disruptions, the energy crisis provoked by Russia's invasion of Ukraine and economic coercion by China have put economic security high on the European Union policy agenda. The question is how exactly the EU should 'de-risk' its external economic relationships without foregoing the benefits of trade. The standard answer is that it should identify product-level trade dependencies, mainly on the import side, and reduce them, mainly through diversification of suppliers, while otherwise maintaining maximum trade integration.

This Policy Brief argues that this answer falls short. First, product-level dependencies cannot be identified reliably even with sophisticated analysis and data. As a result, both 'missed dependencies' and 'false positives' are inevitable. Second, external shocks and coercion could be propagated through exports, productive assets held abroad and financial channels as much as through imports.

The analysis has five main implications:

1. Import de-risking should focus on a few product categories for which the costs of supply interruptions would be unquestionably large. This reduces false positives.

2. De-risking and/or buffers to deal with exports and financial coercion require more attention.

3. De-risking must be complemented by raising resilience against all shocks, whatever their origin. This requires a deeper and broader European single market.

4. De-risking and resilience must be complemented by deterrence.

5. A sufficiently high probability of chronic trade conflict – or one very large conflict – may justify reducing overall integration with a large trading partner, on both the export and import sides.

EU economic security policies have been right to emphasise the reduction of import dependence on chips and critical raw materials, and the creation of a powerful legal instrument to deter coercion (the Anti-Coercion Instrument). In most other respects, there is room for improvement.

Economic risks relate increasingly not just to crises or shocks, but to deliberate economic coercion by foreign governments

1 Introduction

Over a period of just fifteen years, Europe has been confronted with a financial shock that originated in the United States, a pandemic shock that originated in China but could have come from anywhere, and an energy shock provoked by Russia's invasion of Ukraine. These events have prompted a re-examination of efficiency/security trade-offs that arise as a result of international integration, and particularly as a result of specialisation in international trade and the vulnerabilities of global supply chains.

Economists and policymakers have long worried about similar trade-offs. At the most fundamental level, such trade-offs arise from the standard tension between growth and economic crises: higher growth is often accompanied by greater instability. For example, regulation of financial and product markets may prevent or mitigate financial or environmental hazards at the cost of dampening entry and growth of firms. Similarly, in open economies, trade and financial integration may be good for growth, but can expose economies to imported shocks.

The most recent set of concerns – as exemplified, for example, by a series of European Commission (2021, 2022) papers and an associated legislative agenda (see section 4, and McCaffrey and Poitiers, 2024) – differs from these standard preoccupations in two respects.

First, economic risks relate increasingly not just to crises or shocks, but to deliberate economic coercion by foreign governments or even non-governmental entities (such as criminal groups). This is probably the reason why the term 'security' – as opposed to 'stability' or 'resilience' – has become popular to describe the mitigation of economic, rather than just national security threats (we discuss the difference in section 2).

One reason to be concerned with economic coercion is that China, an increasingly powerful and authoritarian country, has been applying coercion regularly in response to political actions by trade partners (for example,

Australia's call for investigations into the origin of the COVID-19 pandemic and Lithuania's decision to let Taiwan open a representative office in Vilnius[1]).

But the concern is not just about China: the policies of President Trump between 2017 and 2020 showed that even one's closest ally can be tempted to leverage its market power and its control of the technical and financial infrastructures of globalisation. The possibility of a second Trump term is now prompting a reflection on the need for Europe to prepare for such a risk (Gonzales Laya *et al* 2024).

Second, recent concerns have focused mostly on trade-related rather than financial vulnerabilities. This reflects the fact that trade-related vulnerabilities have become more prominent as a result of specialisation and the vulnerability of global supply chains that maximise efficiency, but at the cost of creating hidden fragilities.

But the prominence of trade concerns may also reflect a rather myopic reasoning, as the last two or three external shocks that Europe (and, to a lesser extent, the US) has suffered have been trade-related: supply chain disruptions related to COVID-19 and energy price shocks following the Russian invasion of Ukraine.

In line with this concern, we focus mostly on trade-related external economic security. This should not be taken to imply that Europe does not need to worry about financial security. But unlike trade-related security, financial risks continue to be mostly of the financial-stability variety, linked to shocks and financial vulnerabilities rather than coercion. To the extent that financial coercion is a serious concern, it is linked to one main potential source: the United States if President Trump returns (see section 2). In contrast, trade-related external security risks are ubiquitous.

In this Policy Brief we seek to answer two critical questions. First, how should trade-related vulnerabilities be identified, and what trade relationships make Europe particularly vulnerable to shocks and coercion? Second,

how can these vulnerabilities be reduced while minimising the costs of 'de-risking' and reducing the chances of unintended consequences? Four such potential costs come to mind:

- Foregoing some of the gains from trade specialisation and trade openness. This could weigh on European growth and competitiveness, which depend on export specialisation and on importing raw materials and intermediate inputs more cheaply than they could be produced at home (if at all). It could also make it harder to attain emission reduction objectives, by raising the cost of the transition to renewable energy sources. In turn, this could exacerbate social and political divisions related to climate action.

- Becoming more vulnerable to domestic shocks including natural disasters, epidemics and home-grown financial crises – and more generally, to any shock whose consequences would be mitigated by international trade and/or capital flows.

- Damaging international cooperation. This could include European Union cooperation with China on vital matters of common interest, such as climate-change mitigation, as well as respect for the rules of the multilateral trading system. Notwithstanding the damage that the World Trade Organization has suffered over the last decade, these rules continue to be largely respected (Mavroidis and Sapir, 2024).

  An aggressive 'de-risking' of European trade relationships through trade policy tools and subsidies could trigger protectionist reactions from trading partners, particularly if measures violate WTO rules. It could also become an excuse for protectionists in the EU, who might use economic-security arguments to further special interests.

- Damaging cohesion within the EU. EU countries differ in their trade structures and their dependence on specific export and import markets. As a result, attempts to de-risk trade may have net benefits for some

and net costs for others. If de-risking becomes a source of division, it may be counterproductive, as internal divisions in the EU are partly what an adversary – whether China, Russia or President Trump – might try to exploit (and indeed, divisions are what these three powers have tried to exploit in the past).

The remainder of this paper summarises as best we can the answers to these questions, drawing on a set of papers collected in Pisani-Ferry *et al* (2024). Section 2 defines what we mean by economic security, and what risks we should be worrying about. Section 3 discusses how these risks should be addressed in principle. What trade relationships require de-risking? Section 4 discusses the instruments. How can protection be built that preserves the benefits of trade? A concluding section summarises the main lessons.

2 Defining risks to economic security

As noted by Bown (2024), economic security remains an emerging concept. At its most abstract level, it can be defined as both preventing bad economic outcomes and making sure that should risks materialise, the damage they cause is minimised. Societies care both about raising expected welfare and about reducing its volatility. Economic security is concerned with the latter.

Defined in this broad way, economic security has been a standard concern of policy- makers for centuries – and not just of economic policymakers, since economic harm can be inflicted by 'non-economic' shocks, including political disruption and wars. The use of state intervention to address these concerns, including industrial policy and trade policy, is similarly nothing new (Kelly and O'Rourke, 2024).

The question, then, is how the concept of 'economic security' differs from those of 'economic crisis prevention' or 'national security'. To the extent that the perceived nature of the risk and risk propagation has changed, it is important to understand how it has changed, to avoid duplication, and to prevent overreaction to perceived new risks when the old risks and risk propagation channels might still be there.

Economists concerned with crisis prevention and mitigation typically focus on risks and vulnerabilities related to the financial system or the structure of production. For example, credit cycles can expose countries to financial crises, which are propagated internationally. Dependence on commodity exports or imports exposes economies to swings in international prices and to disruption to domestic production that relies on commodity imports.

Military and security experts worry about a different type of threat: harm that is inflicted purposely by outside actors, normally nation states, but also terrorist or criminal organisations. Murphy and Topel (2013) widened the definition of national security to include all 'substantial threats' to the safety and welfare of a nation's citizens, eg. including national catastrophes and public health threats.

Defined this broadly, national security would include preparedness and mitigation against any harmful acts conducted by foreign governments or non-governmental organisations with military or non-military means, including economic sanctions, and threats related to physical and information infrastructure.

The recent usage of the term 'economic security' is at the intersection of non-financial economic crises and national security in the broad sense defined by Murphy and Topel[2]. Specifically, it focuses on harm inflicted through international economic relationships – and particularly trade relationships – whether these reflect exogenous shocks (such as COVID-19-related trade disruption) or deliberate actions by foreign governments or non-governmental organisations (Bown, 2024; McCaffrey and Poitiers, 2024; European Commission, 2021, 2022).

These risks are particularly relevant today because of the combination of economic integration through trade and FDI, specialisation, long supply chains and actors willing to engage in coercion through these channels.

It is in this sense that the term 'economic security' will be used in the remainder of this paper. In this definition, achieving economic security involves the prevention and mitigation of:

- Disruption to critical imports, whether accidental or deliberate;

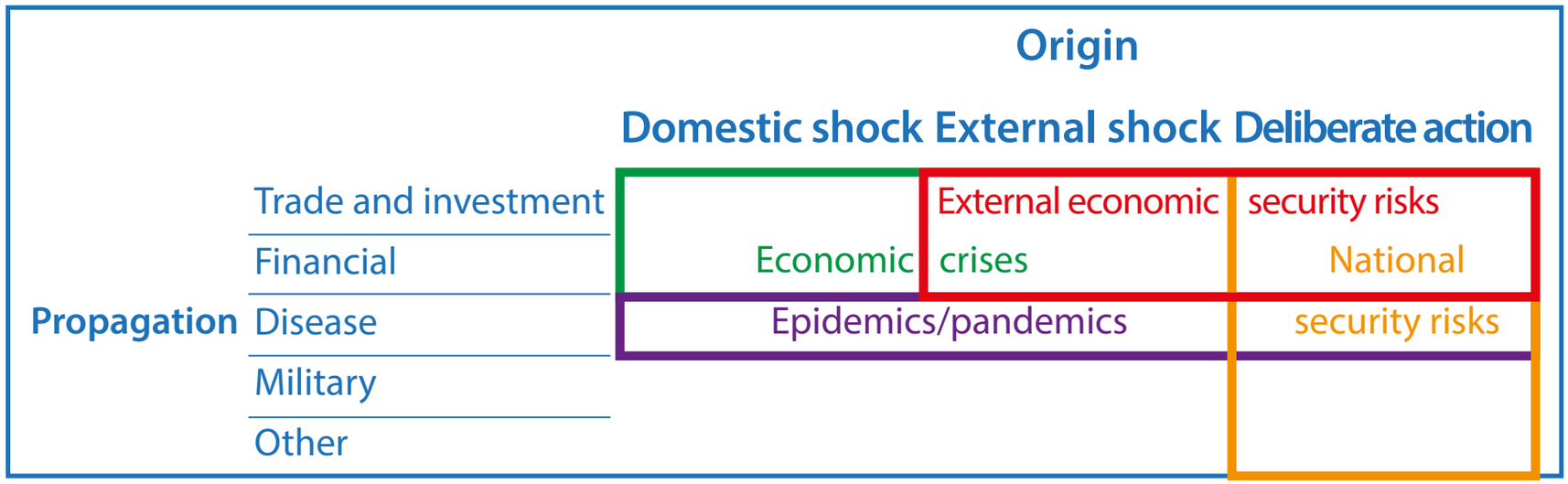
- Economic coercion through restrictions or boycotts on specific exports, along the lines of actions taken by China against Australia; or through pressures on foreign companies even when they produce locally (for example, threats of depriving them from access to the domestic market, restrictions on profit repatriation, or expropriation);

- A broad disruption of global trade at a scale with macroeconomic impact, for example, as a result of geopolitical conflict leading to economic sanctions or a protracted tariff war with a major trading partner. Events that could trigger such scenarios include a Chinese attack on Taiwan, or the re-election of President Trump triggering a sharp deterioration of the political relationship between the US and the EU.

It is important to emphasise that this a narrow – perhaps inappropriately narrow – definition of economic security, for two reasons. First, it disregards the possibility of economic disruptions as a result of domestic shocks, which historically have been a major source of economic crises (Table 1).

Hence, a better term for the type of economic-security risks we discuss would be 'external economic security'. This terminology reminds us that there could be trade-offs not just between economic security and economic growth, but also between external economic security and security from domestic shocks. International integration may increase exposure to the former but offers protection against the latter.

Table 1. Varieties of welfare threats and propagation mechanisms



Note: The columns in Table 1 define the origin of a bad event – an exogenous shock originating at home or abroad (pro- duction disruption, natural catastrophe, transportation or infrastructure disruption, confidence shock) or a deliberate action by a foreign government or a non-governmental entity. The rows define the propagation channel: economic activity related to trade or finance, disease, military action or other (for example, through IT infrastructure).
Source: Bruegel.

Second, the narrow definition largely ignores external economic security risks through financial channels. However, international finance – including the international payments system and the confiscation of financial assets located in foreign jurisdictions – is an obvious instrument of economic coercion and economic sanctions, as shown by G7 sanctions against Russia since its full-scale invasion of Ukraine.

The main reason why financial risks do not feature prominently in the recent literature on European economic security is that Europe is much less likely to be on the receiving end of such sanctions, given the control exerted by the US and its allies over international finance.

But this could rapidly change if President Trump is re-elected in the United States and decides to use financial coercion against Europe for whatever reason (for example, to force Europe to align its foreign or commercial policies with those of the United States, as was the case when the US threatened EU firms with 'secondary sanctions' for violating US-imposed sanctions on Iran).

A broader analysis of European economic security should take into account such financial economic risks and how to mitigate them. For now, the remainder of this paper focuses on trade and investment-related risks.

These are particularly relevant for the relationship with China, but could also become relevant in the event of a return of President Trump and a revival of US tariffs against Europe, whether imposed for mercantilist or political reasons.

3 What to de-risk

Firms have incentives to avoid becoming dependent on one or a small number of suppliers or customers, particularly when those suppliers or customers are vulnerable to high risks out- side their control, including politically motivated interference.

Yet, as Mejean and Rousseaux (2024) have pointed out, the firms' private interest in security may not be enough to take care of the collective EU security interest. Firms often fail to realise the extent to which suppliers or customers are themselves subject to risks, simply because they do not know the entire value chain.

Firms also do not internalise the potential costs of supplier or customer dependency on the entire value chain, and ultimately the welfare of citizens. If a supplier relationship rep- resents a critical link in that chain, the social costs of that link failing may far exceed the private costs to the firm. This argument, which is broadly consistent with the evidence presented by Bown (2024), can justify policy-led de-risking.

But what areas of trade require de-risking? How can policymakers tell when trade dependencies are excessive, in the sense that the economic security risks of trade outweigh its benefits, both for efficiency and growth and as protection against domestic disruption? The ideal way to answer this question would be through a firm-level model of trade and supply relationships, both across borders and within the EU.

The model would 'know' who trades with whom, how specific inputs enter each stage of production, and to whom firms sell. It would also have information about the ease of switching suppliers if a supplier fails or sharply raises its prices. Such a model could be used to stress test European economies in relation to specific supply chain or customer risks.

Where large effects are found, it would be used to identify trading relationships worth de-risking. Unfortunately, such a model does not exist and may never exist because of data limitations. We are therefore constrained by the available information and should make the best of it.

3.1 Critical goods and the risk of import disruption

Suppose we were mainly interested in risks related to import disruption. This would be the case if exports are either well diversified or go mainly to countries that one would not consider to be major sources of shocks.

In that case, the following approach might be a close substitute for the perfect model. Using the most disaggregated data possible, one should identify products for which:

1. A large share of EU consumption relies on imported inputs;

2. Foreign supply of these goods is highly concentrated;

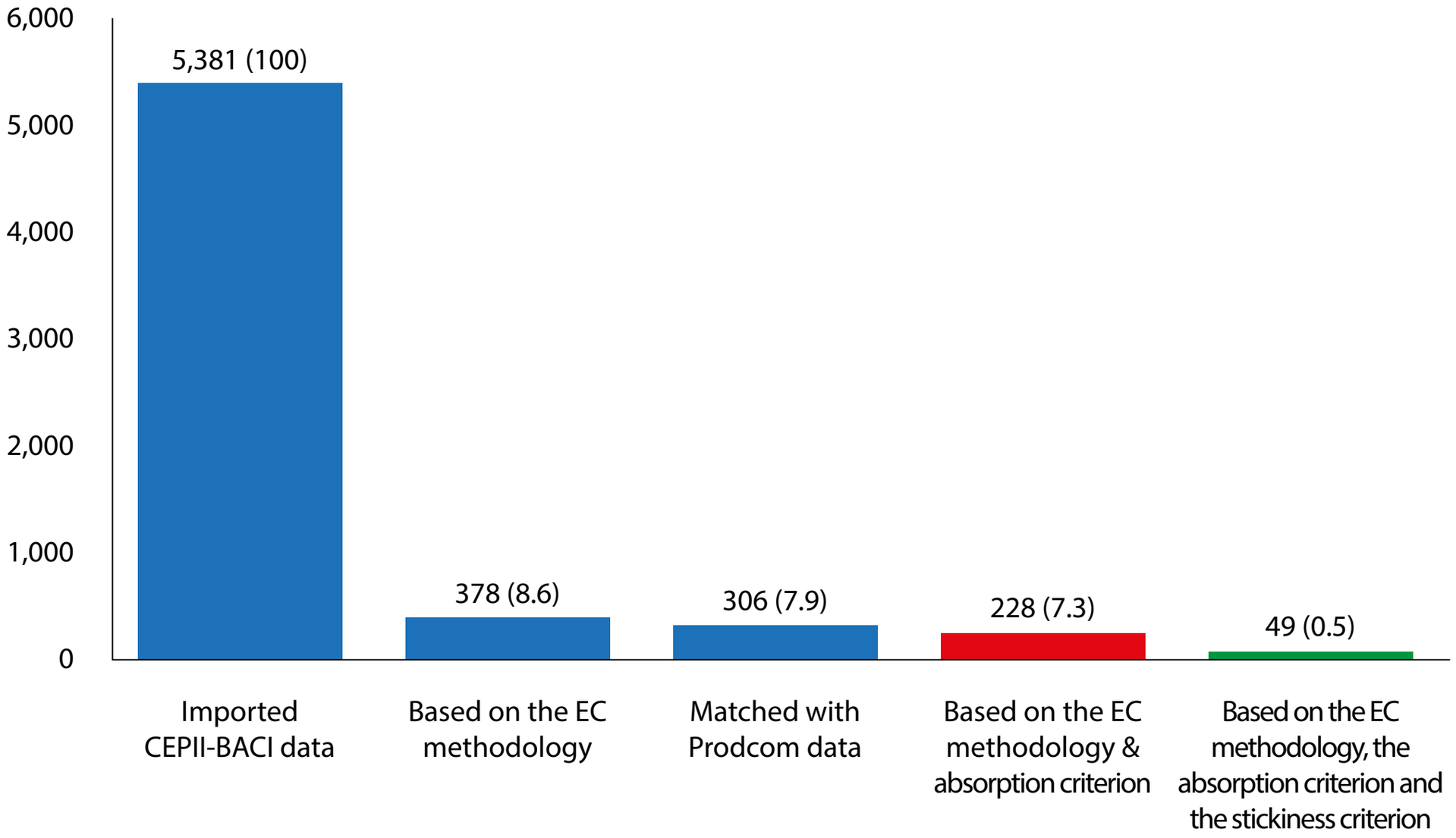
3. Finding alternative suppliers in the event of a disruption is difficult, and

4. Disruption to supply would have high economic costs. Unlike criterion 3, this criterion reflects the substitutability of products in either consumption or production, as opposed to the substitutability of supplier relationships.

Products that meet all four criteria would be prime candidates for de-risking. This approach, which builds on work undertaken by the European Commission (2021), approximately describes the approach taken in Mejean and Rousseaux (2024). Their main innovation relative to the work of the Commission and other authors is step 3, which they implement by eliminating products for which 'relationship stickiness' – the typical duration of firm-supplier relationships – drops below a specific threshold.

For example, if the stickiness threshold is set at the sample median, the number of products for which the EU should consider itself import-dependent drops from 378 to just 105, and to just 49 if the 75 percent least relationship-sticky products are eliminated (Figure 1). Focusing only on upstream intermediate products – for which an export ban would affect many supply chains and hence have high economic costs – would reduce the list further, to just 21 products. For 12 of these, the main supplier is China.

Figure 1. Number of products for which the EU is import dependent



Note: The figure shows the numbers of products for which the EU is import-dependent according to various methodologies, starting with that of the European Commission (2021) (second blue bar) and adding the criteria proposed by Mejean and Rousseaux, based on the ratio of imports over domestic absorption (red bar) and the degree of product stickiness (green bar). Numbers in brackets refer to percentage of value of EU imports.
Source: Mejean and Rousseaux (2024).

To these, Mejean and Rousseaux (2024) suggested adding a small number of 'critical goods' that, if insufficiently supplied, 'can result in human losses and other severe non-economic consequences'. These would include between two and 19 pharmaceutical products, depending on where the substitutability cut-off is set, as well as inputs to the green transition.

Interestingly, most of these inputs – including most critical raw materials, which have been among the main justifications for the drive to de-risk imports, particularly from China – currently fail one or several of Mejean's and Rousseaux's dependency tests.

While highly relationship-sticky, batteries and their components, hydrogen technologies, rare earth metals and solar panels fail the concentration test, and most components of solar panels fail both the concentration test and the relationship-stickiness test.

Yet, Mejean and Rousseaux urged caution with respect to these products, on the grounds that demand for them is developing so fast that the structure of EU imports during 2015-2019, on which concentration indices and import needs are based, may be a poor proxy for trade dependencies in the future.

Mejean and Rousseaux's work represents the most exhaustive analysis so far to identify dependencies on the basis of ranking critical imports with respect to concentration and relationship substitutability, and deciding on thresholds above or below which concentration is deemed too high or substitutability too low. Precisely because it is more thorough and comprehensive than previous attempts in this literature, Mejean and Rousseaux (2024) illustrates the intrinsic limitations of this approach.

- We have so far no systematic way of telling which imports are genuinely critical. Focusing on upstream products and pharmaceuticals may miss other products (such as computer chips), the accidental scarcity

of which would cause large economic or non-economic losses. Meanwhile, some upstream products and pharmaceuticals might not be critical if they can be substituted by other products.

The European Commission's (2021) approach of designating whole 'ecosystems' (sectors, such as health, energy, digital, electronics and aerospace) as critical, seems even more problematic, both because many products in these sectors are not in fact critical and because products outside these sectors that may well be critical could be missed (for example, most of Mejean and Rousseaux's upstream products).

- As both Mejean and Rousseaux (2024) and Bown (2024) emphasised, data limitations imply that import dependence measures do not reflect indirect exposure. If the EU has an import exposure to a country that is itself import dependent on China for this product (or an important intermediate input), then direct import dependence on China might significantly understate total import dependence.

- The final lists can be very sensitive to how the cut-offs are set, which is somewhat arbitrary. For example, whether relationship substitutability thresholds are set at the twenty-fifth, fiftieth or seventy-fifth percentile adds or subtracts large shares of products from the sample.

- Supplier relationships in normal times tend to be relatively long (25 and 19 months, respectively, for the seventy-fifth and fiftieth percentiles in Mejean and Rousseaux's sample). This implies that unless replacement duration is significantly shorter in a crisis, an import interruption could be very damaging even for products that are relatively non-relationship-sticky in normal times.

  But the impact of a forced interruption on the replacement period could go both ways. Firms seeking to replace suppliers under duress would have incentives to do so much faster than in normal times. However,

finding new suppliers when many other firms are trying to do so could take longer and/or result in price jumps for scarce supplies, which could be very damaging.

3.2 Risk from export disruptions and from decoupling

Another problem is that an approach focused on reducing dependence on critical imports does not consider disruptions to exports, which could equally have a macroeconomic impact if they were highly concentrated in any one destination country.

For example, 20 percent of EU exports got to the US, 13 percent to the United Kingdom and 9 percent to China; while 41 percent of UK exports go to the EU, 21 percent to the US and 5 percent to China. Furthermore, just as import dependency numbers ignore indirect exposures, so do export shares. For example, direct UK export dependency to China is only 5 percent, but the UK's indirect exposure via the EU alone could be larger if UK products are part of the value chains of goods ultimately destined for the Chinese market.

While demand shocks via exports are a standard risk of trade integration, geopolitical conflict can take this risk to an entirely new level. First, hitting the exports of specific industries through import bans, high tariffs or social-media campaigns can be a form of geopolitical coercion. As reported by Bown (2024) and McCaffrey and Poitiers (2024), there are numerous examples of Chinese coercion of this type.

This type of coercion is typically not macroeconomically critical, but may seek to exploit the lobbying power of groups that are hurt, as well as internal divisions (in the case of the EU, this may include divisions across member states). Second, deliberate economic sanctions can of course have a much greater impact than swings in export demand triggered by normal economic fluctuations, or even than an economic crisis in a trading partner.

Baqaee *et al* (2024) simulated the impact of a decoupling from China in a trade model with 43 countries and 56 sectors, in the form of a complete stop in trade between a 'Friends' bloc comprising the G7 countries, Spain, the Netherlands and an artificial country comprising the rest of the EU, and a 'Rivals' bloc including China and Russia, on the assumption that trade continues both within these blocs and with the rest of the world.

As might be expected, the short-term effects are substantial, with German output calculated to decline by 3-5 percent of GDP. At the same time, the simulations suggest that the cost of a complete decoupling from China would be relatively low if done slowly over time: around 1.25 percent of GDP for Germany and Japan, while the US and the remaining European countries would suffer in the range of 0.47 percent to 0.69 percent of GDP.

The intuition behind this result is that the welfare costs of an end to trade integration between China and the 'Friends' group are mitigated by the fact that the Friends continue to trade with each other and with the 'Neutrals', and that these groups are sufficiently large and diverse to preserve most of the gains from trade.

3.3 Putting it all together

Combining the insights of Baqaee *et al* (2024) and Mejean and Rousseaux (2024) with the assumption that external economic risks include not only exogenous shocks to trade but also coercion, and possibly a wider trade disruption involving China, leads to the following conclusions.

First, there is a strong case for de-risking concentrated exposures to critical imports, by either diversifying supply or making preparations to mitigate disruption. However, identifying such products turns out to be very difficult, mainly because it is hard to assess the criticality of products, ie. the welfare losses inflicted by a shortage or price spike. While we know that some products are critical – chips, energy, some pharmaceuticals, some minerals and some upstream inputs – we do not know what other products are critical.

A good way to start is by de-risking the products known to be critical. Because we don't know how long it would take to find new suppliers in a crisis, or how price sensitive these imports might be to a loss of the main supply source, products known to be critical should be de-risked even if their relation- ship stickiness in normal times is fairly low.

The identification of such products obviously needs to take into account the costs as well as the benefits of de-risking. Take the example of solar panels and their components, often cited as a prime de-risking candidate because of their importance in the green transition and China's overwhelming global market share (63 percent, according to Mejean and Rousseaux, 2024).

However, the short-term economic costs to the EU of a complete stop in solar panel imports from China would be tiny (hitting mostly installation services, while leaving the solar capacity unchanged). Unlike imported gas from Russia, disruption to solar panel imports from China would have no direct impact on the energy supply, although it would affect the increase in installed energy capacity and would raise the cost of replacing panels that become obsolete.

Hence, the main benefit of de-risking Chinese solar panel imports would be insuring against a (possible) disruption to the energy transition to renewables, which could sharply raise solar-panel prices. This needs to be weighed against the certain price impact of a decision to diversify away from Chinese solar imports and purchase panels from more expensive sources, which will slow the green transition.

Second, the de-risking of trade dependencies cannot be the only layer of protection against import disruption, because it will never be possible to identify and de-risk all critical products. Beyond trade de-risking, it is hence essential to strengthen the resilience of European economies against import shocks, whatever their source. This is

an argument for a better-functioning and more flexible single market, and for the broadening of international trade relationships through free-trade agreements with friendly countries.

Third, it is important to de-risk export dependencies as well as import dependencies. For specific products, this could be done in three ways: by deterring coercion (as the EU's new anti-coercion instrument, discussed in the next section, attempts to do); by offering EU producers incentives to diversify export destinations, particularly to reduce concentrated exposures to China; and through insurance mechanisms that reduce ex post the impact of export disruptions to specific products.

The latter must of be designed in a way that avoids moral hazard, ie. does not encourage concentrated exposures ex ante. We return to possible instruments for export diversification and ex-post protection in the next section.

Fourth, there is a role for deterring coercion, rather than just reducing vulnerability to it. This is because de-risking of export and import dependencies will never be complete – and should not be complete, given that de-risking needs to be weighed against the benefits of trade specialisation and continuation of trade with China and other countries that may use coercion.

Fifth, there is the question of whether the EU should reduce its overall trade integration with China to soften the blow of sudden trade disruption triggered by a geopolitical confrontation. According to Baqaee *et al* (2024), the cost of a gradual reduction in trade integration with China would be small for most EU countries, even if trade integration is reduced all the way to zero.

Even for Germany, where the cost of complete decoupling from China would not be small, the cost of a partial reduction of trade integration – for example, reducing export and import shares by one third – would be small if

pursued gradually. On this basis, policy measures to encourage a pre-emptive reduction in trade integration would be justified if all three of the following conditions are met:

1. The probability of a very costly sudden trade disruption is considered to be sufficiently high, and

2. Firm-level diversification of trade is not, by itself, sufficient to engineer this pre-emptive de-risking;

3. Targeted (ie. firm- or sector-level) export diversification efforts do not have a substantial impact in terms of reducing aggregate import dependency.

There is significant uncertainty around each of these points. With regard to points two and three, Bown (2024) found that trade diversion triggered by US tariffs on China and Chinese retaliation has further increased EU trade integration with China. With fresh US legislation directed against Chinese imports, such as the Inflation Reduction Act, this effect might continue.

At the same time, the combination of a heightened sense of the risks created by concentrated exposure to China and the structural slowing of the Chinese economy might push in the other direction. Furthermore, targeted de-risking efforts may have an aggregate impact, particularly if they reduce concentrated exposures to China in major sectors for the EU economy, such as the car industry.

Finally, it is important to highlight two trade-related economic-security concerns that are the intellectual cousins of the risks identified and quantified by Baqaee *et al* (2024) and Mejean and Rousseaux (2024), but are not directly discussed in those papers.

The first is the obvious risk, already mentioned in section 2, of a broad disruption to European trade with the United States in the event of a return of Donald Trump to the US presidency[3]. Given the much larger share of US imports and exports in European trade, this could hit Europe even harder than a disruption to trade with China.

While Baqaee *et al* (2024) did not directly simulate such a shock, this is suggested by their 'EU autarky' scenario, which has substantial costs even in the long run, ie. even when phased-in slowly (a permanent consumption loss of 9 percent of GDP). It follows that de-risking the trade relationship with the US by reducing trade integration might makes sense only if an even more catastrophic sudden decoupling from the US is viewed as likely.

However, a disruption to trade with the US would likely take the form of a (limited) tariff war rather than a trade embargo. This argues against a pre-emptive reduction in trade with the US. Instead, the EU must be politically prepared to fight a trade war with the US, if and when a returning President Trump decides to start such a war.

A second related concern is that exposures to China and other countries that might engage in coercion against EU firms could take the form of asset expropriation – in particular, expropriation of production sites. By removing an important source of foreign revenue and profits, this could impact EU firms in much the same way as an import prohibition.

However, the risk would show up ex ante in the form of a concentration of profit sources, rather than concentrated exports, and the remedy could involve diversification of production sites and profit centres, rather than diversification of exports, as along with increases in capital buffers.

Summing up, our analysis results in five main calls for European policy action:

1. Reduce import dependency for critical products;

2. Diversify foreign revenue sources and/or strengthen firm resilience against potential disruption to foreign demand, asset expropriations or payment controls impeding profit repatriation;

3. Deepen the EU single market and make it more flexible;

4. Deter economic coercion of any kind, whether through imports or exports, or through other means such as expropriation;

5. Possibly, limit overall trade dependency (and particularly export dependency) on China, at the aggregate level.

Achieving these objectives requires policies that are effective, that balance costs and benefits, and that minimise risks of unintended consequences. We next examine what such policy might look like concretely, starting with those the European Commission has already started implementing.

4 How to de-risk

As the outbreak of COVID-19 revealed dangerous vulnerabilities and called for a reassessment of the EU's international economic relations, rising pressure from the US under the Trump presidency and the increasingly aggressive behaviour of the Chinese government focused the attention of European policymakers on the threat of economic coercion and prompted a redefinition of the toolkit with which they could respond.

The EU took a series of major initiatives to strengthen its economic resilience and to equip itself to better counter malicious behaviour by economic partners (Box 1).

Box 1. Additions to the European external economic security policy toolkit

The EU has adopted or is discussing a series of new initiatives, which complement standard trade defence instruments[4] (anti-dumping or anti-subsidy duties consistent with the World Trade Organisation Agreement on Subsidies and Countervailing Measures, for which the EU has developed procedures that are in the process of being strengthened):

The Foreign Subsidies Regulation[5] (FSR, in force since July 2023) introduced new tools to tackle foreign subsidies that cause distortions and undermine the level playing field in the areas of mergers and acquisitions and procurement (see Anderson, 2020).

The European Chips Act[6] (in force since September 2023) is intended to bolster Europe's competitiveness and resilience in the semiconductor sector by supporting large-scale manufacturing projects via somewhat more permissible subsidy rules compared to a conventional Important Projects of Common European Interest (IPCEIs, investment projects involving crossborder collaboration and state aid from several EU countries). It also entails measures aimed at mapping and monitoring the semiconductor supply chain to assess ex-ante risks of potential import disruption but also and envisions broader powers for the Commission to act in a crisis, including as common purchasing body (see Poitiers and Weil, 2022).

The Net Zero Industry Act (NZIA)[7] and related parts of the Temporary Crisis and Transition Framework[8] (TCTF) are intended to strengthen the European ecosystem of clean-tech manufacturing. The NZIA includes measures intended to accelerate permitting, while the TCTF allows member states to provide subsidies to clean tech manufacturing projects which can match subsidies of third countries under certain conditions (see Tagliapietra *et al* 2023).

The Critical Raw Materials Act[9] (CRMA) aims to tackle the issue of highly concentrated imports of certain raw materials that are of strategic importance. It seeks to boost domestic mining, refining and recycling of such raw materials through accelerated permitting procedures as well as measures related supply chain monitoring, stockpiling and improving the recyclability of CRMs (see Le Mouel and Poitiers, 2023).

The Health Emergency Preparedness and Response Authority (HERA)[10] that was launched in September 2021 has as part of its mission to improve the resilience and availability of medical supplies. It aims to achieve this mission by identifying key supply chain bottlenecks and addressing them through measures such as coordinated stockpiling and joint procurement.

The Anti-Coercion Instrument (ACI, in force since December 2023) is intended to provide to the EU a wide range of possible countermeasures when a third country exercises coercion. It gives the EU extensive powers to deploy countermeasures in response to an act of foreign coercion, including the imposition of tariffs, restrictions on trade, services and intellectual property rights, and restrictions on access to foreign direct investment and public procurement.

The Internal Market Emergency and Resilience Act[11] (IMERA, formerly Single Market Emergency Instrument, on which agreement was reached between the Parliament and the Council in February 2024) aims at ensuring continued access to critical goods and services. Although primarily intended to respond to COVID-type emergencies, it also covers disruptions to the single market triggered by conflicts, such as the war in Ukraine.

Commission initiatives on inward and outward investment screening and the coordination of export controls were proposed in January 2024. The coordination mechanism for inbound investment screening is in place since 2020, but it mainly commits member states to put an investment screening into place. The 2024 economic security package includes an update of this scheme, but remains vague on the prospect of outbound investment screening.

Limitations notwithstanding, the EU has assembled an impressive package that expresses a change of attitude. Considerable effort has gone into addressing critical import dependencies, giving the European Commission powers to deter coercion (the Anti-Coercion Instrument, application of which must be triggered by a majority in the Council), and preventing a breakdown of the single market in an emergency (Internal Market Emergency and Resilience Act, IMERA). However, these efforts fall well short of meeting the policy objectives listed at the end of section 3.

First, and most obviously, export dependencies have been largely neglected. Aside from the intention to negotiate additional trade agreements with friendly nations, there is no instrument to encourage export diversification and/or reduce concentrated export dependence on China.

Second, instruments to address import dependencies remain imperfect and incomplete:

- While the European Chips Act, Critical Raw Materials Act (CRMA) and Health Emergency Preparedness and Response Authority have plausible economic-security justifications, the Net Zero Industry Act covers a broad swathe of goods that mostly fail to meet the definition of critical good proposed in section 3[12].

  Many other goods that might be critical, such as the upstream products with high import concentration identified by Mejean and Rousseaux (2024), remain outside the scope of any of these acts. There is no framework for identifying goods that may be genuinely critical, but are not part of any of the four identified product categories.

- EU-level instruments to reduce dependency on these goods are for the most part weak. EU-level funding for industrial policy directed at expanding EU capacity is small (Chips Act) or non-existent (CRMA). Trade policy

instruments rely mainly on increasing market or investment access for EU companies via new or expanded trade agreements.

The main channel through which these acts operate is by giving EU countries greater leeway to subsidise investment in the areas covered by these acts. While this may lead to occasional successes (investment in a critical area that would otherwise not have happened), there is no governance structure to ensure that critical dependencies are reduced in a timely way.

Furthermore, the approach mostly benefits EU countries that have the fiscal resources to provide large subsidies, and large incumbents, which have the clout and scale to lobby for subsidies and participate in IPCEI consortia.

Third, the Commission has so far missed the opportunity to rally members states behind the push to increase resilience by deepening the single market. This would help the EU resist external shocks and coercion – whatever the source and the channel – by allowing faster re-direction of trade and supply.

Banking and capital markets union would raise economic security both by funding new productive capacity and by improving automatic risk-sharing, better risk sharing across intra-EU borders would in turn make the EU more cohesive, and would make it harder to exploit internal divisions.

A more systematic attempt to strengthen economic security could involve the following elements.

1. A process for identifying and regularly reviewing critical import dependencies, based on the criteria developed in section 2, and better data (Mejean and Rousseaux, 2024; Bown, 2024). Better data may require

Table 2. Economic security objectives and available instruments

Objective	Available Instruments	Problems
Reduce import dependency for critical products	Important Projects of European Interest (IPCEIs) European Chips Act Critical Raw Materials Act Net Zero Industry Act and related sections of the Temorary Crisis and Transition Framework for State Aid Health Emergency Preparedness and Response Authority (HERA)	Imperfect match between critical products and targetted products. Lack of cost-benefit analysis Weak EU-level instruments Weak governance - actions and funding rely mostly on member states and lobbying by large firms
Diversify concentrated export exposures at the firm level	None, except for intention to negotiate additional free trade agreements with 'friends'	Lack of instruments leaves EU vulnerable to coercion
Deepen the single market and make it more flexible	Internal Market Emergency and Resilience ACT (IMERA)	No economic security-motivated deepening agenda
Deter economic coercion	Anti-Coercion Instrument	Council majority required to allow the Commission to deploy ACI powers
Limit overall trade dependency on China's market	None, except for intention to negotiate additional free trade agreements with 'friends'	Economic cost of sudden decoupling may deter appropriate action by the EU

Source: Bruegel.

more systematic due diligence on the part of European firms in relation to their supply chains, from an economic-security perspective.

2. Stronger governance and better funding for a competition-friendly EU-level industrial policy. This could involve:

> i. An institution similar to the US Advanced Research Projects Agencies (ARPA) to develop technology in areas that are identified as critical (Tagliapietra et al, 2023; Pinkus *et al* 2024).
>
> ii. Where the technology exists already, allocation of production or investment subsidies through auctions (along the lines of auction mechanisms that are currently used to tender renewable energy capacity).

These mechanisms would not necessarily require large funding. US ARPA budgets are relatively modest (in the single digit billon range), while the auction process could be co-funded by EU countries, along the lines of the 'Auctions as a Service' concept proposed by the Euro- pean Commission in relation to climate goals (European Commission, 2023).

3. The use of WTO-consistent trade instruments to incentivise import and export diversification. These could include:

> i. On the import side: countervailing duties, justified by the presence of a foreign subsidy, that are focused on an area in which there is a critical import dependency on the country that is responsible for the subsidy;

ii. On the export side, a duty levied on EU exports to countries for which export exposure is considered excessive. The latter could be politically difficult, but would be fully consistent with WTO rules[13].

4. As an alternative to export taxes, requiring exporters that are highly dependent on a specific export destination to buy publicly provided political risk insurance that would defray the costs of ex-post public support in the event of coercion (and would discourage exports to the destination in question).

5. Incentivising European firms that are highly dependent on production and profits in foreign jurisdictions to diversify production, structure their operations or hold capital to enable them to survive an expropriation (or controls that impede profit repatriation).

6. To further increase the deterrence value of the ACI, allowing the Commission to trigger retaliation under the ACI without requiring confirmation by a majority of member states.

7. Preparing for economic coercion through financial channels rather than just trade channels. While European firms have not recently been at the receiving end of such coercion, this may change if Donald Trump returns to the White House.

8. Invigorating the single market for economic security rather than just for efficiency reasons. ■

Jean Pisani-Ferry is a Senior Fellow at Bruegel, Beatrice Weder di Mauro is President of the Centre for Economic Policy Research, and Jeromin Zettelmeyer is Director of Bruegel

Endnotes

1. See, for example, The Economist, 'China punishes Australia for promoting an inquiry into covid-19', 21 May 2020; and Andy Bounds, 'Lithuania complains of trade 'sanctions' by China after Taiwan dispute', Financial Times, 3 December 2021.

2. The European Commission (2023) uses a definition which also includes "risks related to physical and cyber security of critical infrastructure" and "risks related to technology security and technology leakage". We would classify this as part of national security (within the 'other' category in Table 1) rather than economic security.

3. Trump has announced that he would implement a 10 percent across-the-board tariff. This would affect EU exports significantly, in addition to US importers. See Charlie Savage, Jonathan Swan and Maggie Haberman, 'A New Tax on Imports and a Split From China: Trump's 2025 Trade Agenda', New York Times, 26 December 2023.

4. See European Commission, 'Trade defence', undated.

5. See European Commission, 'The Foreign Subsidies Regulation in a nutshell', undated.

6. See European Commission, 'European Chips Act', undated.

7. See European Commission, 'Net-Zero Industry Act', undated.

8. See European Commission, 'Temporary Crisis and Transition Framework', undated.

9. See European Commission, 'Critical Raw Materials Act', undated.

10. See European Commission, 'Health Emergency Preparedness and Response (HERA)', undated.

11. Final compromise text agreed in February 2024 available at https://data.consilium.europa.eu/doc/document/ST-6336-2024-INIT/en/pdf.

12. Namely, photovoltaic and solar thermal, onshore wind and offshore renewables, batteries and storage, heat pumps and geothermal energy, electrolysers and fuel cells, sustainable biogas and biomethane, carbon capture and storage (CCS) and grid technologies.

13. Article XI of the 1994 General Agreement on Tariffs and Trade prohibits quantitative export restrictions (with certain exceptions) but permits "duties, taxes or other charges". See https://www.wto.org/english/res_e/publications_e/ai17_e/gatt1994_art11_oth.pdf.

References

Anderson, J (2020) 'Not all foreign investment is welcome in Europe', Bruegel Blog, 10 November.

Baqaee, DJ Hinz, B Moll, M Schularick, FA Teti, J Wanner and S Yang (2024) 'What if? The Effects of a Hard Decoupling from China on the German Economy', chapter 4 in J Pisani-Ferry, B Weder Di Mauro and J Zettelmeyer (eds) Paris Report 2: Europe's Economic Security, CEPR Press.

Bown, C (2024) 'Trade policy, industrial policy, and the economic security of the European Union', chapter 5 in J Pisani-Ferry, B Weder Di Mauro and J Zettelmeyer (eds) Paris Report 2: Europe's Economic Security, CEPR Press.

European Commission (2021) 'Strategic dependencies and capacities', SWD(2021) 352 final.

European Commission (2022) 'EU strategic dependencies and capacities: second stage of in-depth reviews', SWD(2022) 41 final.

European Commission (2023) 'Auctions-as-a-Service for Member States, Concept Note', Directorate- General Climate Action.

González Laya, A, C Grand, K Pisarska, N Tocci and G Wolff (2024) 'Trump-Proofing Europe: How the Continent Can Prepare for American Abandonment'.

Kelly, M and KH O'Rourke (2024) 'Industrial policy in the shadow of conflict: Lessons from the past', chapter 2 in J Pisani-Ferry, B Weder Di Mauro and J Zettelmeyer (eds) Paris Report 2: Europe's Economic Security, CEPR Press.

Le Mouel, M and N Poitiers (2023) 'Why Europe's critical raw materials strategy has to be international', Analysis, 5 April, Bruegel.

Mavroidis, P and A Sapir (2024) Key New Factors likely to shape the new EU Trade agenda in the next term, study for the European Parliament International Trade Committee.

McCaffrey, C and NF Poitiers (2024) 'Instruments of economic security', chapter 5 in J Pisani-Ferry, B Weder Di Mauro and J Zettelmeyer (eds) Paris Report 2: Europe's Economic Security, CEPR Press.

Mejean, I and P Rousseaux (2024) 'Identifying European trade dependencies', chapter 3 in J Pisani-Ferry, B Weder Di Mauro and J Zettelmeyer (eds) Paris Report 2: Europe's Economic Security, CEPR Press.

Murphy, KM and RH Topel (2013) 'Some Basic Economics of National Security', American Economic Review 103(3): 508–11
Pinkus, D, J Pisani-Ferry, S Tagliapietra, R Veugelers, G Zachmann and J Zettelmeyer (2024) Coordination for EU competitiveness, study for the European Parliament Economic and Monetary Affairs Committee.
Pisani-Ferry, J, B Weder di Mauro and J Zettelmeyer (eds) (2024) Paris Report 2: Europe's Economic Security, CEPR Press.
Poitiers, N and P Weil (2022) 'Is the EU Chips Act the right approach?' Bruegel Blog, 2 June.
Tagliapietra, S, R Veugelers and J Zettelmeyer (2023) 'Rebooting the European Union's Net Zero Industry Act', Policy Brief 15/2023, Bruegel.

This Policy Brief is based on chapter 1 of Pisani-Ferry et al (2024). We thank Chad Bown, Julian Hinz, Morgan Kelly, Conor McCaffrey, Isabelle Mejean, Kevin O'Rourke, Pierre Rousseaux and Moritz Schularick, as well as Shekar Aiyar, Alicia García-Herrero, Petros Mavroidis, Francesco Papadia, André Sapir, Fiona Scott Morton, Nicolas Véron, Lennard Welslau and Guntram Wolff for helpful discussions and comments. We are particularly grateful to Niclas Poitiers for his contribution to the survey of EU economic security instruments that appears in section 4 of this paper. This article is based on a Bruegel Policy Brief Issue n°07/24 | May 2024.

The rocky road to EU accession

The Western Balkan countries are moving towards European Union accession. Armin Steinbach examines the obstacles and lessons from the Eastern Partnership

Executive summary

Th e Western Balkan countries and the countries of the Eastern Partnership are moving towards European Union accession at different speeds. We explore whether and how the variable speed towards EU accession can be traced to different legal regimes governing European integration: Stabilisation and Association Agreements (SAA) for the Western Balkan countries, and Deep and Comprehensive Free Trade Agreements (DCFTA) for the countries of the Eastern Partnership (EaP).

We find that DCFTAs apply more lenient conditionality to intra-regional cooperation. They subject non-tariff barriers to a more explicit regime than the Western Balkan SAAs. The DCFTAs also off er a more rigid and comprehensive approach to the approximation of laws than the SAAs, and the DCFTAs are more inclusive with regard to the role of civil society.

However, there is no indication that the differences in legal governance have translated into stronger economic performance in the EaP countries or greater integration with the EU, compared to the Western Balkans.

The Western Balkan countries remain significantly more integrated than the EaP countries with the EU in trade terms, while convergence with the EU has been stagnating both for the Western Balkan and the EaP countries. Economic shortcomings in the Western Balkan still need to be addressed.

Conditionality attached to both integration into the EU single market and EU funding should be nuanced; the eradication of non-tariff barriers should be prioritised both inter-regionally and intra-regionally between Western Balkan countries; the need for stronger EU investment in the region is reinforced by geopolitical concerns about

Chinese investments coming without EU-type conditionality attached; and governance should give a stronger role to civil society.

In order to address the shortcomings in SAAs, a pragmatic solution is to use the existing governance framework under the SAAs.

The importance of EU single market membership to West Balkan economic prospects cannot be overstated

1 Introduction

Until the Russian invasion of Ukraine, the European Union pursued a two-track approach to its south-eastern and eastern European neighbours. The EU accession prospects of the Western Balkan (WB) states (Albania, Bosnia and Herzegovina, North Macedonia, Montenegro, Kosovo and Serbia) were more promising than those of their eastern counterparts – in particular Georgia, Moldova and Ukraine – which were associated with the EU through its Eastern Partnership (EaP).

Georgia, Moldova and Ukraine declared they wanted to join the EU in the mid-2000s, but for a long time the EU preferred alternative models: first the European Neighbourhood Policy (in 2004) and then the EaP (in 2009). But though the EU pursued an integration model in relation to the EaP that did not aim at EU accession, Russia's war against Ukraine triggered a change to this two-track approach.

Suddenly, the process, at least with Ukraine, Geogia and Moldova (which are the reference point of comparison with the WB in this paper), turned into an accession process, ushering in the initiation of accession negotiations with Ukraine and Moldova in December 2023.

The three eastern European states had practically no waiting time before being accepted as candidate countries right after application (Box 1). This contrasts with the Western Balkans, with either, as for North Macedonia, a decade of waiting for the opening of accession negotiations because of resistance from some EU member states or, as for Serbia, a decade of dragging negotiations because of democratic backsliding.

As the progress report in Box 1 shows, given that WB applications to accede to the EU date back as far as 2004, the accession process has advanced much more slowly than for the EaP countries that applied only in 2022.

Yet, new impetus has spilled over to the WB, as the EU opened accession talks with Albania and North Macedonia in July 2023 and with Bosnia and Herzegovina in March 2024, while Kosovo officially submitted its membership application in 2022.

The new 'reversed order' of accession, with Ukraine seemingly outpacing the WB since 2022, adds to a dissatisfaction with the WB accession process that has long been growing. Among WB countries, the dominant perception was that the EU promise of WB membership was not credible, while the EU felt persistently concerned about the lack of *"genuine domestic reforms"* and remaining political rifts in the region (Dabrowski, 2022).

Ukraine's rapid move towards accession raises the question – notwithstanding the political accelerator for Ukrainian accession arising from the Russian assault – whether there are lessons to be learned from the new 'front runners'[1].

With the relationship between the EU and Ukraine, Georgia, and Moldova now governed by a different set of agreements and governance, this paper explores possible differences between the relationships the two blocs have with the EU.

It has been argued – but not analysed in depth – that the Deep and Comprehensive Free Trade Agreements (DCFTA) led to Ukraine, Georgia and Moldova being better integrated with the EU in terms of their access to its markets, than the Stabilisation and Association Agreements (SAAs) did for WB countries (Blockmans, 2018). The DCFTAs form part of the countries' Association Agreements with the EU and supplement and deepen their integration into the EU internal market.

Our analysis explores more deeply the comparison between the two groups of agreements. Clearly, we consider the pre-war situation and as such exclude that war-related geopolitical factors changed the accession pace of EaP countries, and of Ukraine in particular.

Specifically, we seek to better understand the differences in regimes and access to the EU internal market. First, we systematically assess and compare the substantive, procedural and institutional differences between the eastern European AA/DCFTAs and the WB SAAs with respect to their potential in offering access to the EU internal market.

Despite large similarities between the agreements, we find considerable differences in legal governance related to conditionality, non-tariff barriers of trade, trade in services, foreign direct investment (FDI) and the approximation of laws. We extend the comparative analysis to shortcomings in the governance and implementation process of the relevant SAAs and working plans.

Second, in view of the differences, we explore the extent to which they may have had an impact on economic performance in terms of convergence with the EU, trade in goods and services, non-tariff barriers, FDI and what measures should be implemented to overcome the existing shortcomings.

These could be implemented either by modifying the WB SAAs or through modifications to the level of technical implementation. We caution against claiming a causal effect in terms of the differences in legal governance leading to Ukraine to obtain the status of accession negotiations so rapidly (geopolitical reasons are likely to trump the modest performance of Ukraine, for example).

Our analysis comes at a critical time. Political sentiment in some WB countries, particularly Serbia and North Macedonia, blames the EU for slow accession, while democratic backsliding and authoritarian regimes in the WB is leading to China and Russia, as underpinned by an influx of Chinese FDI (Figure 7), to be seen as alternatives to moving closer to the EU, with the EU portrayed as just one among the external players in the region (Vulović, 2023).

The new Growth Plan (European Commission, 2023a) and the draft Reform and Growth Facility for the Western Balkans (European Commission, 2023b) seek to revive WB integration. While additional funding for the region will be made available, the new proposal brings a demanding degree of conditionality, increasing the pressure for domestic reforms (in line with the EU Copenhagen, or accession, criteria), and setting additional intra-regional integration as cumbersome preconditions, both for internal market access and funding eligibility.

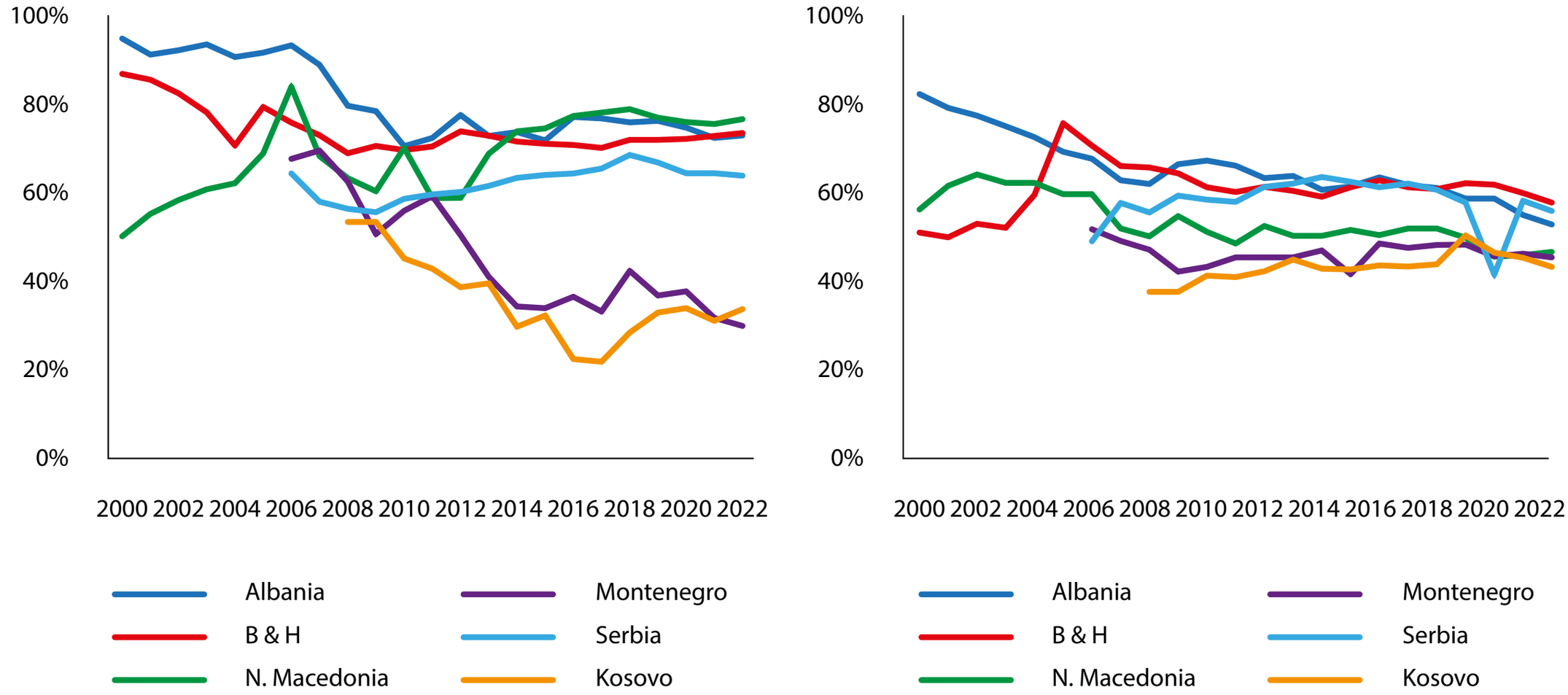
Yet, the current negotiations of a roadmap for Ukraine's accession to the EU may offer a new momentum for the WB states to integrate further into the EU single market, by underlining the mutual benefits. The new geopolitical reality enhances the significance of the EU's enlargement policy, but for it to materialise, it requires modification of the current regime governing market access, financial investment and governance.

We focus on access to the single market both from the perspective of substantive market access and governance of the implementation. The EU is the key trading partner of the Western Balkans, with WB goods exports to, and imports from, the EU in 2022 amounting to €37 billion and €48 billion respectively (equating to simple averages of approximately 59 percent and 49 percent of their respective trade totals; Figure 1). Services trade between the two is also significant, with exports to and imports from the EU amounting to approximately €8.5 billion and €7.5 billion respectively for the same year (Figure 6)[2].

However, the WB share of exports to and imports from the EU27 has been constant in average over the last twenty years. Since the sequential entry into force of SAAs since 2004 there has not been a significant increase in trade integration with the EU. In turn, the share of the EU as an export destination for EaP goods has on average increased (Figure 1b).

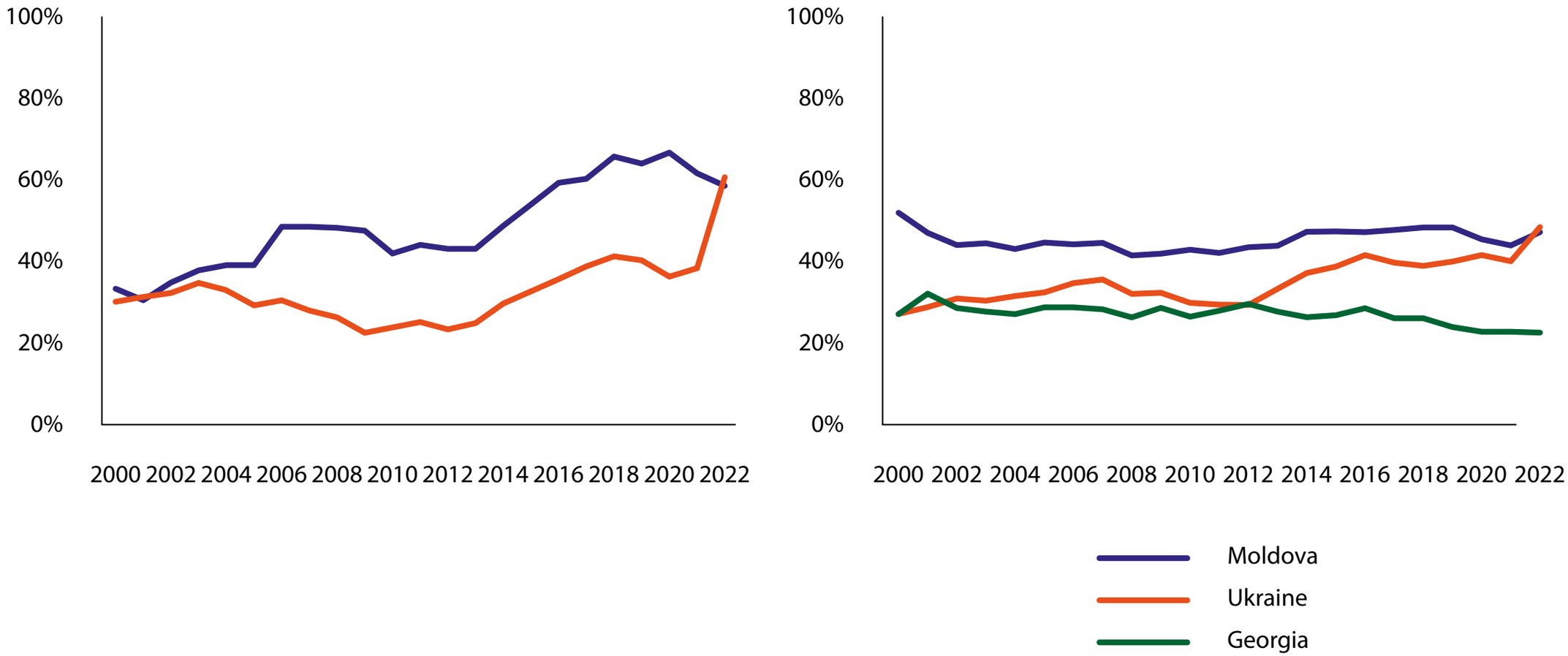
At the same time, the rate of convergence of the Western Balkans countries was described in the new Growth Plan as *"not satisfactory"* and *"holding back their progress on the EU track"* (European Commission, 2023, p.1).

Figure 1a. The EU as an export destination (left) and import source (right) for WB goods (% of total exports and imports respectively)



Source: Bruegel based on IMF Direction of Trade Statistics.

Figure 1b. The EU as an export destination (left) and import source (right) for EaP goods (% of total exports and imports respectively)



Source: Bruegel based on IMF Direction of Trade Statistics.

As illustrated in Figure 2, both regions have struggled with GDP per capita convergence to the EU27 average, recording moderate gains between 2011 and 2021. WB countries had higher initial GDP per capita level than the EaP countries (by approximately 10 percentage points of average EU27 GDP) but caught up less quickly up to 2021. In 2022, Ukraine and Moldova recorded reversals of their previous growth trends, because of Russian aggression against Ukraine.

The stagnating share of the EU27 in trade with the WB, and the moderate pace of convergence, provide the economic motivation for our analysis and for exploration of a possible connection to the legal regime set out in the SAAs.

Based on our comparative legal and institutional analysis, we identify a number of differences between the agreements the EU concluded with the eastern European countries and the WB. Yet while differences in the legal governance of DCFTAs and SAAs would suggest WB economic underperformance compared to the EaP, because of a legal framework limiting WB integration into EU internal market in comparative perspective with the DCFTAs, this is not supported by the available economic evidence.

While these differences are significant deficiencies and should be addressed, we hasten to say that there is no compelling evidence that remaining shortcomings can causally been traced to the different legal treatments.

In any case, taking the DCFTAs as an example, the remaining constraints in the SAAs and in the new Growth Plan should be lifted to untap further potential for WB convergence with the EU internal market.

Figure 2a. GDP per capita in PPP (percent, EU27 = 100)

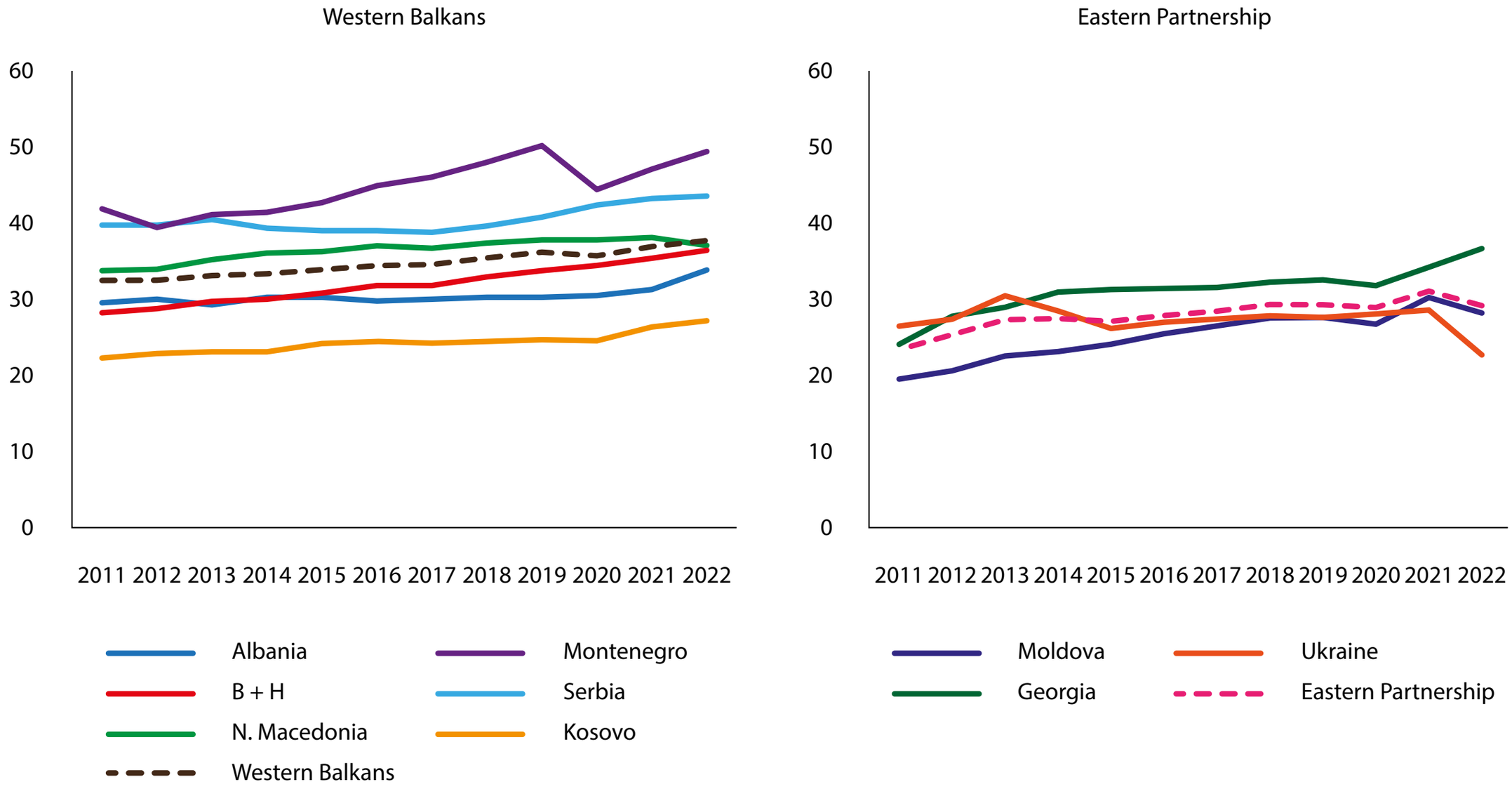
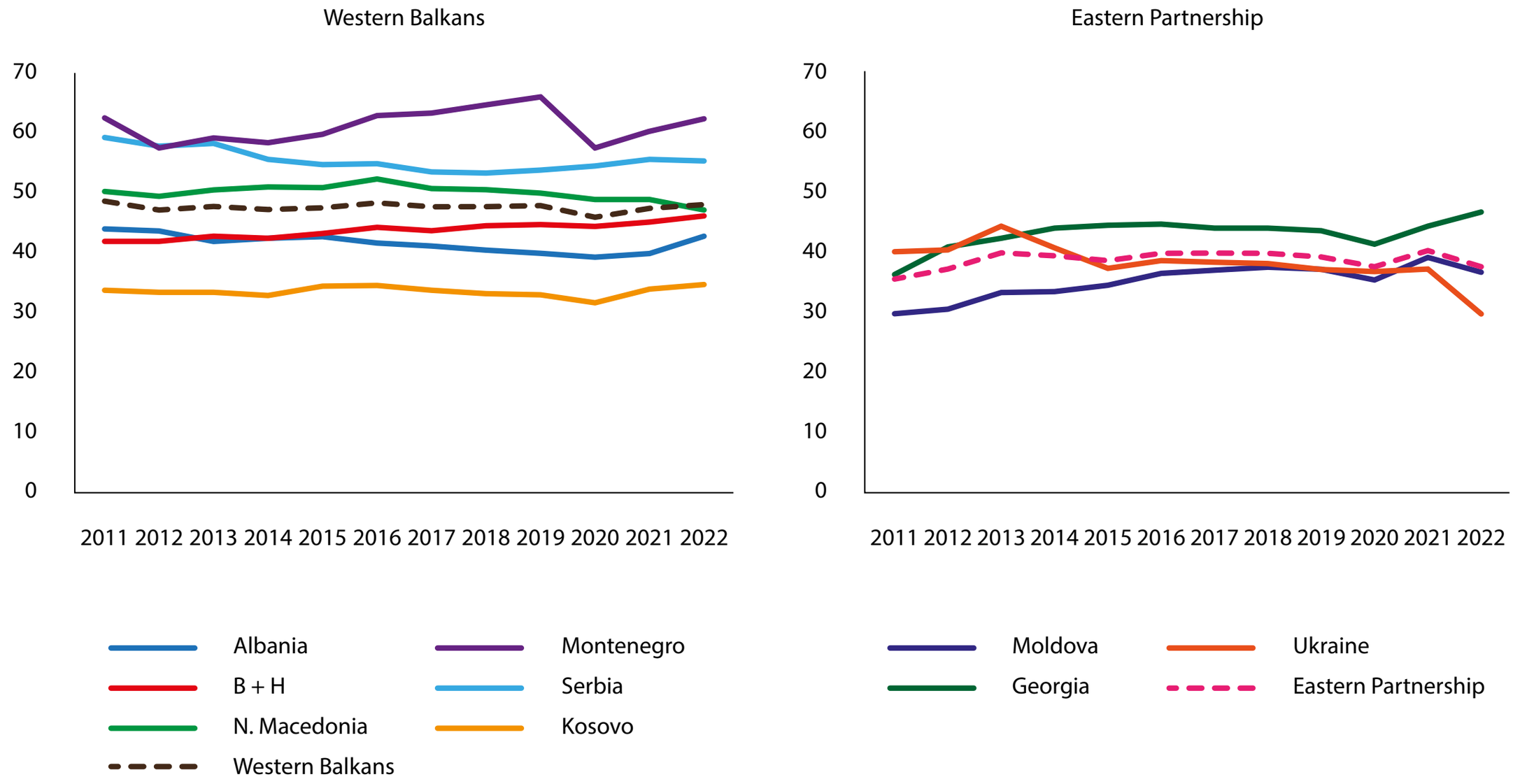


Figure 2b. GDP per capita in PPP (percent, 10 central and eastern European countries = 100)



Note: The Western Balkans and Eastern Partnerships dashed lines are simple averages. For an insight into convergence in the regions in general, a weighted approach to account for population may be more appropriate. However, the relevant metric for accession is the convergence of the countries in question, not the regions as a whole. These averages are only included for ease of comparison.
Source: Bruegel based on World Bank World Development Indicators.

Box 1. The nature and state of play of the Accession talks[3]

The EU accession process involves five main steps[4]. First, a country must apply to the Council of the EU to become a member. Article 49 of the Treaty on the European Union (TEU) stipulates that any European country that respects and commits to the values of the EU as expressed in Article 2 TEU can apply, and this is the stage that Kosovo is currently at.

The second step is a positive assessment of the Commission recommending the granting the candidate status. Third, candidate status is approved based on a unanimous decision of the European Council, which is what happened for Georgia for instance in December 2023. However, this does not necessarily mean that formal negotiations have been opened.

The fourth step is the accession negotiations, which begin with a detailed examination (screening) carried out by the Commission, together with the candidate country, of each policy field (chapter), to determine how well the country is prepared. This initial screening exercise of the EU's acquis serves to identify levels of preparedness in each policy field (which Albania and North Macedonia completed in 2023).

If completed satisfactorily, negotiations ensue focusing on six different thematic clusters, each consisting of various chapters; these negotiations take place at intergovernmental conferences (Montenegro, for instance, has opened negotiations on all chapters and closed three).

Fifth and finally, the process concludes when all chapters have been closed and an accession treaty is approved unanimously by the European Council and receives the consent of the European Parliament. Each EU country must also ratify the treaty according to its constitutional procedures (Dabrowski, 2014).

Country	Stage of process (early 2024)	State of play	Next step(s)
Western Balkans			
Albania	Applied for membership in 2009; candidate country since 2014; accession negotiations began in 2022[5].	The screening meetings (ie. prior to accession negotiations entailing analytical examination of the EU acquis) were completed in November 2023.	First negotiation cluster will begin once the roadmaps identifying rule of law and public administration reforms are assessed and approved[6].
Bosnia and Herzegovina	Applied for membership in 2016; candidate country since 2022; accession negotiations opened in March 2024.	The Commission noted positive steps towards meeting key priorities and opening negotiations following the awarding of candidate country status, but recent rule of law developments have proved a barrier.	Preparation of the negotiating framework.

Kosovo	Applied for membership in 2022; currently a potential candidate country[7].	The European Reform Agenda was adopted in 2016 and updated in 2021 between the Commission and Kosovo to guide the implementation of SAA reforms. Due to a lack of de-escalatory measures regarding rising tensions with Serbia, the EU froze various cooperation and funding mechanisms in 2023 (European Commission, 2023d).	The frozen measures are temporary and will be reversed if and when authorities take satisfactory de-escalatory steps and implement commitments related to Serbia. The next steps of the accession process are unclear.
Montenegro	Applied for membership in 2008; candidate country since 2010; accession negotiations began in 2012.	Since 2012, all negotiating chapters have been opened, with three closed. The enlargement methodology was revised in 2021 to place more emphasis on fundamental reforms and reinvigorate the process.	Further progress on the rule of law chapters is necessary before any others are provisionally closed.

North Macedonia	Applied for membership in 2004; candidate country since 2005; accession negotiations began in 2022.	The screening meetings were concluded in December 2023.	First negotiation cluster will begin once the roadmaps identifying rule of law and public administration reforms are assessed and approved[8].
Serbia	Applied for membership in 2009; candidate country since 2012; accession negotiations began in 2014.	Since 2014, 22 negotiating chapters have been opened, with two closed. The enlargement methodology was revised in 2021 to place more emphasis on fundamental reforms and reinvigorate the process.	The rate of progress in the rule of law chapters and in the normalisation of relations and de-escalation with Kosovo dictate the pace of negotiations.
Eastern Partnership			
Georgia	Applied for membership in 2022; candidate country since December 2023; accession negotiations yet to begin.	Due to progress on the 12 identified priorities since the application was made, candidate country status was granted on the understanding that nine steps would be taken.	Progress must continue on the nine steps detailed in the November 2023 Communication[9] on enlargement.

Moldova	Applied for membership in March 2022; candidate country since June 2022; Council decided to open accession negotiations in December 2023.	In the June 2022 Commission Opinion (European Commission, 2022a) on Moldova's application recommended to grant candidate status on the understanding that nine steps were taken. As of November 2023, six of the nine steps were completed.	Accession negotiation framework will be adopted once the three recommendations in the November 2023 Communication[10] on enlargement are completed. Screening began in January 2024[11].
Ukraine	Applied for membership in March 2022; candidate country since June 2022; Council decided to open accession negotiations in December 2023.	June 2022 Commission Opinion (European Commission, 2022b) on Ukraine's application recommended to grant candidate status on the understanding that nine steps were taken. As of November 2023, six of the nine steps were completed.	Accession negotiation framework will be adopted once the four recommendations in the November 2023 Communication[12] on enlargement are completed. Screening began in January 2024.

2 Comparing DCFTAs and the Western Balkan SAAs in terms of EU market integration

This section highlights differences between the legal regimes governing market access for the eastern European countries of Ukraine, Moldova and Georgia (on basis of DCFTAs) and the applicable framework under the Western Balkan SAAs. Differences are explored in relation to five benchmarks: conditionality, non-tariff barriers to trade, trade in services, movement of capital and the approximation of laws.

Annex I provides a comprehensive comparative assessment of the relevant agreements and the applicable rules, while this section discusses some of the marked differences. What facilitates the comparison (while highlighting the stark differences between the regimes) is a large degree of homogeneity in agreements within each group – within DCFTAs and Western Balkan SAAs. For the purpose of making comparisons, the Serbia SAA[13] will be the reference point for the WB SAAs, while the Ukraine AA/DCFTA[14] is referred to to exemplify the agreements the EU concluded with the eastern European partners.

2.1 Regional integration as conditionality

One core distinguishing feature between the DCFTA and the WB SAAs is the degree of conditionality attached to intra-regional integration as a precondition for further access to the EU internal market.

Most recently, this emphasis has been reiterated in the draft New Growth Plan, which, as an extension of the WB SAAs, makes single market access conditional not only on political and economic domestic structural reforms, but on the progress made in intra-regional market integration.

The Serbia SAA emphasises regional cooperation by requiring the WB country to *"enhance its cooperation"* and to *"implement fully the CEFTA"* (Article 14 Serbia SAA) – the Central European Free Trade Agreement governing trade relations between the WB states.

The Serbia SAA further requires the conclusion of additional bilateral conventions with WB countries that foster political dialogue, establish free trade, cooperation in justice affairs and provide free market access more globally (Article 15 Serbia SAA).

This conditionality has been constantly upheld in the EU's policy on the WBs, with the most recent draft Growth Plan tying access to EU internal market benefits and the release of funds under the draft Reform and Growth Facility (the financial assistance vehicle of the plan) (European Commission, 2023b) to a wide set of reforms.

This extends not only to traditional conditionality securing the Copenhagen criteria, including democracy, rule of law and human rights (which apply to WB and EaP countries alike). In the case of WB, the political conditionality also extends to requiring Serbia and Kosovo to normalise their relations and comply with the relevant agreements governing reconciliation, and to negotiate the Comprehensive Agreement on normalisation of relations (European Commission, 2023b, Article 5).

Importantly and in addition, the EU requests economic intra-regional integration as precondition and conditionality attached to access to the EU single market. For example, the Commission envisages making access to EU financial support through its draft Reform and Growth Facility (European Commission, 2023b) conditional ex ante on the implementation of the Common Regional Market Action Plan.

This Plan is the outcome of the Common Regional Market Initiative of the WB countries, which builds on the CEFTA framework (and thus connects to the conditionality embedded in the SAA). The Plan requires, inter alia, the development of a regional digital market, which requires investment in broadband internet access, 5G and digital services.

The Plan also foresees expansion of green lanes at the border to cut waiting times. Hence, the extended conditionality regime allows the EU to make internal market access and access to funding conditional on WB ex-ante investment in these areas.

This conditionality contrasts with the absence of mandatory regional cooperation under the DCFTAs. The agreements are silent on this type of intra-regional conditionality. Specifically, the Ukraine-DCFTA provides for *"regional stability"*, stipulating a vague obligation for Ukraine, Moldova and Georgia to *"intensify their joint efforts to promote stability, security and democratic development in their common neighbourhood"* (Article 9 DCFTA Ukraine).

The main conditionality in the Ukraine-DCFTA is the approximation of the relevant EU law by Ukraine along with the Copenhagen criteria, which must be respected by all EU aspirants. However, the DCFTAs lack the intra-regional layer of conditionality that the EU, in relation to the WB, has increasingly insisted on.

Not only are the DCFTAs lenient on regional integration as a requirement, the question is also whether the EU's persistent insistence on regional economic cooperation is an adequate requirement. Intra-regional conditionality is plausible if it seeks to alleviate political rifts between Serbia and Kosovo, and societal tension and political blockages in decision-making (European Commission, 2023a; Ghodsi *et al* 2022). But the economic intra-regional conditionality referred to above appears much more ambivalent.

On one side, creating a common regional market for goods, services and labour within the Western Balkans offers opportunities for increased trade – according to one estimate[15], regional economic integration in the Western Balkans could generate up to 2.5 percent of GDP growth, should the level of integration reach the level of that of the European Free Trade Association (EFTA), while it could even generate up to 7 percent should it reach the EU's level of integration.

The most ambitious initiative negotiated in this regard is the creation of the Common Regional Market[16] as an outcome of the Berlin Process, launched in 2020[17]. It foresees WB intra-regional freedoms of goods, services, capital and people, including aspects relating to digital, investment, innovation and industry policy.

On the other side, barriers to intra-regional economic integration lie in the lacking physical infrastructure and persistent inequality in the WB. In particular, lack of public investment in roads, digital infrastructure, railways and energy have been identified as limiting factors (Ghodsi *et al* 2022).

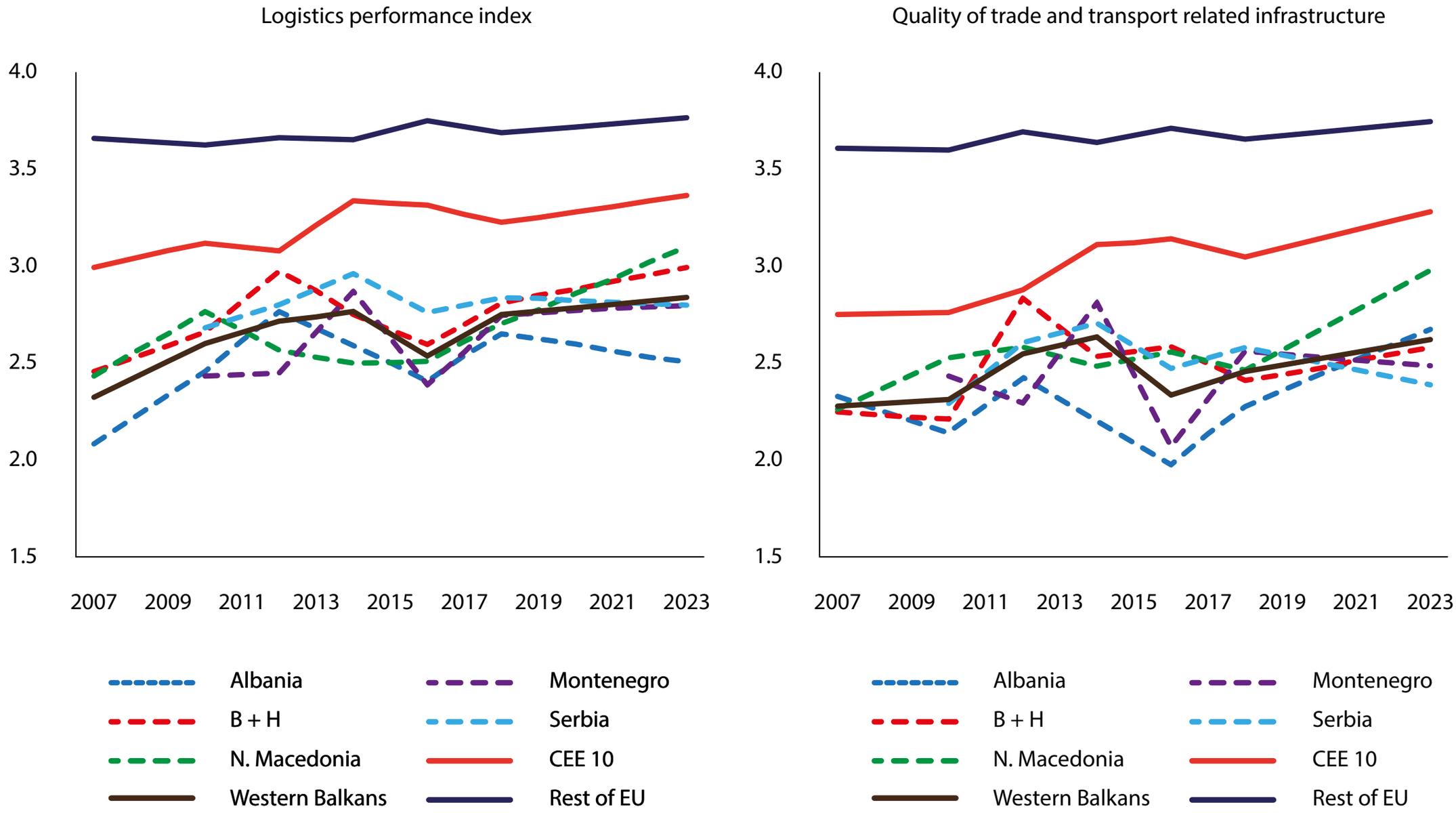
The Commission itself noted in its November 2023 Communication on enlargement that *"there is a strong need to upgrade infrastructure; investments should be… consistent with the priorities agreed with the EU"* (European Commission, 2023c, p.11).

Panel B of Figure 3 highlights the limited progress achieved on improving the trade-related intra-regional infrastructure and in closing the gap with the EU, using the broader logistics performance index[18] (Figure 3, Panel A), similarly showing low levels of convergence.

Even the central and eastern European EU members (a more adequate group for comparison with WB countries) seem to have been more successful in improving trade-related infrastructure by reducing the gap with other EU members. However, convergence has not been better across the same indicators for the EaP countries (see Annex 4).

The connection to conditionality is that with limited public investment in infrastructure identified as one persistent barrier to regional integration in the WB[20], the EU should not implement ex-ante conditionality on WB public

Figure 3. Logistics and trade-related infrastructure



Note: Data is available for 2007, 2010, 2012, 2014, 2016, 2018 and 2023. Data for Serbia, Montenegro and Georgia unavailable for 2007. Data for Albania is unavailable for 2014. Data for Kosovo unavailable throughout. WBs is a simple average of the relevant countries. CEE 10 and Rest of EU refer to the simple averages of the central and eastern European countries that joined the EU in the 2000s[19] and the other 17 EU countries, respectively. See Annex 4 for the same exercise for EAP countries.
Source: Bruegel based on World Bank Logistics Performance Index.

investments in digital infrastructure or crossborder trade facilities, as set out in the Common Regional Market Action Plan (eg. lanes at borders or customs procedures).

The EU should fund these 'win-win' investments, which are beneficial to the WB and the EU alike, rather than blocking EU internal market access because of the lack of these investments. This concerns in particular crossborder infrastructure and networks that are often underfinanced because of a mismatch between costs and benefits and that are, under EU internal market standards, typically eligible for funding. WB infrastructure should be prioritised accordingly. Conditionality attached to these kinds of projects is not a sensible approach.

In fact, intra-regional crossborder transport infrastructure has significant positive spillovers, such as the potential to reduce income disparities across the EU and its neighbouring regions.

In this regard, it is positive that the draft Growth Plan implies revising the trans-European transport framework (TEN-T), in order to include a new corridor crossing the Western Balkan region (Western-East Mediterranean corridor), and the EU's recent Economic and Investment Plan for the Western Balkans offers financing of rail transport[21].

However, conditionality of the new Growth Plan should be relaxed for these infrastructure projects more generally and the involvement of European Investment Bank and the European Bank for Reconstruction and Development funding in the investment should be further facilitated (Ghodsi *et al* 2022).

Finally, conditionality should also be rethought in light of geopolitical rivalry. EU conditionality contrasts with Chinese investment in the region without strings attached, which makes Chinese FDI more attractive.

Again, the legal comparison of WB SAAs with the DCFTAs shows that the latter offer a more explicit acknowledgement of internal market integration. The Ukraine AA is explicit about its objective of bringing Ukraine into the EU internal market (Article I (d) of the Ukraine-DCFTA), while such an explicit recognition of this objective is absent from the Serbia SAA, in which language is limited to *"gradually develop a free trade area between the Community and Serbia"* (Article 1 (f) Serbia SAA).

While more assertive language in the agreements does not guarantee more favourable economic outcomes, specifying the objective in the agreement can bind the institutions under the SAA to work towards that goal.

2.2 Trade in goods and non-tariff barriers

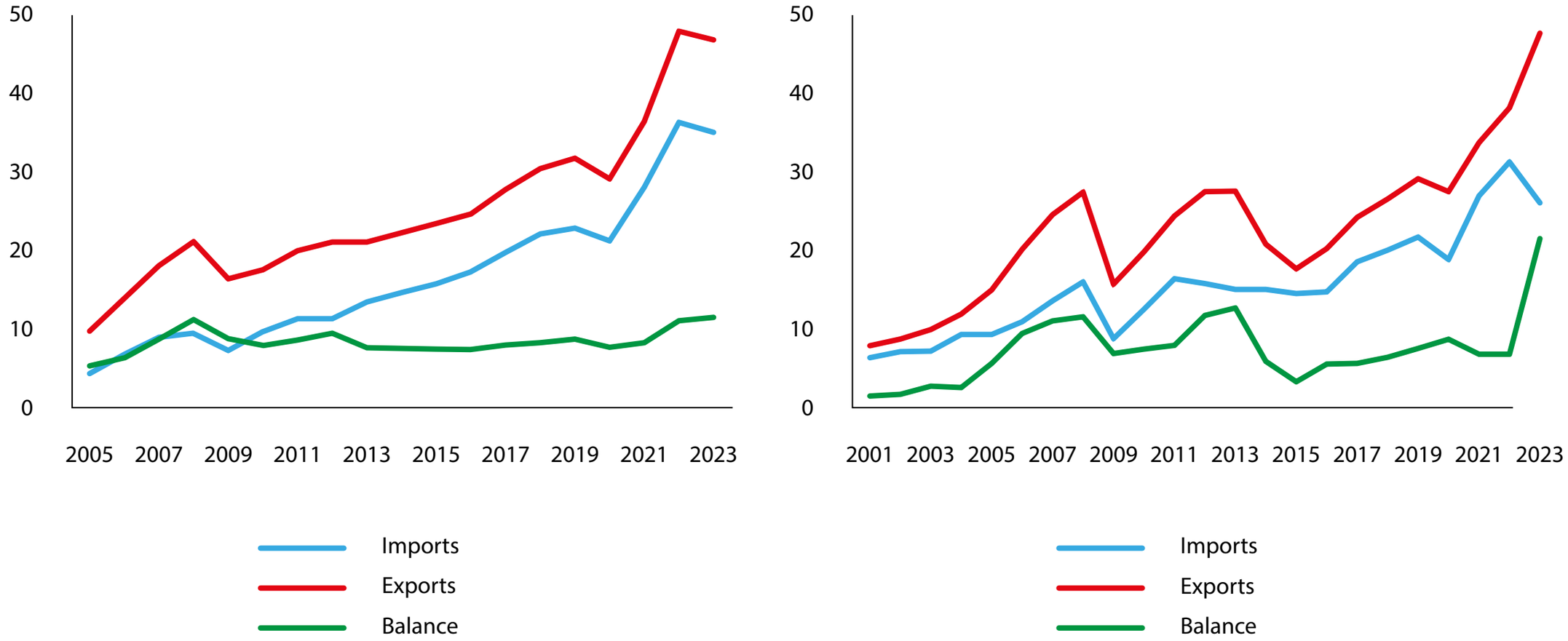
The EU-Ukraine association agreement has been praised by European Commission officials as *"the most ambitious Agreement that the EU has ever developed with any partner"*[22].

Indeed, by integrating the DCFTA into the Association Agreement, the integration of Ukraine, Georgia and Moldova into the EU has been propelled through wide-reaching market access and regulatory approximation, ushering in increased trade with the EU.

How do the agreements facilitate market integration? The WB SAAs have eliminated tariff barriers with the EU to a great extent, and trade with the region has grown by almost 130 percent over the past 10 years.

Figure 5 confirms that trade between the EU and WB has grown in absolute terms (though did not further increase the already high levels in relative terms, Figure 1), and there is no indication of being outpaced by the Eastern Partnership countries. Yet, non-tariff barriers (NTBs) remain significant – both barriers with the EU and within the Western Balkans region.

Figure 4. EU27 trade in goods with WBs (left) and EaP countries (right), € billions



Note: See Annex 2 for data disaggregated by country.
Source: Bruegel based on Eurostat (DS-018995).

NTBs can generally be associated with technical regulations, customs procedures, licensing requirements and other regulatory obstacles, all of which limit trade through increased costs, delays and administrative burdens.

For example, the waiting and processing time only at crossing points in CEFTA states generates between €250 million and €300 million in costs annually (World Bank, 2015). While reliable data on the scope of NTBs is limited, some proxies indicate their presence.

For instance, World Bank Trading Across Barriers data points to higher costs, both financial and in terms of time taken, associated with border and documentary compliance for importing goods to the Western Balkan countries, than to the EU or high-income OECD countries (Annex 5). While the same data limitations make it difficult to identify non-tariff barriers in EaP countries, the consensus is that they also pose challenges to trade in these countries.

Comparative legal analysis of the treatment of NTBs reveals a more detailed legal regime in the Ukraine DCFTA in three respects. First, the Serbia SAA does not foresee a non-discrimination rule regarding non-tariff measures, while the Ukraine DCFTA established a national treatment rule (Article 34).

It has been argued that the current legal reference to freedom of goods in the SSA should be interpreted in line with EU law and would thus suffice to ban non-tariff barriers (Sretić, 2023).

Second, the Ukraine DCFTA explicitly addresses technical barriers to trade (TBTs), in particular the *"adoption and application of technical regulations, standards, and conformity assessment procedures"* (Article 53).

Again, the Serbia-SAA is silent on the treatment of technical barriers to trade. The CEFTA addresses TBTs and provides for a governance structure to minimise them (Article 13). There have been further attempts to address

NTBs in the WB intra-regional integration process. For example, the Common Regional Market (CRM) has established green lanes at borders within the region.

Through better exchange of customs data before goods arrive at crossing points, the transit times for goods have greatly reduced (European Commission, 2023a). The draft Growth Plan, while requesting alignment with EU standards, does not foresee a regime to address further eradication of NTBs.

Yet overall the lack of salience of TBTs in the SAAs does not correspond to the significance of this source of impediment to market integration. Estimates suggest that a three-hour reduction in waiting times is the equivalent of a 2 percent reduction in tariffs (Del Mar Gomez *et al* 2023).

The OECD has considered the trade reducing effects of being outside the single market associated with TBTs and sanitary and phytosanitary measures (SPS) measures, suggesting these costs amount to 50 percent of the ad-valorem equivalent of measures on goods imported into the European Union from third countries (RCSPI, 2023). We infer that NTBs remain under-addressed at the level of the SAA agreements between WB countries and the EU.

Reducing NTBs is pivotal. Slow customs procedures are often the result of lacking infrastructure. For example, electronic payment of duties and charges and pre-arrival processing are essential infrastructure elements, lacking in all CEFTA economies. Serbia and Montenegro are reported not to offer the option of paying the fees for exports online (GIZ, 2022).

As argued above in relation to crossborder infrastructure and networks, infrastructure facilitating customs procedures should qualify for EU funding without (or with limited) conditionality, because the positive intra-regional economic effects are significant. The EU should allocate financial resources to the modernisation of such

facilities, in particular infrastructure that facilitates the payment of duties, taxes and other fees for the importation process.

In addition, mutual recognition also helps to reduce waiting times caused by scanning procedures and sample testing. The EU has created separate lanes with WB countries, and the same practice should be applied between WB countries (GIZ, 2022).

Again, where EU funding could facilitate this, there should be unconditional support for expanding joint crossing point facilities and establishment of separate lanes.

Likewise, concerning intra-regional commerce with 'mutual recognition' having proved itself as a motor for fostering intra-EU trade, WB countries should pursue recognition of conformity assessments procedures across the CEFTA region. The CEFTA provides the framework for this both in the field of SPS measures and NTBs more generally, but the available legal space under the agreement for eradicating NTBs (Articles 12, 13 CEFTA) should be exploited further.

In particular, Article 13 para. 4 CEFTA paves the way for WB countries to implement *"mutual recognition of conformity assessment procedures"*, offering a powerful tool for eliminating non-tariff barriers.

Finally, the EU should see advantages for itself not only in liberalising access to the internal market but also in outbound investment into the WB region. Access to the EU internal market and EU-financed crossborder infrastructure would reduce WB dependence on geopolitically risky partners.

For example, given Serbia's persistent dependence on Russian energy supplies, the EU should integrate the WB into its energy internal market by fostering the construction of electricity and gas connections – in the EU's own best interest and without conditionality.

At a time when economic security is becoming so important, helping to integrate the WBs into the supply chain could be very useful and help reduce dependencies. The Trans-Balkan electricity corridor is a good example[23], but further energy-oriented EU investments efforts could be directed to financing solar-energy capacity in the Western Balkans or wind and hydropower projects (Ghodsi *et al* 2022).

The EU can also do more to provide loan guarantees and investment incentives for private firms to invest in infrastructure in the region, in addition to tying this to reform and green agenda benchmarks. With EIB and ERBD expanding targeted loan guarantees to firms investing in these areas, the investment potential would be increased (Ghodsi *et al* 2022).

The draft Growth Facility aims at accelerating the green transition towards decarbonisation and to boost innovation, particularly for SMEs and in support of the green transition, yet no reference is made in the draft Facility to technological and industrial support to that end.

Energy-related infrastructure is an important policy field in view of the politically controversial energy dependence of WB countries on Russia (in particular Serbia). However, the CEFTA agreement is silent on issues of infrastructure, energy or gas supplies, leaving untapped a natural area of cooperation.

While integration into Europe's energy markets is part of the goals under the Serbia SAA (Article 109), there is no provision for translating these goals into substantive market access and specific cooperation obligations.

By contrast, the Ukraine DCFTA offers a comprehensive and substantive regime on energy, covering, inter alia, prohibition of trade-restrictive measures and striving for the emergence of energy markets (Article 338).

As long as there is no integration into EU energy markets in the WB, trade in energy will be constrained significantly by insufficient investment in transmission infrastructure and production capacity. China and Russia are likely to fill a void left by the EU, using state-driven investments in essential infrastructure in the WB (Stanicek, 2022).

Against this background, a proposal worth exploring on the level of implementation is to integrate the Western Balkans fully into the EU emissions trading system (ETS), which would accelerate the energy transition in the WB and be a significant new source of funding (Egenhofer, 2023).

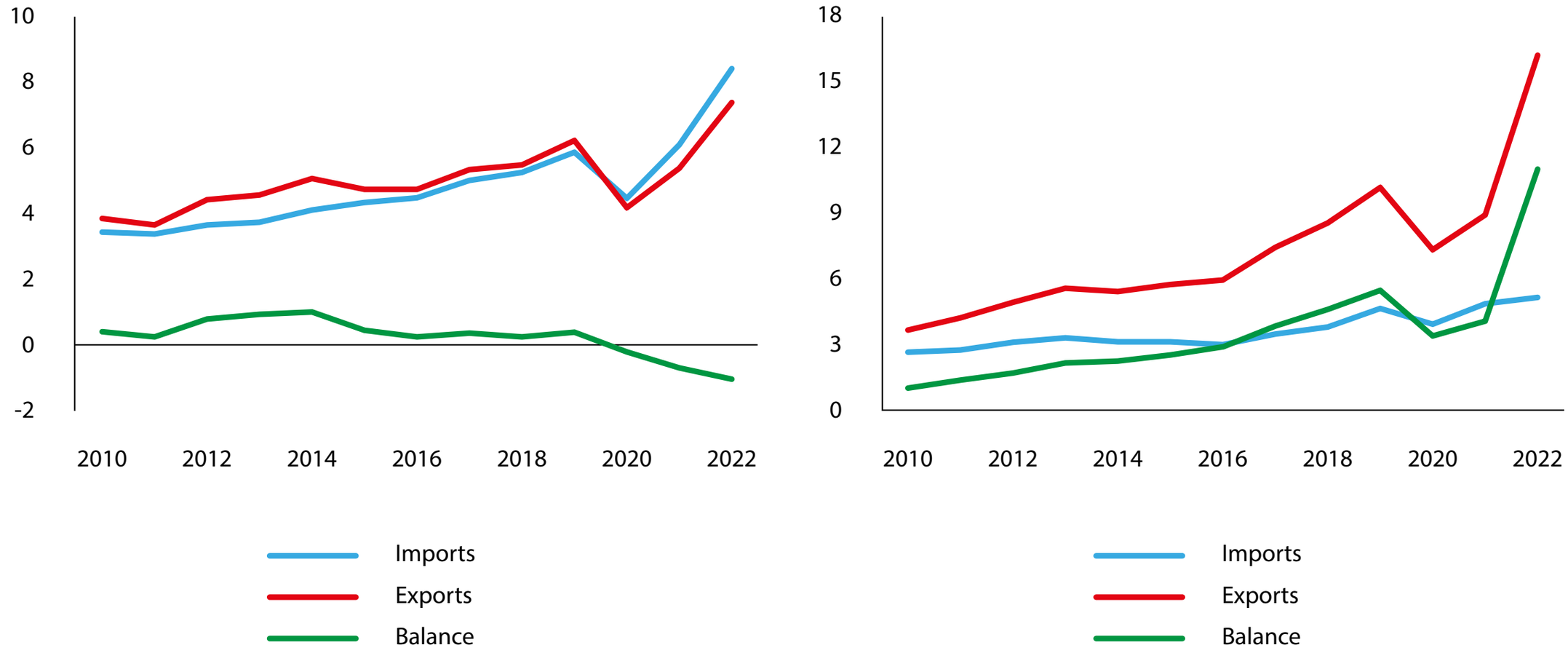
2.3 Freedom of services

From a comparative perspective, data on trade in services shown in Figure 6 indicates that WB services trade with the EU has grown less quickly than goods trade (compare with Figure 4). Also, EU services exports have grown more quickly with the EaP than with the WB, though from a very low basis.

One reason for this may be associated with the shortcomings in unleashing the potential of services, which can be illustrated by the inferior treatment of services in the Western Balkans SAAs compared to the Ukraine DCFTA. The EU-Ukraine DCFTA establishes a non-discrimination standard for Ukrainian services provided in the EU.

Specifically, these services must be granted *"treatment no less favourable"* than EU domestic services (Articles 93, 94). While this does not apply to all services, it extends to an extensive list of services. Consequently, the available evidence on Georgia supports the idea that its services sector has been expanded, with exports more than doubling in size since the entry into force of the DCFTA between 2014 and 2019 (Akhvlediani *et al* 2022).

Figure 5. EU27 trade in services with the Western Balkans (left) and EaP (right), € billions



Note: Data for Kosovo is not available. Data is presented from the perspective of the EU. See Annex 2 for data disaggregated by country.
Source: Bruegel based on Eurostat (bop_its6_det).

The Serbia SAA does not stipulate a no-discrimination principle similar to the Ukraine DCFTA. The Serbia SAA provides that the EU may not take measures that are *"significantly more restrictive"* than the situation before the Serbia SAA. It also provides procedurally for the EU and the WB to engage in *"steps to allow progressively the supply of services."*

Yet, this procedural potential has not so far been exploited, while substantive law liberalisation of services remains weak compared to the non-discrimination rule under the DCFTAs. Even the CEFTA does not provide unconditional liberalisation of services on intra-regional level.

The legal comparison points at the absence of rules providing for substantive discrimination prohibitions and the lack of regulatory harmonisation. This contrasts with the non-discrimination clearly spelled out in the agreement on trade in goods. Regulatory harmonisation (or mutual recognition) would be particularly beneficial in core service areas of the region, such as travel and transportation (RCSPI, 2023).

2.4 Capital movement

The EU accounts for approximately 60 percent of the current FDI stock in the Western Balkans[24], but there is no indication that FDI is treated more favourably in either the Western Balkan or the countries of Eastern Partnership.

The rules laid down in the relevant agreements indicate a high degree of capital movement freedom. Established through a ban on discrimination, capital movement is guaranteed both in the WB (Article 63 Serbia SAA) and in the Ukraine (Article 145 Ukraine DCFTA). Both types of agreements explicitly extend the free movement of capital to direct investments.

However, specific relevant sectors enjoy less-favourable treatment in the WB. For the financial sector, for example, DCFTA agreements offer an elaborate regime to promote the access of European investment in the Eastern partnership countries.

Access is granted to payment systems (Article 132 Ukraine DCFTA), regulatory approximation is required (Article 133) and bans on discrimination exist (Article 128). By contrast, the WB SAAs emphasise that financial services are subject to significant restrictions (Articles 54, 56 Serbia SAA).

Figure 6 shows that, much like for trade, EU FDI in the two regions is mainly into Serbia and Ukraine respectively (however, see Annex 3 for a breakdown of EU FDI into the various countries as a share of their GDP)[25].

The evidence suggests that FDI could be driven, more than the other freedoms we have discussed, not only by the openness of market access but by factors beyond the absence of barriers to moving capital. This is also evidenced by the experience of Bulgaria and Romania.

Both saw a one-time surge in FDI after accession to the EU, but have remained at pre-accession levels since. Rather, factors associated with state-driven investment and geopolitical competition have significant effects on FDI in the WB. The EU has historically been the dominant investor in the WB (See Annex 3).

In any case, a legal regime that secures non-discriminatory treatment of capital movement does not offer a complete picture on possible vulnerabilities related to FDI. This is so because state-funded, non-EU foreign investment increasingly outcompetes EU private investment. Some research points to a growing Chinese investment footprint in the region, especially in Serbia (Vulović, 2023; Bykova *et al* 2022), which seems to be driven

Figure 6a. EU27 FDI stock in the Western Balkans (€ billions)

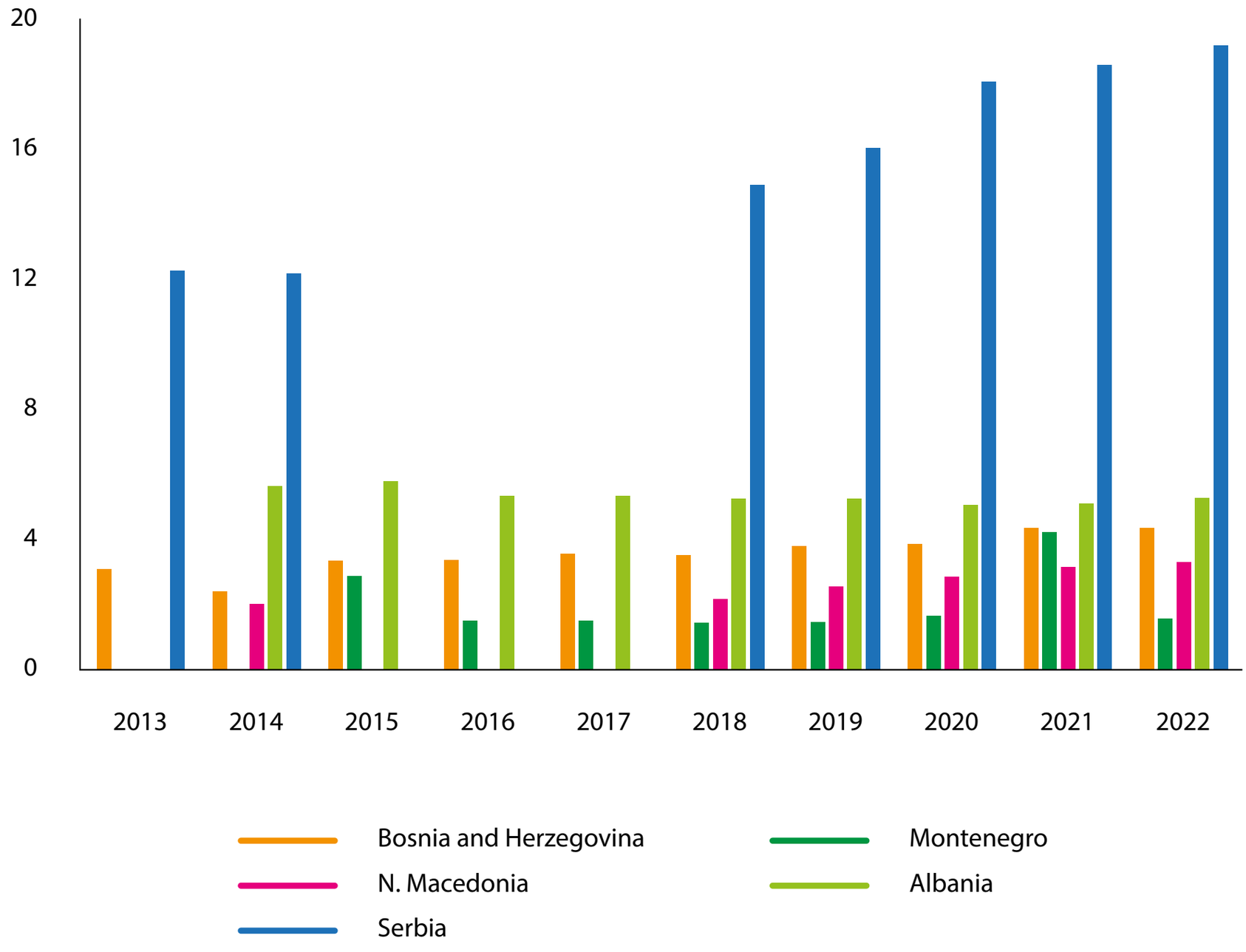
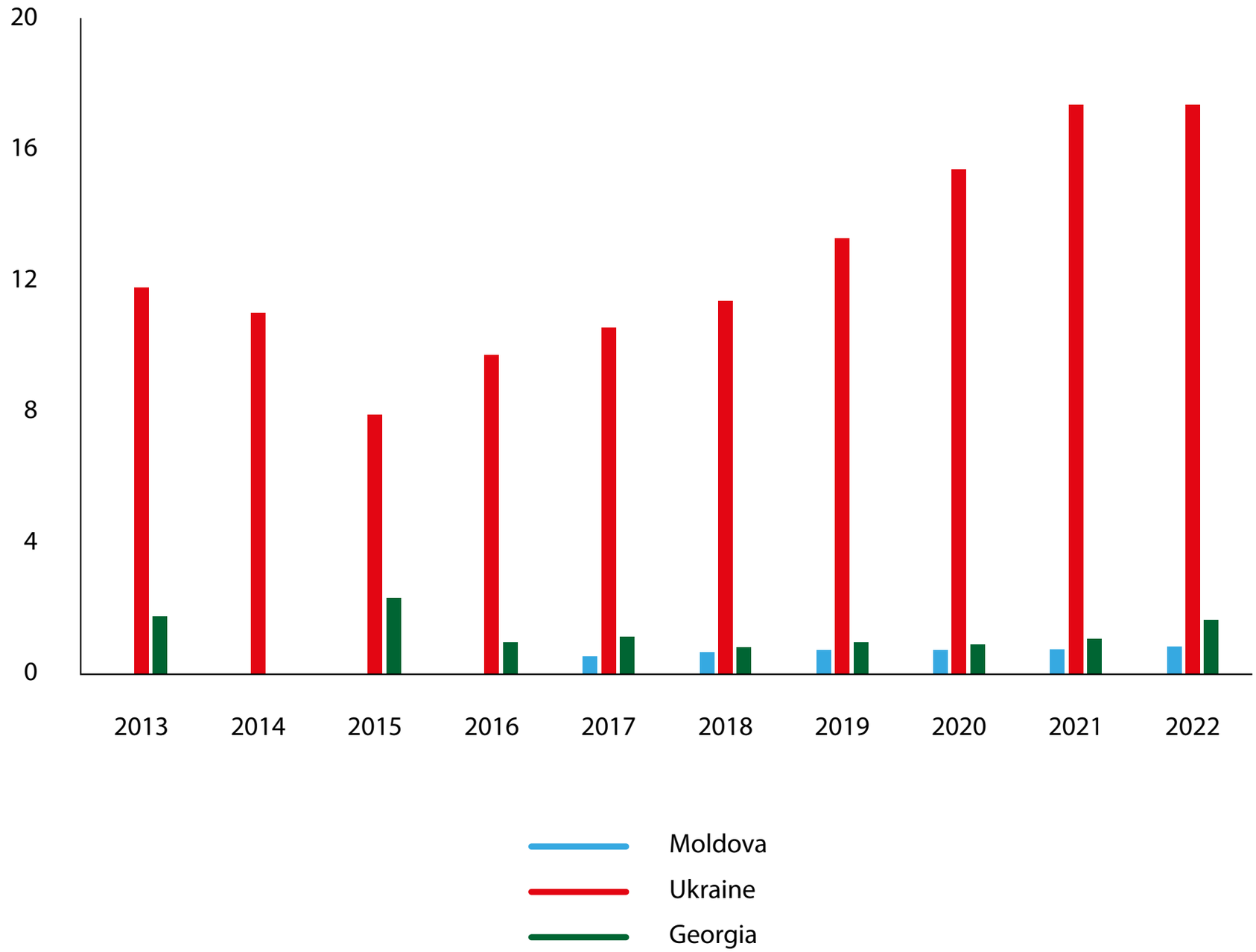
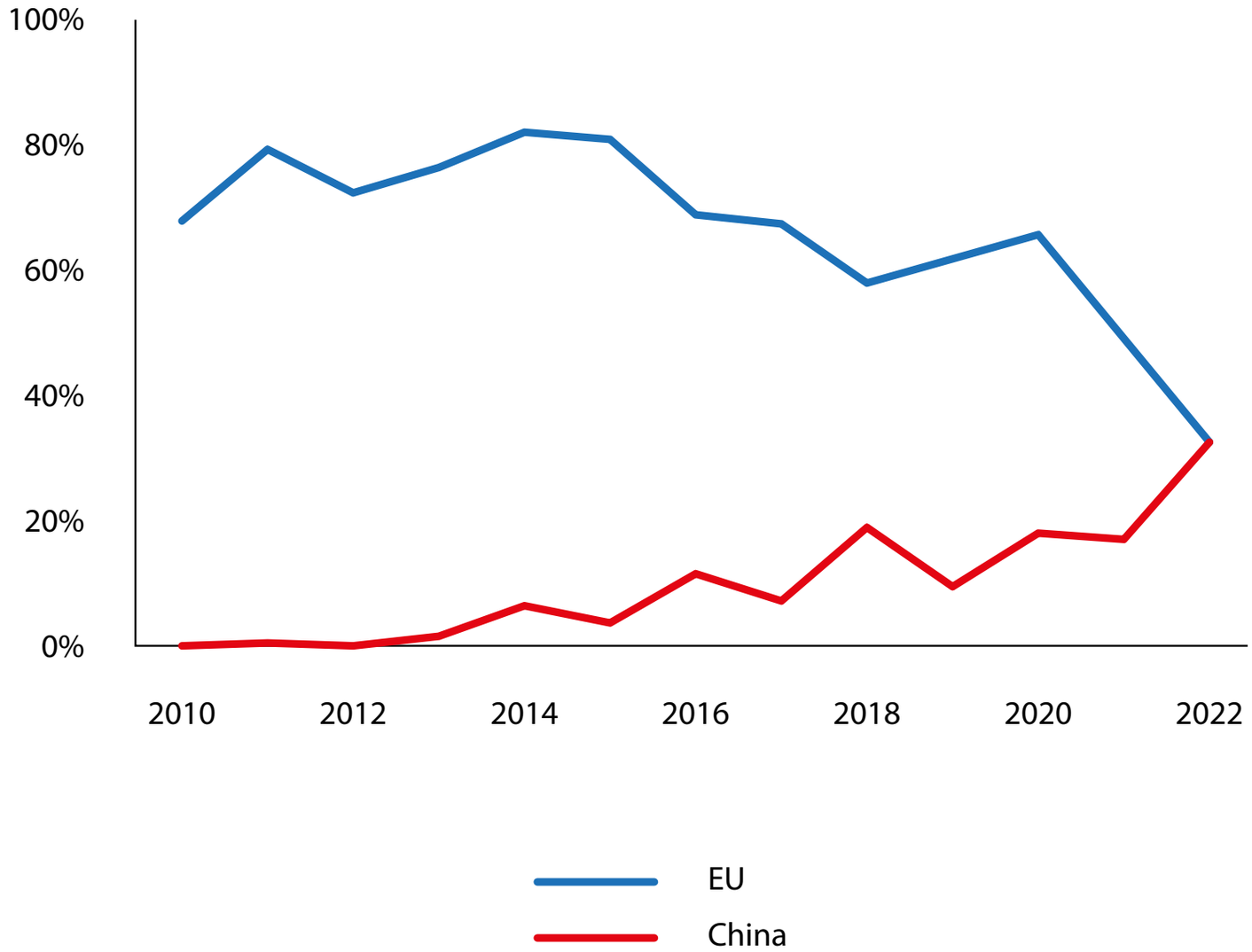


Figure 6b: EU27 FDI stock in the EaP (€ billions)



Note: The lack of data in some years is due to data not being reported by Eurostat for confidentiality purposes.
Source: Bruegel based on Eurostat (bop_fdi6_geo).

Figure 7. Share of net FDI flows to Serbia, 2010-2022



Note: The variable reported is the share of the EU27 and China net FDI in overall net FDI in Serbia. Net FDI is calculated as the difference between assets (Serbian residents' investments abroad) and liabilities (non-residents' investments in Serbia). Over this period there was consistently a larger inflow of investment into Serbia than outflow. This figure shows the share of that net inflow of FDI that comes from the EU27 and China.
Source: Bruegel based on National Bank of Serbia[26].

by state-owned investors or by state-guaranteed finance linked to contract guarantees for Chinese companies (Ghodsi *et al* 2022).

Indeed, this increase in Chinese investment in Serbia is supported by China's growing share in net FDI flows to Serbia (Figure 7).

Dependence on countries perceived (from a European perspective) as geopolitical rivals increases the WB's vulnerability to geopolitical turmoil. A high EU share of FDI in turn should align EU and WB interests.

Furthermore, from the EU perspective, FDI in WB is self-serving, as one element of a 'de-risking' strategy, put in place by incentivising European firms to shift production closer to home, with the Western Balkan as one region in which geopolitical competition takes place.

As mentioned above, the ERBD and EIB can play an important role in promoting EU FDI in the region and in maintaining the FDI-based ties between the EU and WB, thus sidelining investment from geopolitical rivals. Through these institutions, the EU should develop and enhance the capital market in the region, in particular by stimulating investment by smaller firms in the region (Ghodsi *et al* 2022).

Both EU outbound investment promotion and inbound investment control can play roles here. Outbound EU investment to WB has positive implications (both for the EU and WB countries) beyond market opportunities and should be promoted through available incentivising instruments, while WB inbound investment control becomes increasingly important in light of the state-driven and strategic investment of China and Russia in the region.

The existing EU inbound investment control regime should be treated as relevant acquis that should enjoy priority in implementation in the WB. This would help to identify (and divert) state-driven acquisitions that could ultimately increase WB dependence and vulnerability.

Within the WB bloc, this implies that EU and WB countries must develop regional guidance on screening mechanisms that respond to FDI in line with the EU investment control regime.

2.5 Approximation of laws

Another comparative imbalance between the WB and eastern European countries are their variable commitments on the approximation of laws. While the EU generally makes the adoption of the acquis an ex-ante precondition for access to the internal market, there are significant differences in how this obligation is put in place substantively and in governance structure.

Approximation of laws forms an essential element of the SAAs, which provide for seamless access to the internal market for goods originating from WB countries based on a sufficient alignment of national rules with the Union acquis.

Specifically, the WB SAAs *"recognize the importance of the approximation to that of the Community"* (Article 72 Serbia-SAA) and they provide for a governance structure that aims at promoting the approximation process.

What is missing beyond this general obligation is a more detailed enumeration of specific legal texts to be adopted and by when. Likewise, CEFTA provides a governance structure on *"harmonization of technical regulations and standards"* in the field of TBTs (Article 13 of CEFTA) but remains silent on substantive obligations and concrete legal texts.

This contrasts with the extensive approach on the approximation of laws under the DCFTA agreements, which specify the approximation of laws for individual policy areas (rather than one single encompassing global obligation).

In the DCFTAs, the agreements are much more explicit, with the listing of hundreds of directives and regulations that the Eastern partnership countries are required to implement.

Take public procurement as a specific example. The Georgia DCFTA provides for a gradual approximation of public procurement legislation in Georgia with the Union public procurement acquis based on the specific EU procurement law (Article 141 Georgia DCFTA), and it requires further approximation with the Union's public procurement acquis (Article 146 Georgia DCFTA).

In essence, while the WB SAAs rely on a procedural framework to pursue approximation of law (through cooperation), the DCFTA agreements, in addition to a procedural framework, specify substantively the specific approximation obligation.

Evaluation of the Georgian experience shows that the gradual approximation to EU norms in public procurement improved the already reformed system (Akhvlediani *et al* 2022).

The higher degree of specificity in terms of the obligation to approximate the laws is also a result of a continuous practice of amending the SAAs. The Ukraine SAA has been modified and extended by new or revised Annexes to the SAA around ten times since 2018, while the Serbia AA has been amended in the same time period only once.

One reason for this difference could lie in the more compelling approximation ambition in the EaP SAAs. For example, the Ukraine SAA contains special approximation provisions for the areas of sanitary and phytosanitary and animal welfare legislation, as well as for telecommunications – these specific approximation obligations have been used to amend and further develop the Ukraine SAA. In turn, the Serbia SAA is limited to a general approximation provision but largely lacks more specific obligations.

3 Comparative assessment of governance deficiencies

While integration into the internal market is primarily an issue of substantive requirements on market access, governance is essential in implementing effectively the commitments under the agreements.

The governance structure common to SAAs typically involves an SAA Council as political body, with high-level representatives of both the EU and the country in question, tasked to supervise and evaluate the integration process. A Stabilisation and Association Committee composed of high-level civil servants supports and prepares the work of the SAA Council. Sub-committees involving civil servants meet at technical level throughout the year to discuss and monitor progress on specific subject areas covered by the SAA.

There is also a joint SA Parliamentary Committee, involving members of the national parliament and of the European Parliament, from across the political spectrum. These joint institutional structures manage the process by jointly overseeing the implementation of the SAA.

3.1 Political dialogue and civil society

With the WB as a region characterised by multiple historical and contemporaneous internal political tensions (Domi, 2023), the political dialogue as a reconciliatory and inclusive element for integration of the WB into the EU single market is key when it comes to effective implementation of the agreements.

The EaP countries and the WB have established structures of political dialogue that serve to address political and technical issues impeding implementation and deepening cooperation. Dialogue can take place at different political and technical levels between the EU and the region (Annex 1).

Building on the general governance institutions mentioned above, a number of additional formats subsequent to the initial governance under the SAAs have been initiated. Intra-regional governance is put in place through the Regional Common Council (RCC) Secretariat under the Regional Common Market initiative, in cooperation with the CEFTA Secretariat.

The different institutions perform different functions, either inter-regionally to foster convergence with the EU, or intra-regionally between WB countries.

A core difference and shortcoming of the WB structures, compared to the relationship between the EU and the EaP countries, is the absence of civil-society involvement in the framework of implementing the agreements.

Civil society plays an important role in various ways: civil society is a carrier of expertise feeding into implementation of commitments; civil society is key in identifying and eliminating barriers to trade; it collects relevant information to provide to the bodies engaging in trade facilitation or rules approximation.

Civil society also has an important and disciplining surveillance function over governmental decision-making. Also, civil society is one of the groups affected by democratic backsliding in some of the WB countries, undermining the ability of civil society to monitor government action.

The sufficient integration of civil society into the governance structure of the SAA (and the EU Growth Plan) can thus be likened to the Copenhagen Criteria for EU accession, for which involvement of civil society without political and administrative pressures is indispensable.

In that respect, the Ukraine DCFTA establishes a comprehensive structure for political dialogue involving civil society. The EU and the DCFTA countries are obliged *"to involve civil society in the implementation of the agreement"*, to encourage mutual exchanges of experiences and multiple other forms of connecting civil society among each other, as well as with decision-makers (Articles 443, 444, SAA Ukraine). It even creates policy-specific civil-society exchanges, such as for trade and sustainability issues (Article 299, SAA Ukraine).

By contrast, the relevant agreements involving the WB are silent on the role of civil society. The WB SAAs do not assign a task to civil society, nor has CEFTA integrated civil society into the implementation process, nor does the Working Programme of the Common Regional Market[27] identify civil society as a relevant contributor to the implementation process.

In addition and likewise, the EU does not seem to attach much value either to civil-society involvement. Its draft Growth Plan foresees a role for civil society only at the evaluation stage, and only as one of many stakeholders (Article 25 of draft Growth and Resilience Facility).

The limited role of civil society in implementing the WB SAA is insufficient and forgoes benefits, both from the perspective of relevant expertise as well as a source of legitimacy and acceptance.

Again, Georgia can be referred to as a positive example in this respect. The Georgia SAA established a Civil Society Platform, which enables civil-society organisations from both sides to monitor the implementation process and prepare their recommendations to the relevant authorities.

Specifically, the Georgian National Platform of the Eastern Partnership Civil Society Forum was established in 2015 as a consultative body under the Association Agreement. It brings together up to 200 organisations, among them civil-society organisations, employee organisations, trade unions and associations.

Not only does this platform perform a bottom-up process of providing insight, but it also assures the monitoring of the AA/DCFTA's implementation by producing recommendations to the Association Council and the relevant authorities of both parties (Akhvlediani *et al* 2022).

3.2 The DCFTA Trio format as role model?

There is no shortage of political bodies created under the agreements and involved in the process. Association Agreements, CEFTA, the Common Regional Market Initiative – bodies abound, yet they remain deficient. CEFTA's governance structure lacks the enforcement capacity that other trade agreements with similar scope of ambition have.

CEFTA is designed in intergovernmental fashion, it has not created institutions endowed with competences to make legislative proposals, nor does it exercise adequate supervision over the implementation of the agreement.

While the CEFTA Secretariat is largely limited to providing technical and administrative support to the CEFTA Joint Committee and Bodies, the latter are plagued by the need to decide by consensus and are riddled by political controversies over the representation of Kosovo (RCSPI, 2023).

To some extent, the Common Regional Market initiative sought to create the missing element. The RCC Secretariat created under this framework (including countries such as Turkey and Greece) coordinates and monitors the Action Plan in close cooperation and consultation with CEFTA Secretariat.

While dialogue, reconciliation and cooperation characterise the work of the RCC, its success is limited because of the participation of countries beyond the WB, including the geopolitical rival Turkey, which limits the possibility for this governance framework to focus on the specific concerns of the WB countries in relation to the EU.

Drawing from the experience of the EaP countries, there is a need for a political framework dedicated to the joint WB endeavour for EU accession. The 'new frontrunners' – Ukraine, Georgia and Moldova – motivated but disappointed about the slow accession process, created an Associated Trio format in 2021 to push harder to *"enhance their political association and economic integration with the EU"*, in line with their European aspirations[28].

The Trio format was complementary to the multiple other formats and bodies established under the Eastern Partnership, but it was complementary in a productive way by offering an agenda for the dialogues between the 'Association Trio' and the European Commission, in addition to the DCFTA-related issues, one that deepened cooperation in areas including transport, energy and green economy, even if the Trio has its own shortcomings and the war in Ukraine has hampered the effectiveness of this institution.

Taking the Trio format of the DGFCA countries as role model, it is worth exploring an equivalent body as a complementary element to the multiple existing formats and bodies of the Western Balkan. While WB states maintain their individual agreements with the EU, there is no sufficiently visible format that focuses on the joint WB concerns in pursuing EU accession.

Just as the Trio format of DGFCA countries established ad-hoc trilateral consultations to discuss specific issues in the framework of their integration with the EU, a similar institutionalisation could promote the concerns of the WB beyond the SAAs and the Growth Plan framework.

Such a framework could establish 'Trio' coordinators within the Ministries of Foreign Affairs, and coordinate meetings at expert, senior official and, when appropriate, ministerial levels.

The Open Balkan Initiative (OBI) could be a first step in this direction. Intended to intensify the economic integration between three WB countries (Albania, North Macedonia and Serbia), this initiative could grow further to become a representative body that represents WB interests in relation to the EU.

The initial motivation for the OBI arose from fatigue with the sluggish EU integration process, but it could become a productive forum by accelerating intra-regional economic integration, political cooperation in the areas of infrastructure and transport, and the fight against organised crime and terrorism (Semenov, 2022).

There is the potential that the EU finds a counterpart able to speak with one voice for WB countries. Yet, in its current setup, the OBI is not able to compensate for one of the core deficiencies of the cooperation frameworks under CEFTA and the Common Regional Market, which is the absence of an independent institution tasked with overseeing and implementing agreements, and which ensures consistent implementation across countries and alignment with the EU acquis (RCSPI, 2023).

4 Conclusions

The importance of EU single market membership to WB economic prospects cannot be overstated. This analysis sought to highlight differences between WB SAAs and DCFTAs and lessons to learn from the DCFTA process. It showed that the DCFTAs apply a more lenient approach to intra-regional cooperation.

Also, the DCFTAs subject non-tariff barriers to a more explicit regime than WB SAAs; rules governing trade in services incorporate a stronger non-discrimination standard; and the DCFTAs offer a more rigid and comprehensive

approach to the approximation of laws than the WB countries. It is the latter point in particular that underscores the different integration models underpinning the WB SAAs and the DCFTAs.

The WB SAAs were initially concluded with the prospect of addressing the adoption of the acquis during the subsequent accession negotiations (which then turned out to be delayed), rendering SAAs in some aspects less ambitious.

In turn, conclusion of the DCFTAs with the EaP countries was seen as a substitute for EU accession, which explains the (in parts) greater degree of trade liberalisation in the EaP countries than in the WB, and the more assertive stance of these agreements in particular on approximation issues.

There is no indication that the differences in legal governance have translated into a stronger economic performance in the EaP countries compared to the WB. From a comparative perspective, the analysis suggests that dubbing Ukraine and other EaP countries as the 'new frontrunners' appears premature if not misleading. Rather, they can be dubbed 'quickstarters', reflecting their rapid pace in moving from application status to candidate status and accession negotiations.

The WB remains significantly more integrated in trade with the EU than the EaP countries, while convergence with the EU has been stagnating both for the WB and the EaP. While not underperforming compared to the EaP countries, economic deficiencies in the WB nevertheless exist and should be addressed.

Conditionality attached to both internal market and EU funding should be nuanced; above all, in relation to economic intra-regional integration, it should not impede the necessary investments. The eradication of non-tariff barriers should enjoy priority both inter-regionally with the EU and intra-regionally between WB countries.

The EU's levers for promoting investment in the region should be further enhanced, a demand that is further reinforced by geopolitical concerns about Chinese investments coming without EU-type conditionality attached, and thus creating a tempting alternative for WB countries that have been increasingly disappointed with the slow progress in EU accession.

The question is whether and how the identified shortcomings in the agreements should be addressed. One avenue is to seek amendments of the SAAs and adjust according to the shortcomings identified in this analysis, which implies bargaining with the EU on amending the SAAs on a country-by-country basis. Such a formal amendment approach is likely to undermine the negotiation stage of EU accession (into which five out of six WB states have entered).

Amending the SAAs with a view to aligning them with the DCFTAs would in the WB region be perceived as a (disappointing) substitute for EU accession. An alternative would be to seek an agreement that is complementary to the existing ones, concluded between WB countries (negotiating in unity) on the one side and the EU on the other side.

This approach would be in line with the above exploration of a joint body as a counterparty to the EU. However, the existing and persistent intra-regional political tensions make a sufficiently homogenous stance, as a precondition for crafting a joint agreement, an unlikely prospect.

A third and more pragmatic solution would be to use the existing framework to the greatest extent possible. For example, regulation of trade in services gives leeway to the SAA Council to *"take the measures necessary to progressively"* liberalise the supply of services (Article 59 Serbia SAA). In addition, the SAA Council has sufficiently

wide procedural leeway to widen the scope of interaction with civil society and to create space for civil society in the implementation of the SAAs (Article 120 Serbia SAA).

In turn, the EU is more flexible in unilaterally adjusting its policies on the WB. It could nuance the conditionality embedded in its draft Growth Plan and the draft Growth Facility, and it can extend its tools to foster investment in the regional infrastructure, and thus contribute to stronger convergence by the region. ■

Armin Steinbach is a Non-Resident Fellow at Bruegel and Jean Monnet Professor for EU law and economics at HEC Paris

Endnotes

1. Lisa O'Carroll, 'As Ukraine and others queue to join, is EU ready for enlargement?' The Guardian, 31 August 2023.

2. Services data is missing for Kosovo.

3. Based primarily on European Commission (2023c) and the latest relevant Reports and Conclusions from the European Commission and Council, available for each country; other sources referenced as appropriate.

4. For more details, see 'Treaty on European Union — Joining the EU'.

5. Despite Council agreement to begin negotiations with Albania and North Macedonia in March 2020, the process only began for each country in July 2022.

6. See European Commission news article of 8 December 2023, 'Screening meetings completed as part of screening process with Albania and North Macedonia'.

7. Meaning that it "should be offered official candidate status when it is ready"; see https://neighbourhood-enlargement.ec.europa.eu/enlargement-policy/steps-towards-joining_en.

8. See footnote 6.

9. See point 16 in European Commission (2023c).

10. See point 15 in European Commission (2023c).

11. Based on media reports; see for instance Alexandra Brzozowski, 'EU Commission to start screening process for Ukraine, Moldova after 'surprise' delay', Euractiv, 17 January 2024.

12. See point 14 in European Commission (2023c).

13. See 'Stabilisation and Association Agreement with Serbia'.

14. See Association Agreement between the EU and Ukraine.

15. See Majlinda Bregu, Secretary General of the Regional Cooperation Council, speaking at the 10th Belgrade Security Forum, 22 October 2020.

16. See 'The Western Balkans Common Regional Market – a catalyst for deeper regional economic integration and a stepping stone towards EU Single Market'.

17. See https://www.berlinprocess.de/.
18. Which also includes factors such as the efficiency of the clearing process and the ability to track and trace consignments. For more details see https://lpi.worldbank.org/.
19. Bulgaria, Czechia, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia.
20. As well as political tensions and institutional factors, for example.
21. See European Commission news article of 13 December 2023, 'European Commission announces additional €680 million investment package for the Western Balkans under the Economic and Investment Plan'.
22. Christian Danielsson, Director-General for Neighbourhood and Enlargement Negotiations, speaking on 3 March 2020. See Strategeast, 'EU welcomes Ukraine's progress in implementing the Association Agreement and the Deep and Comprehensive Free Trade Area', 4 March 2020.
23. See EU Projects in Serbia, 'The Trans-Balkan electricity corridor'.
24. See Council of the EU, 'The EU: main investor, donor and trade partner for the Western Balkans'.
25. FDI data is problematic, given the opacity of the ultimate investor behind the FDI in question. To address these concerns, in Annex 3 we build on the work of Damgaard et al (2019), who used firm-level data to estimate the "ultimate investor economy" in FDI data.
26. See 'Foreign direct investments, by country, 2010-2022 (BPM6)'.
27. Available from: https://neighbourhood-enlargement.ec.europa.eu/enlargement-policy/policy-highlights/common-regional-market_en.
28. See Ministry of Foreign Affairs of Ukraine, 'Association Trio: Memorandum of Understanding between the Ministry of Foreign Affairs of Ukraine, Ministry of Foreign Affairs of Georgia and the Ministry of Foreign Affairs and European Integration of the Republic of Moldova', 17 May 2021.
29. Source and notes are consistent for each figure in this section.
30. Eurostat does not provide services data for Kosovo.

References

Akhvlediani, T, S Blockmans, J Bryhn, L Gogoberidze, D Gros, W Hu, I Kustova and F Shamsfakhr (2022) Ex-post evaluation of the implementation of the Deep and Comprehensive Free Trade Area between the EU and its Member States and Georgia, Centre for European Policy Studies.

Blockmans, S (2018) 'Georgia's European Way: What's Next?' Policy Paper, Eastern Partnership Civil Society Forum, 27 July.

Bykova, A, B Jovanović, O Pindyuk and N Vujanović (2022) 'FDI in Central, East and Southeast Europe', Monthly Report, May, wiiw.

Dabrowski, M (2014) 'EU cooperation with non-member neighbouring countries: the principle of variable geometry', CASE Network Reports 119, Center for Social and Economic Research.

Dabrowski, M (2022) 'Towards a New Eastern Enlargement of the EU and Beyond', Intereconomics 57(4): 209-212.

Damgaard, J, T Elkjaer and N Johannesen (2019) 'What Is Real and What Is Not in the Global FDI Network?' Working Paper WP/19/274, International Monetary Fund.

Darvas, Z, M Demertzis, C Martins, N Poitiers, L Léry Moffat, E Ribakova and B McWilliams (2023) EU trade and investment following Russia's illegal invasion of Ukraine, Study for the European Parliament INTA Committee.

GIZ (2022) Report on Non-Tariff Measures in CEFTA, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ).

Domi, TL (ed) (2023) Western Balkans 2023: Assessment of Internal Challenges and External Threats, New Lines Institute.

Egenhofer, C (2023) 'The solution to phasing out coal in the Western Balkans?' CEPS Policy Brief 2023-04 Centre for European Policy Studies.

European Commission (2022a) 'Commission Opinion on the Republic of Moldova's application for membership of the European Union', COM(2022) 406 final.

European Commission (2022b) 'Commission Opinion on Ukraine's application for membership of the European Union', COM(2022) 407 final.

European Commission (2023a) 'New growth plan for the Western Balkans', COM(2023) 691.

European Commission (2023b) 'Proposal for a Regulation of the European Parliament and of the Council on establishing the Reform and Growth Facility for the Western Balkans', COM(2023) 692 final.

European Commission (2023c) '2023 Communication on EU Enlargement Policy', COM(2023) 690 final.
European Commission (2023d) 'Kosovo 2023 Report Accompanying the 2023 Communication on EU Enlargement policy', SWD(2023) 692 final.*
Ghodsi, M, R Grieveson, D Hanzl-Weiss, M Holzner, B Jovanović, S Weiss and Z Zavarská (2022) 'The long way round: Lessons from EU-CEE for improving integration and development in the Western Balkans', Joint Study No. 2022-06, wiiw and BertelsmannStiftung.
Mihajlović, M, S Blockmans, S Subotić and M Emerson (2023) Template 2.0 for Staged Accession to the EU, Open Society Foundations.
RCSPI (2023) A Proposal for a New Approach to the Western Balkans-EU Relationship, Regional Center for Strategic and Political Initiatives.
Semenov, A (2022) 'Open Balkan: Objectives and Justifications', Comillas Journal of International Relations 24.
Sretić, Z (2023) 'Sectoral integration opportunities in the SAA regime', Issue Paper, June, Open Society Foundations.
Stanicek, B (2022) 'China's strategic interests in the Western Balkans', Briefing PE 733.558, European Parliamentary Research Service.
Vulović, M (2023) 'Economic Relations between the Western Balkans and Non-EU Countries', SWP Comment 2023/C 36, Stiftung Wissenschaft und Politik.
World Bank (2015) The Regional Balkans Infrastructure Study (REBIS) Update, Enhancing Regional Connectivity, Identifying Impediments and Priority Remedies.

This article is based on a Bruegel Working Paper that was produced with financial support from The Open Society Foundations Western Balkans. The author thanks Conor McCaffrey for research assistance, along with Marek Dabrowski, Zsolt Darvas, Heather Grabbe and André Sapir for their valuable comments and suggestions.

Annex 1. Legal comparisonAnnex 2. Trade data

Country (Date of entry into force)	Association Agreement/Deep and Comprehensive Free Trade Area Agreements Ukraine (Association Agreement since 2014, DCFTA since 2016)	Associate Agreement Serbia (2013)	Observed differences between DCFTA and SAA
Regional Cooperation Requirements (ie. necessity to integrate primarily regionally)	Chapter 27 - promote regional understanding; support and strengthen involvement of local and regional-level authorities in crossborder and regional cooperation; strive to develop crossborder and regional elements in various areas; regular dialogue on this matter.	Title III, Art 14: "Serbia shall actively promote regional cooperation. The Community assistance programmes may support projects having a regional or crossborder dimension through its technical assistance programmes.… implement fully the CEFTA"; Art 15: "Serbia shall start negotiations with the countries which have already signed an SAA with a view to concluding bilateral conventions on regional cooperation", main elements: political dialogue, free trade areas, various economic freedoms and cooperation in areas such as justice, freedom and security. "Readiness by Serbia to conclude such conventions will be a condition for the further development of the relations between Serbia and the EU"; Art 16: Pursue regional cooperation with the other States concerned by the SA process;	The language seems stronger for SAAs-matches what Windisch said in his intervention "no access to the single market on any of the 7 pillars will be granted before there is integration on the common regional market."

		Art 17: "Foster its cooperation and conclude a convention on regional cooperation with any country candidate for EU accession in any of the fields of cooperation covered by this Agreement… should aim to gradually align bilateral relations… with the relevant part of the relations between the Community… and that country". Should also start negotiations with Turkey on establishing a free trade area.	
Political dialogue structure (institutional exchange, high level, lower level etc.)	Arts 460-468: Highest level is Summit level, to take place in principle once a year; political and policy dialogue at ministerial to take place at least once a year within the newly established Association Council; Parliamentary Association Committee established. Article 5: As well as the above, there will be regular dialogue at Foreign Minister, Political Directors, Political and Security Committee and expert levels.	Title II, Art 10-13: Political dialogue to be further developed between the parties to support the rapprochement between the EU and Serbia and increase convergence on international issues and security and stability; in addition to the institutions described below dialogue can occur directly between officials representing the Council Presidency or HRVP and those representing Serbia Art 119-125: Stabilisation and Association Council, made up of members of the European Council and Commission and the Government of Serbia, is established and shall meet at regular intervals and when required; the Council is to be supported by an SA Committee; Stabilisation and Association Parliamentary Committee established, consisting of members of the European Parliament and the Parliament of Serbia, to allow them to meet and exchange views.	Slight differences: DCFTAs seem to mandate ministerial meetings, whereas SAAs talk about senior officials.

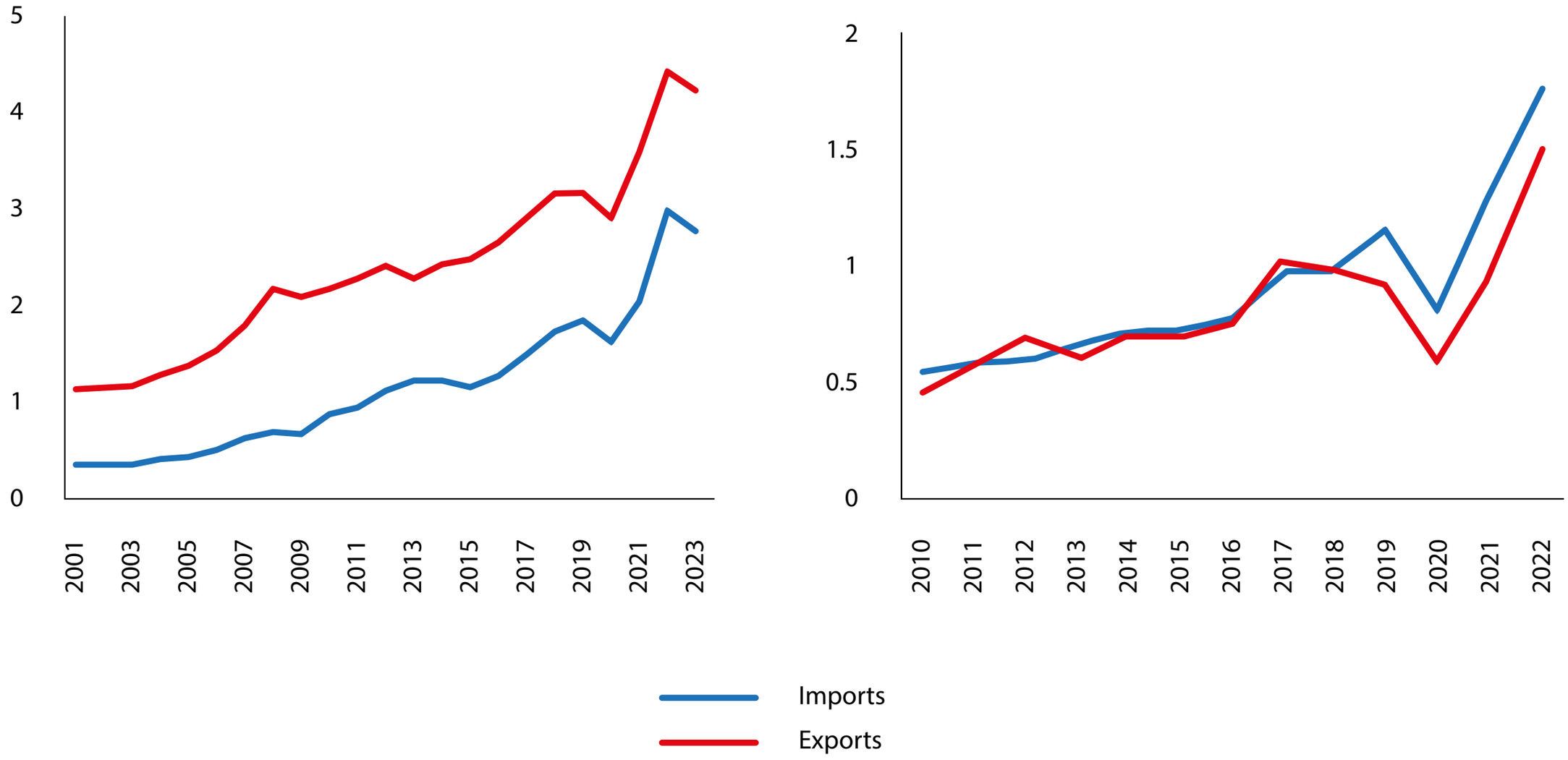
Political dialogue: involvement of civil society	Arts 443 and 444: Promoting dialogue and cooperation between civil society groups in both regions. Arts 469 and 470: Parties will promote regular meetings as representatives of their civil societies; Civil Society Platform established to allow for an exchange of views and to meet with and make recommendations to the Association Council.	No	More of a reference to civil society in the DCFTAs.
Freedom/liberalization of trade in goods	Art 29: Sets out schedule for reduction/elimination of custom duties.	Title IV, Art 18: "shall gradually establish a bilateral free trade area over a period lasting a maximum of six years"; controversial legal interpretation, see Sretic (2023), pg 6-7.	Not significant.
Trade in services	Art 94: In the sectors where market access commitments are inscribed in Annexes… each Party shall grant to services and service suppliers of the other Party… treatment no less favourable than that it accords to its own like service and services suppliers.	Art 59: Liberalisation process-parties undertake to take the necessary steps to allow progressively the supply of services by firms/nationals of the other party, with a review after four years; temporary movement of key personnel allowed to support this; Art 60: "The Parties shall not take any measures or actions which render the conditions for the supply of services by Community and Serbia nationals or companies which are established in a Party other than that of the person for whom the services are intended significantly more restrictive as compared to the situation existing on the day preceding the day of entry into force of this Agreement."	Different form of no discrimination (time vs nationality).

		Art 61: Provisions on transport services specifically.	
Freedom of workers	Art 97-102: Limited freedom of movement for certain classes of workers.	Art 49: Non-discrimination rules. Art 50: Bilateral agreements on access to employment for Serbians should be preserved, improved and possibly expanded to other member states. Art 51: Rules shall be laid down for the coordination of social security systems for Serbian workers, legally employed in the territory of a member state and vice versa.	Not significant.
Freedom of establishment	Art 88: Treatment no less favourable than that accorded to its own legal persons… or to any third-country legal person… whichever is the better;	Art 53: "no less favourable than that accorded to its own companies or to any third country company, whichever is the better."	Not significant.
Freedom of capital	Art 145: Shall "ensure the free movement of capital relating to direct investments made in accordance with the laws of the host country, to investments … and to the liquidation or repatriation of such invested capitals and of any profit stemming therefro". Portfolio investments, financial loans, credits related to commercial transactions also covered. "Ukraine undertakes to complete the liberalisation of transactions on the capital and financial account of balance	Art 63: "With regard to transactions on the capital and financial account of balance of payments, from the entry into force of this Agreement, the Parties shall ensure the free movement of capital relating to direct investments made in companies formed in accordance with the laws of the host country and investments made in accordance with the provisions of Chapter II of Title V, and the liquidation or repatriation of these investments and of any profit stemming there from."	Not significant.

	of payments equivalent to the liberalisation in the EU Party prior to the granting of internal market treatment in the area of financial services… A positive assessment of the Ukrainian legislation on capital movements, its implementation and continued enforcement… is a necessary precondition of any decision by the Trade Committee to grant internal market treatment with respect to financial services." Discussions to take place 5 years after the entry into force to see what still needs to be done.	Free movement of capital relating to credits related to commercial transactions/provision of services, portfolio investment and financial loans and credits are also covered. Serbia should authorise and liberalise the purchase of its real estate by EU nationals so that they ultimately receive the same treatment as Serbians. After four years the SA Council will determine what remains to be done to apply full EU rules on freedom of capital.	
Provisions on non-tariff barriers	Art 34-35: Each Party shall accord national treatment to the goods of the other Party in accordance with Article III of GATT 1994, including its interpretative notes… No Party shall adopt or maintain any prohibition or restriction or any measure having an equivalent effect on the import of any good of the other Party or on the export or sale for export of any good; Art 53-58: Reference cooperation and previous agreement on technical barriers to trade.	Title IV: No explicit mention in trade in goods (though legally controversial, Sretic 2023).	No explicit mention of non-tariff barriers in the SAAs, but the Sretic piece seems to argue they are implicit?

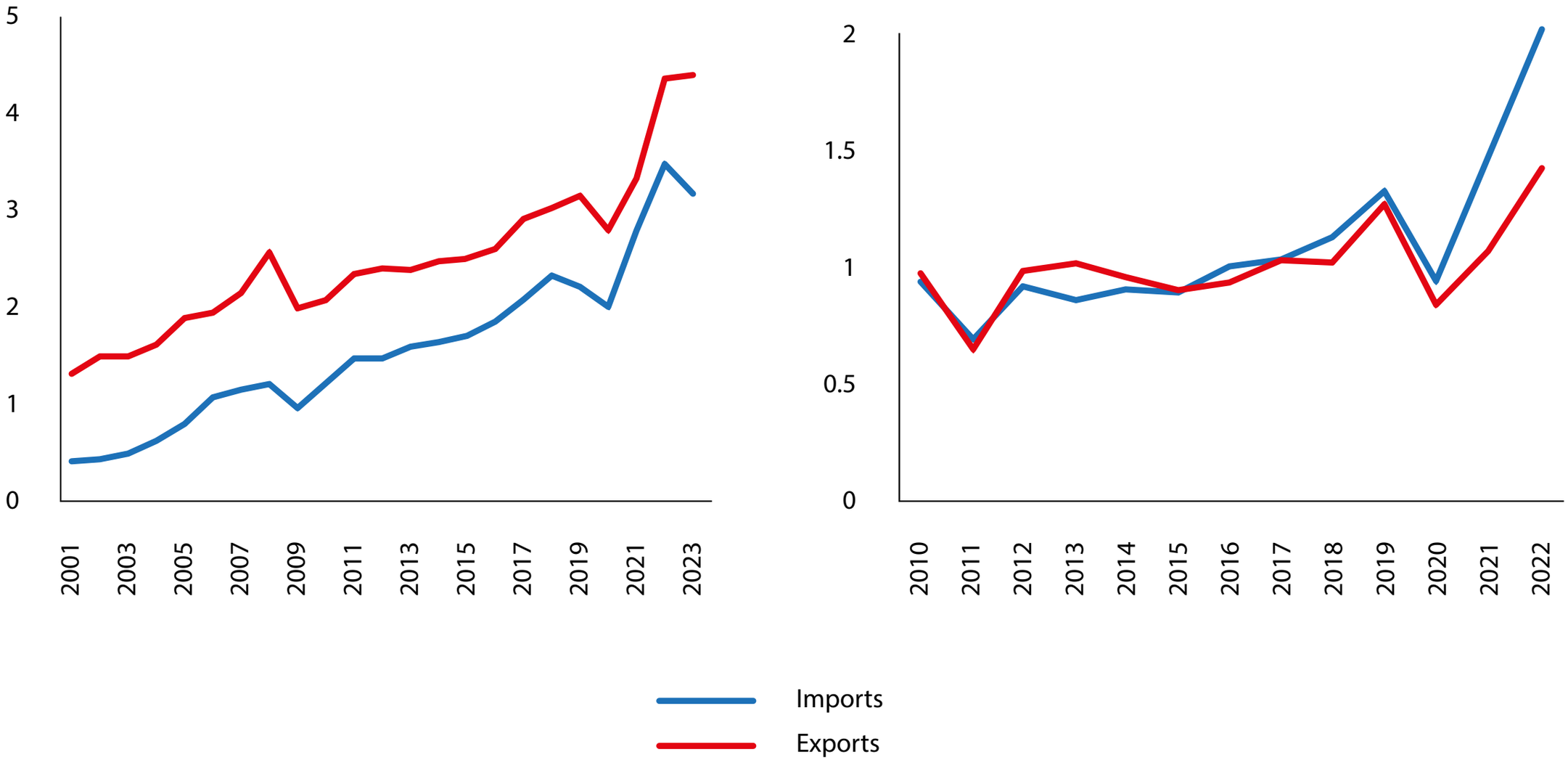
Annex 2. Trade data

EU27 goods (left) and services (right) trade with Albania (€ billions)

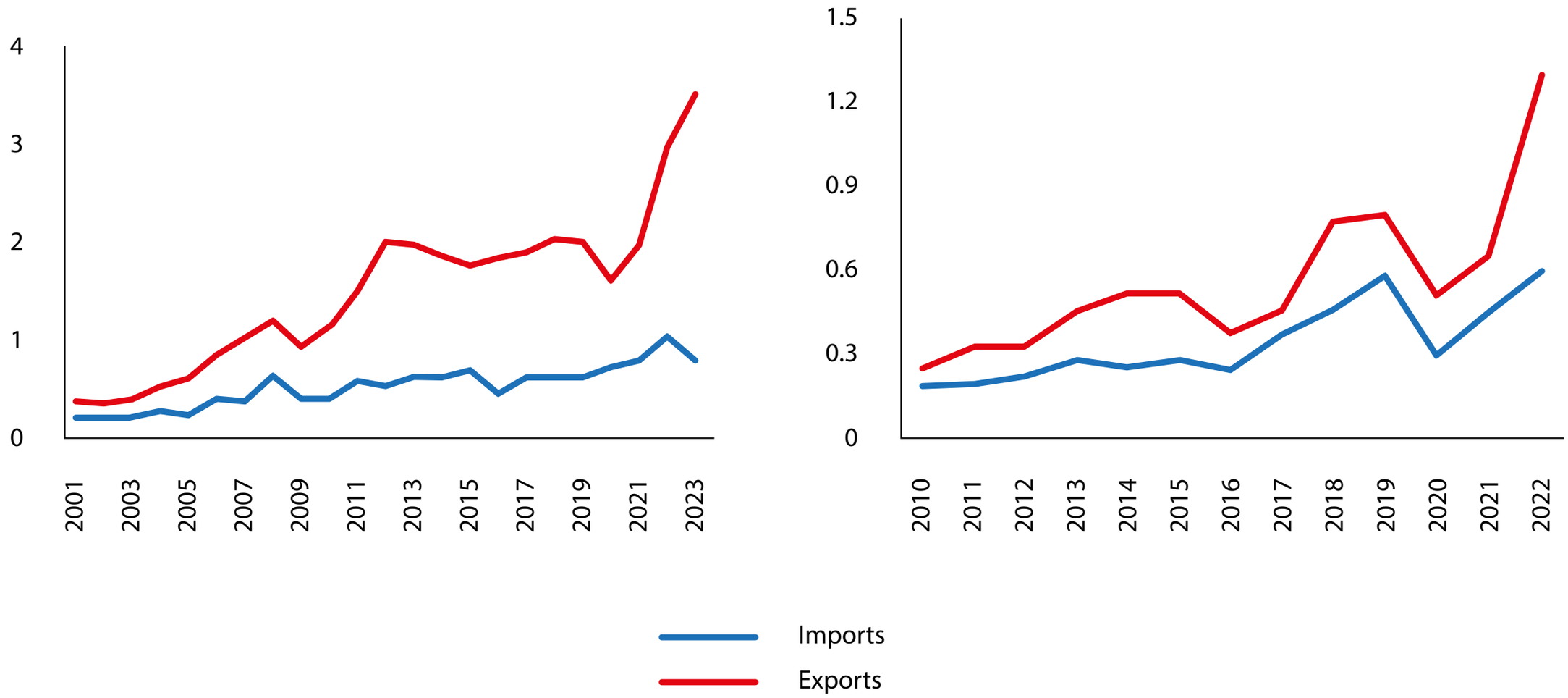


Note: Exports refer to EU exports to Albania and imports the reverse[29].
Source: Bruegel based on Eurostat (DS-018995).

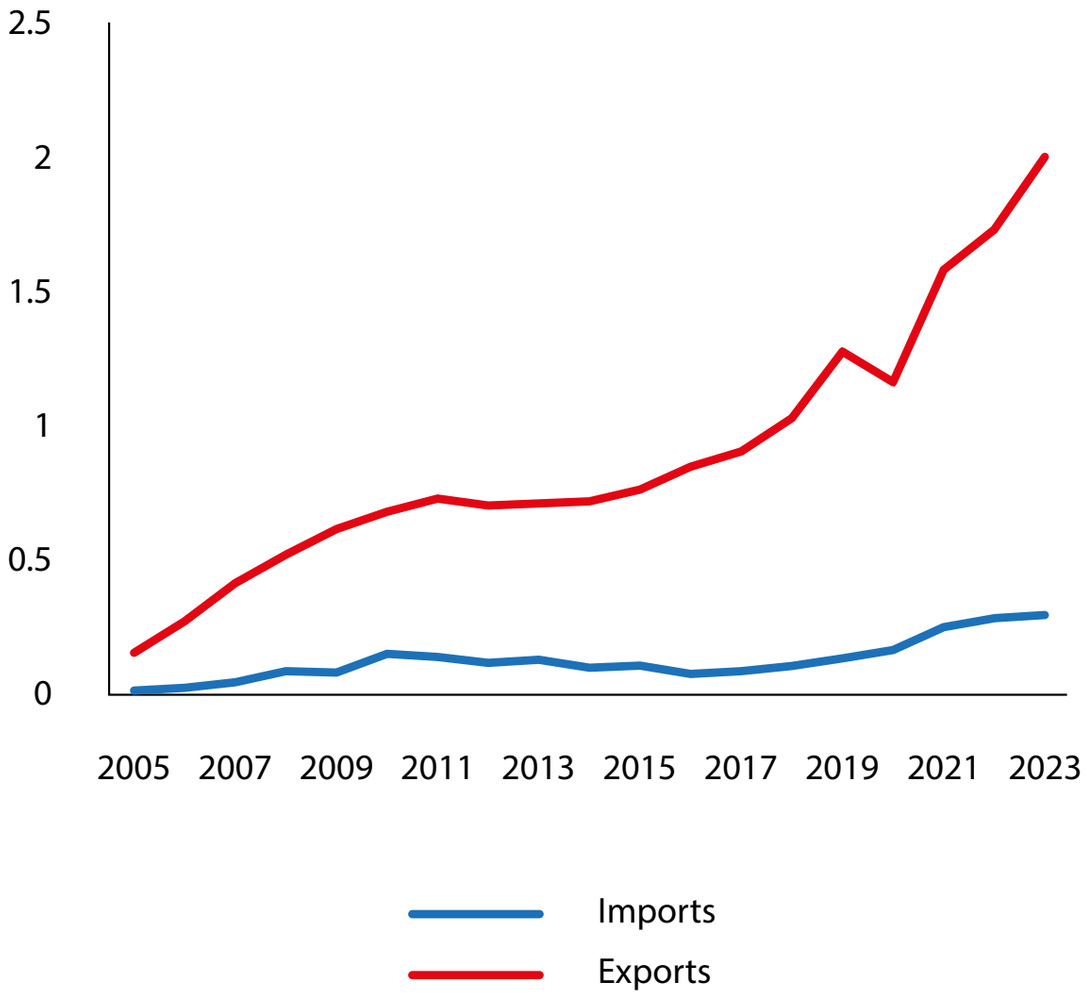
EU27 goods (left) and services (right) trade with Bosnia and Herzegovina (€ billions)



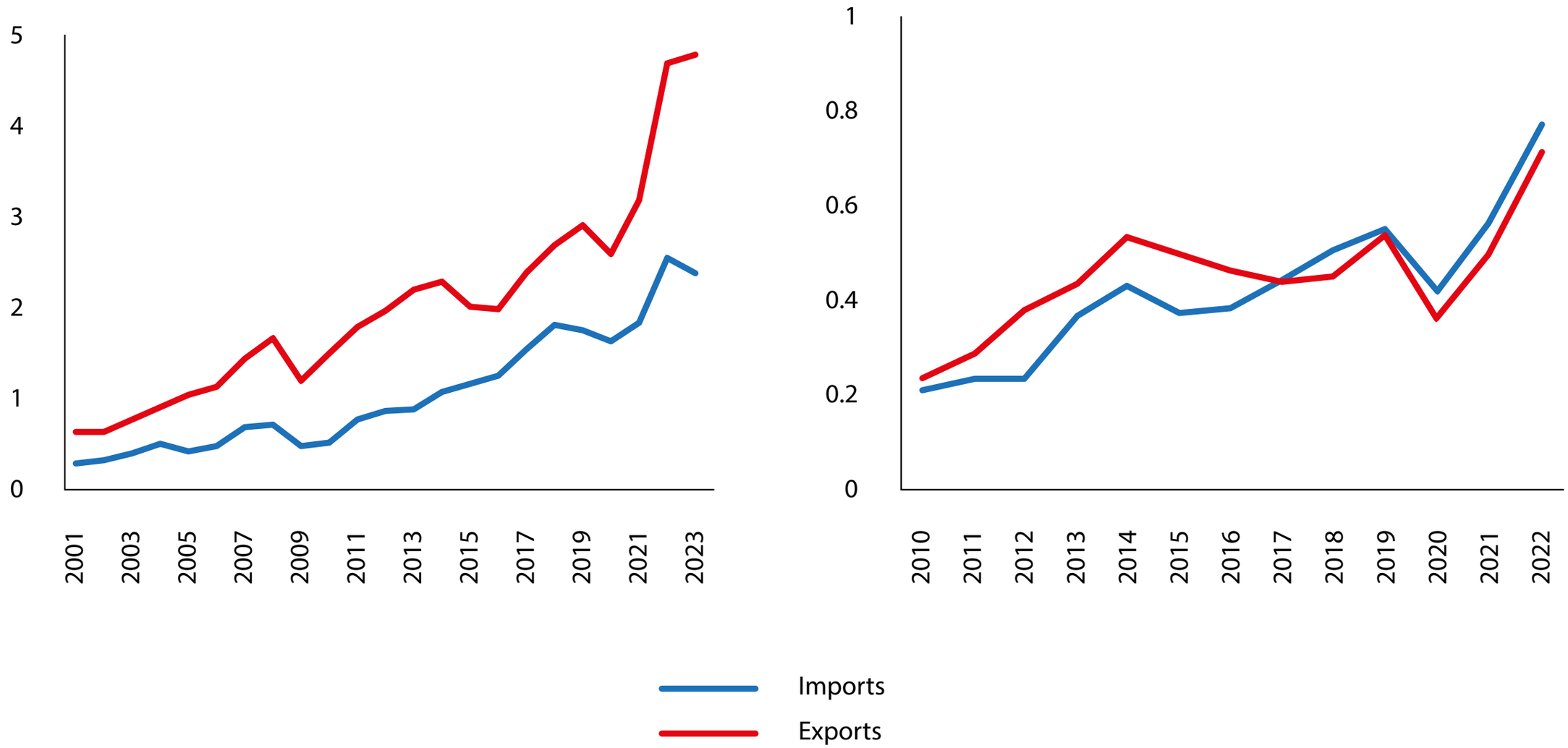
EU27 goods (left) and services (right) trade with Georgia (€ billions)



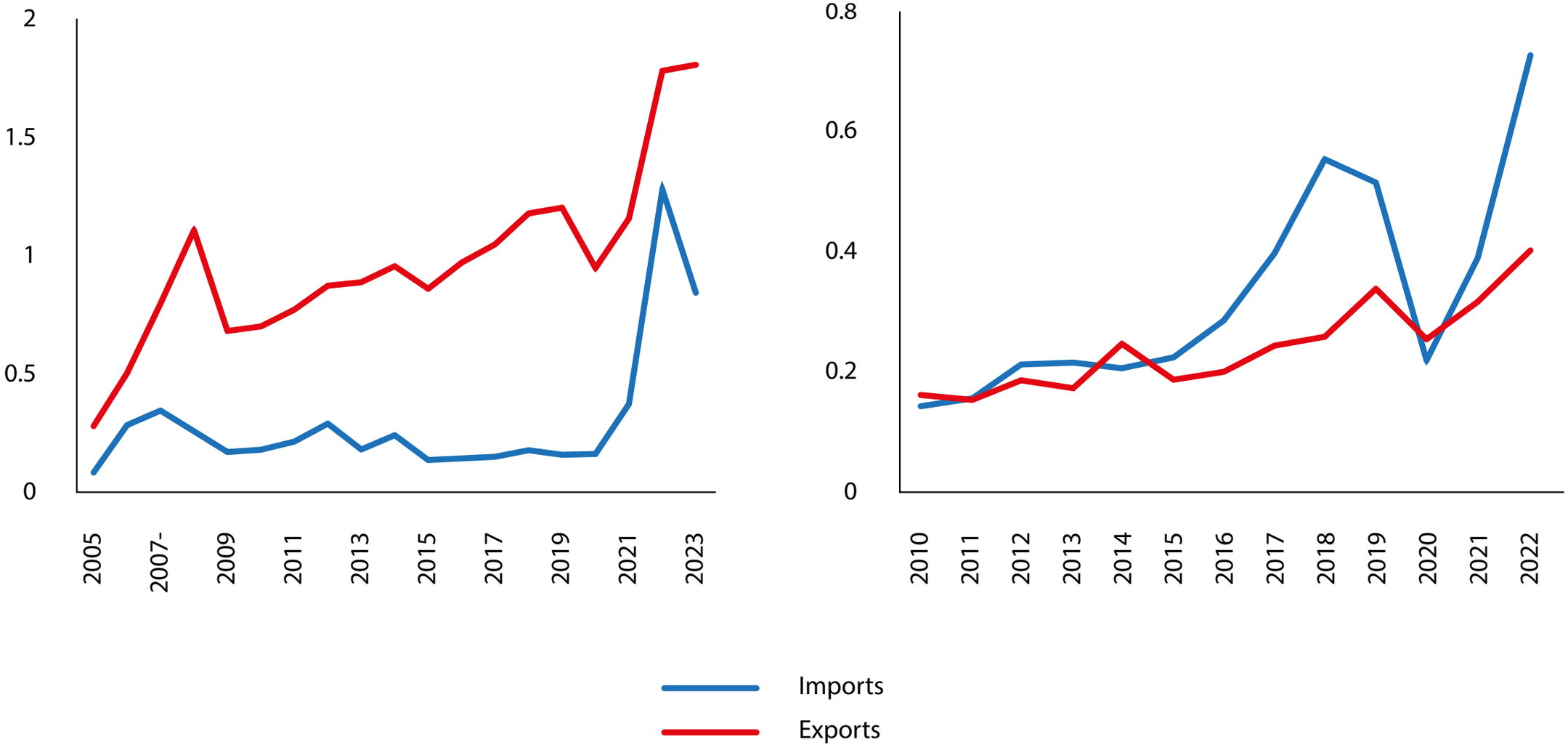
EU27 goods[30] trade with Kosovo (€ billions)



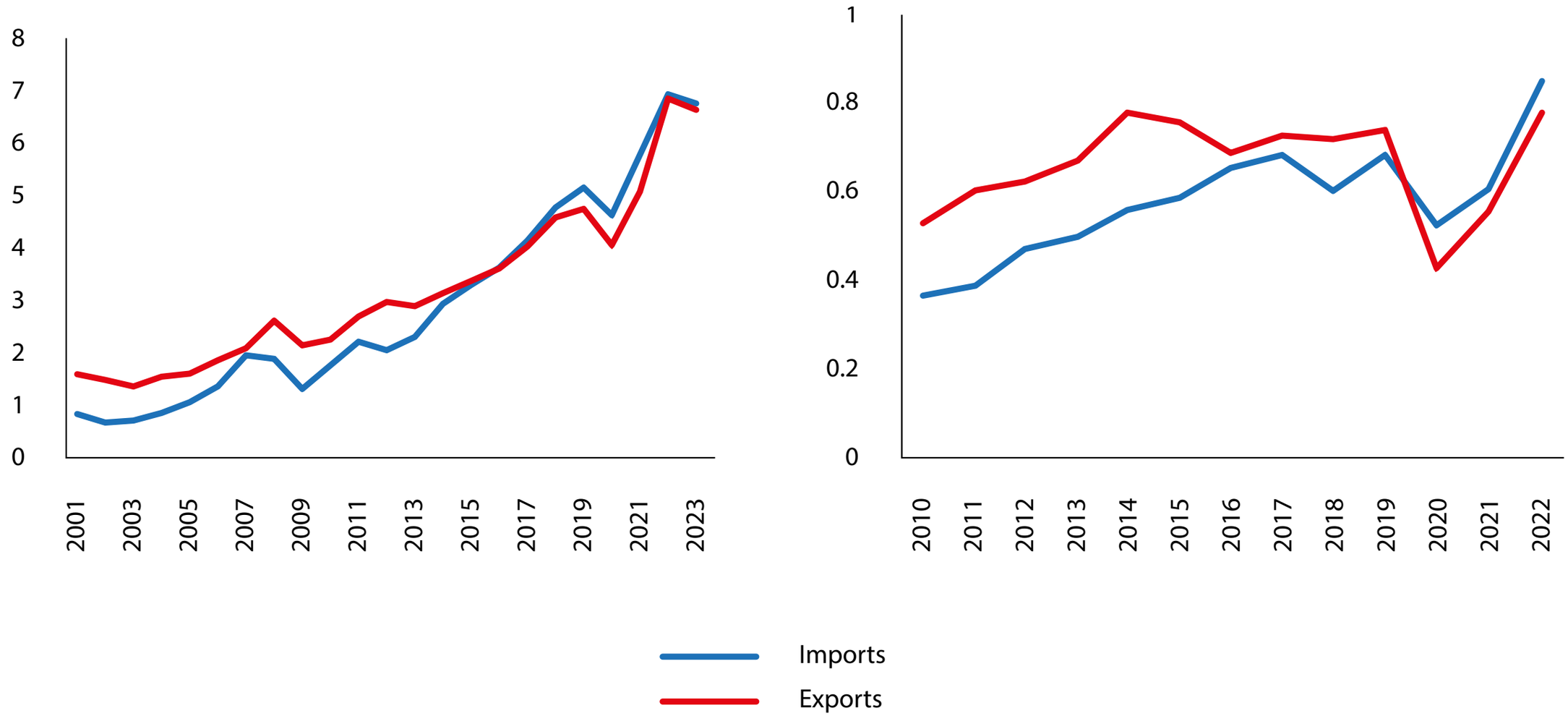
EU27 goods (left) and services (right) trade with Moldova (€ billions)



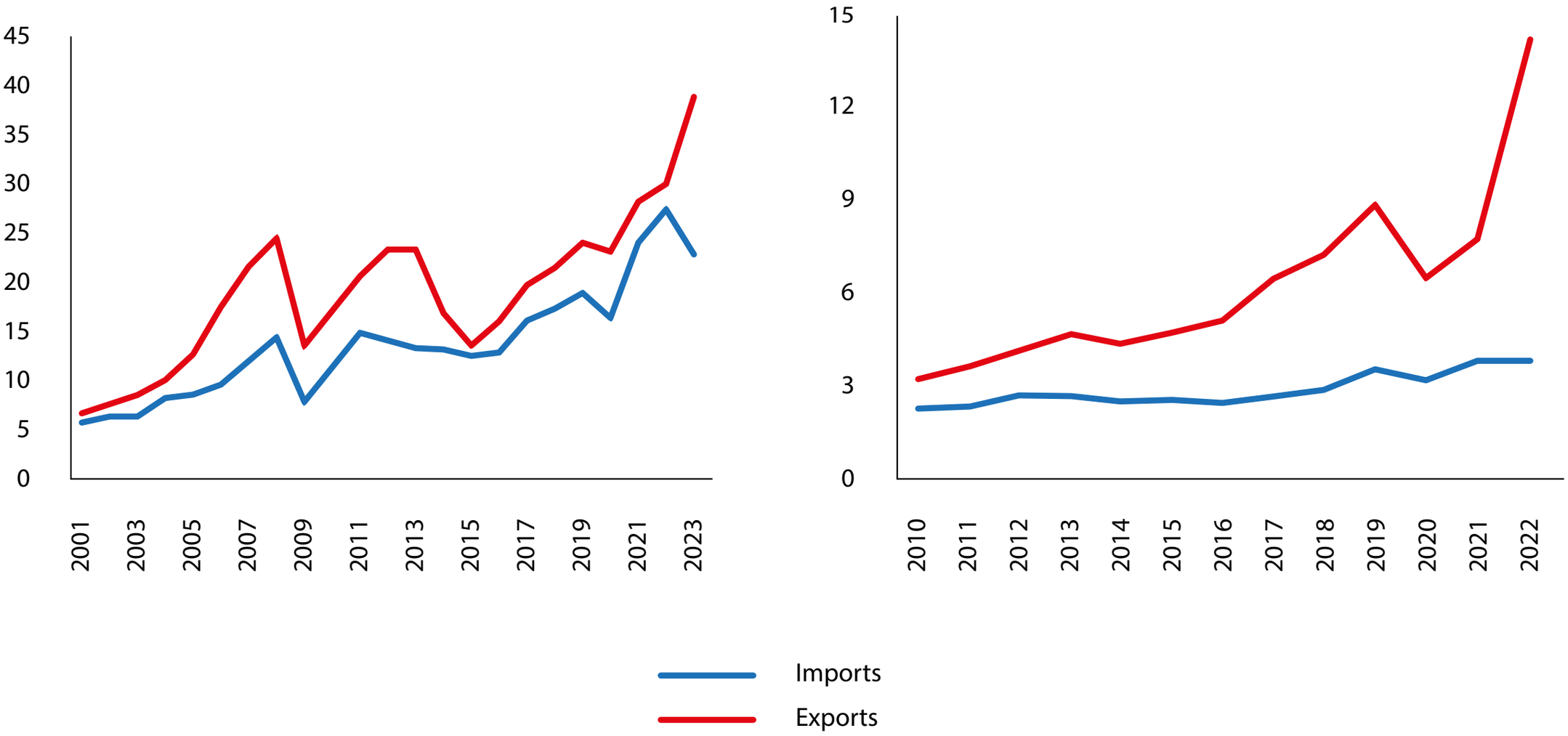
EU27 goods (left) and services (right) trade with Montenegro (€ billions)



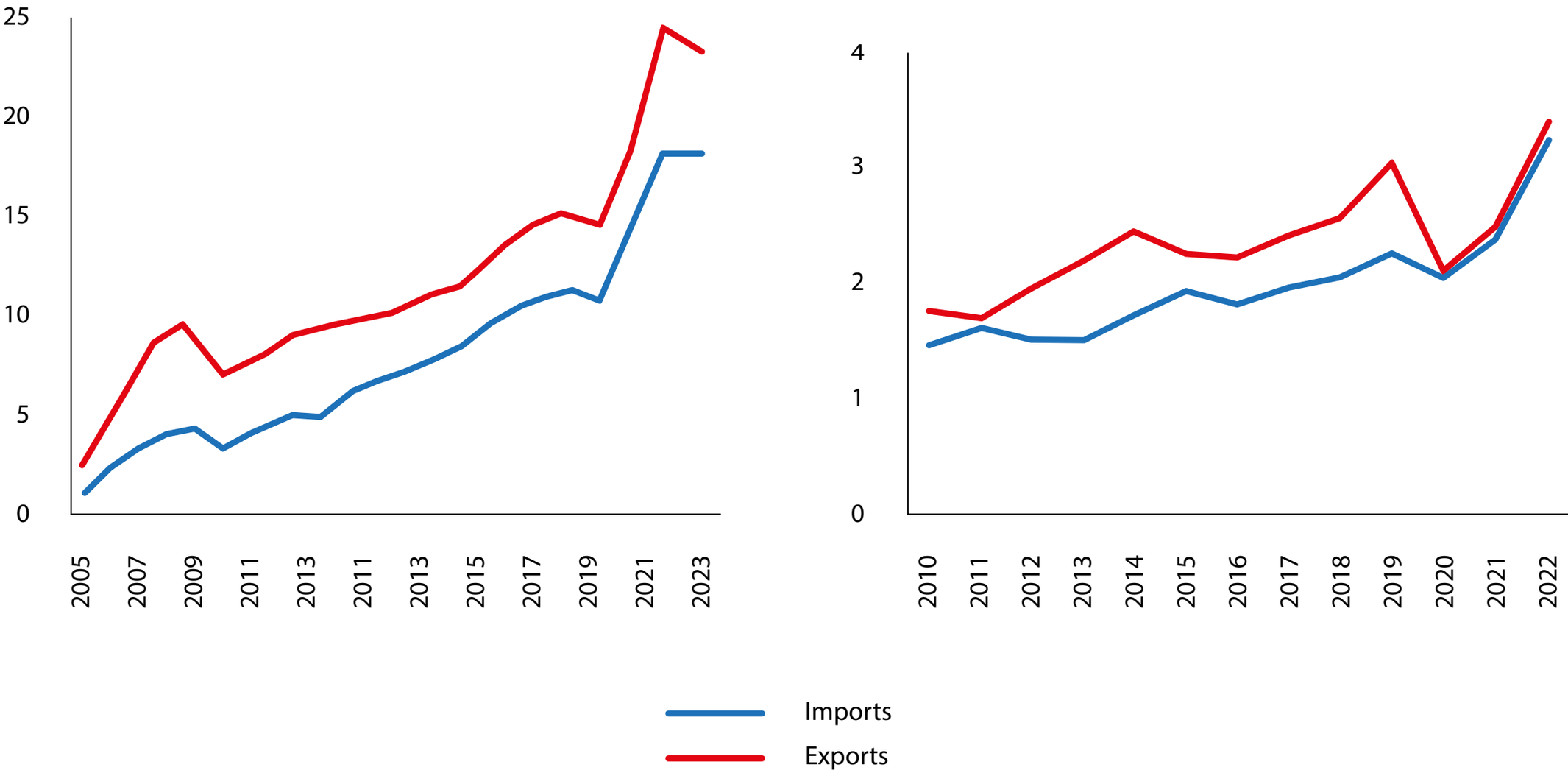
EU27 goods (left) and services (right) trade with North Macedonia (€ billions)



EU27 goods (left) and services (right) trade with Ukraine (€ billions)



EU27 goods (left) and services (right) trade with Serbia (€ billions)



Annex 3. FDI data

Figure 3.1. EU FDI stock as a share of national GDP, Western Balkans

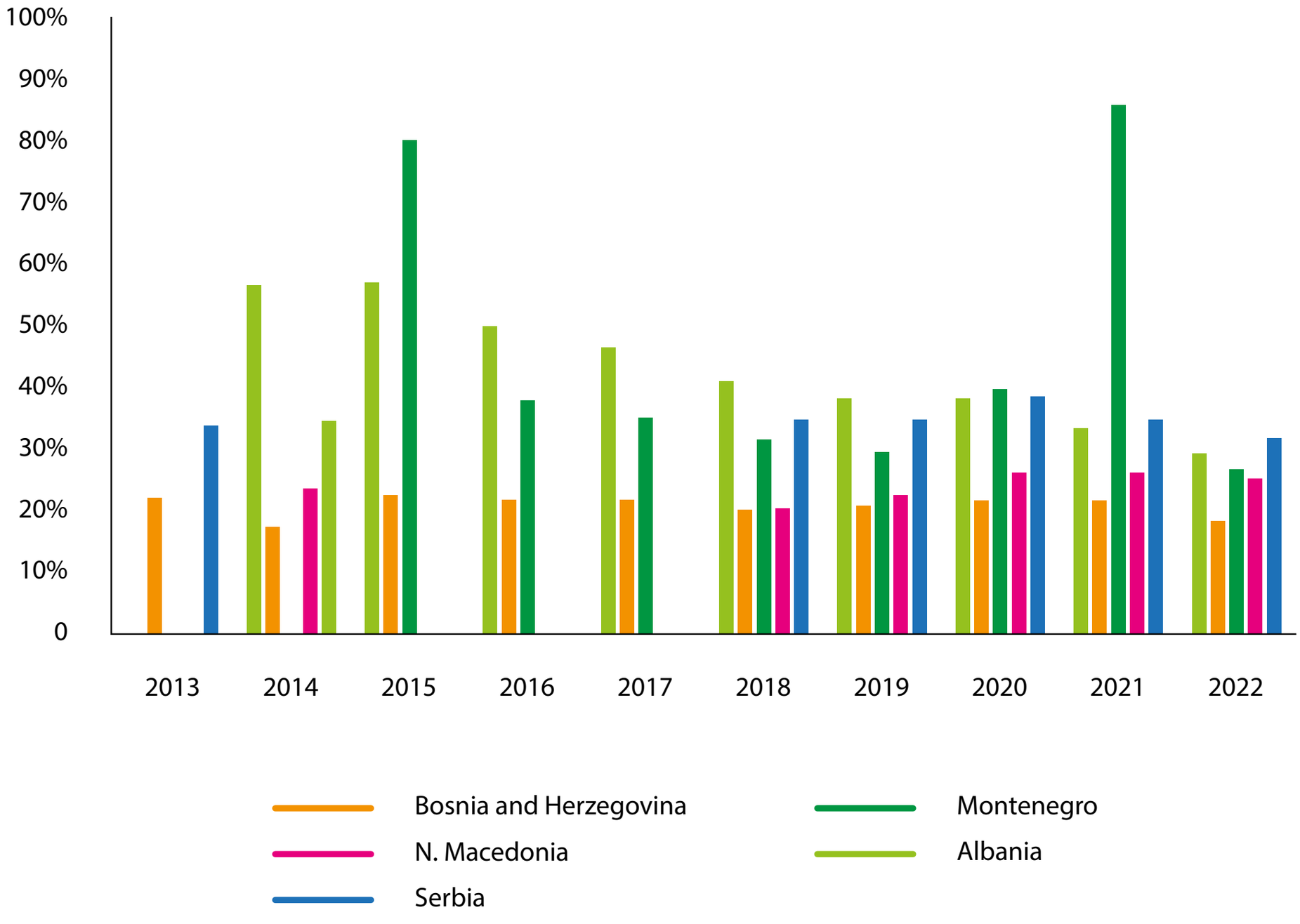
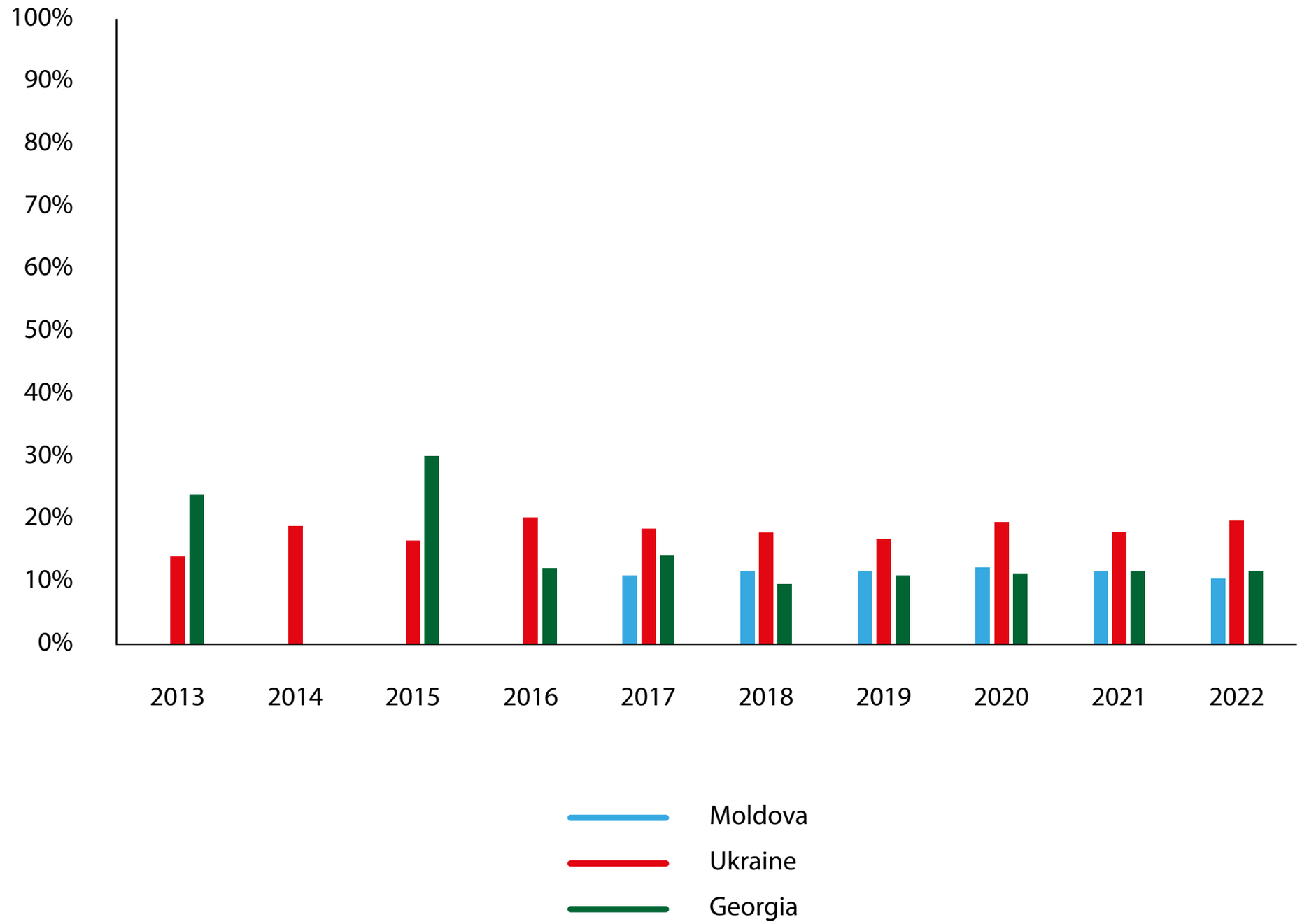


Figure 3.2. EU FDI stock as a share of national GDP, EaP



Source: Bruegel based on Eurostat, World Bank and OECD.

Reporting of FDI data must acknowledge that FDI statistics often mask the true origin of the investment in question, a phenomenon that is exacerbated in the case of the EU given the prominence of certain member states in global tax avoidance (Darvas *et al* 2023). Damgaard *et al* (2019) built a dataset for 2013-2017 that estimated FDI by what they term the *"ultimate investor economy"* (UIE). Over this period, the simple average for the WBs of FDI with the EU as UIE was 45 percent, higher than that of the EaP countries, but lower than the level of trade integration at the same time (the simple average for the EU as a share of total exports for the same period was 59 percent, Table 3.1). An average of 74 percent of the FDI reported as being from the EU across the WB countries actually had the EU as UIE, ranging from 90 percent in North Macedonia to just 50 percent in Montenegro (Table 3.2)

Table 3.1. Share of FDI with the EU as the ultimate investor economy in total reported FDI stock into the WB and EAP countries

Country	2013	2014	2015	2016	2017	2013-2017
Western Balkans						
Albania	43.7%	63.7%	65.0%	53.3%	45.4%	53.5%
B + H	49.7%	50.8%	54.0%	46.6%	47.3%	49.6%
Kosovo	19.0%	21.3%	20.8%	20.3%	20.7%	20.4%
Montenegro	30.6%	32.4%	21.3%	21.9%	21.9%	25.7%
North Macedonia	70.0%	73.0%	68.7%	60.0%	57.3%	65.6%
Serbia	58.8%	57.4%	51.2%	50.9%	45.0%	52.3%
EaP						
Georgia	14.9%	13.3%	10.9%	8.4%	7.5%	10.7%
Moldova	49.0%	48.3%	47.6%	50.6%	52.4%	49.6%
Ukraine	40.6%	34.8%	41.2%	35.1%	40.4%	38.7%

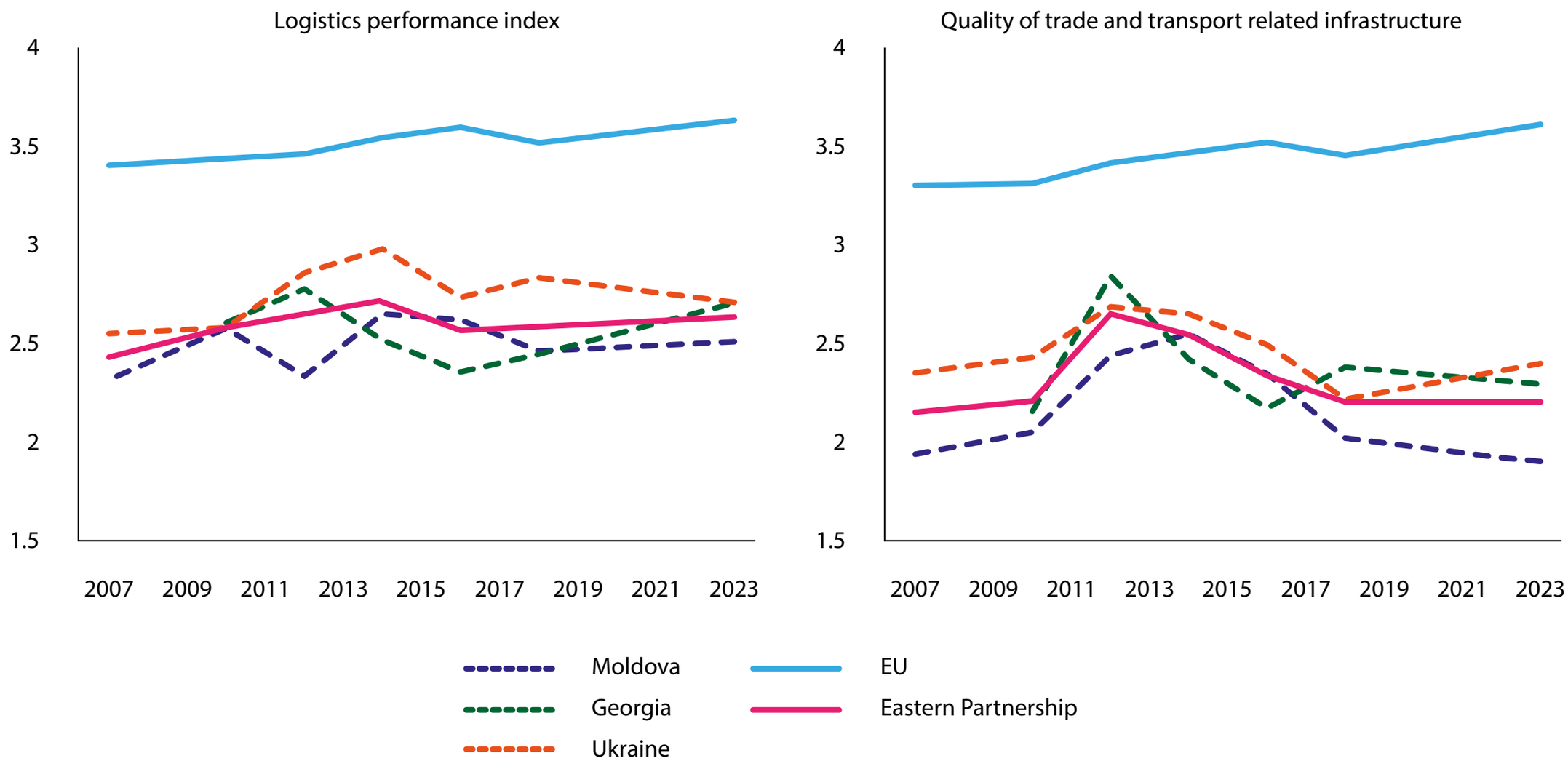
Source: Bruegel based on Damgaard et al (2019) and Darvas et al (2023).

Table 3.2. FDI stock with the EU as the ultimate investor economy as a share of the reported EU FDI stock in each country

Country	2013	2014	2015	2016	2017	2013-2017
Western Balkans						
Albania	91.3%	91.9%	93.0%	90.1%	81.3%	88.9%
B + H	84.7%	88.3%	92.9%	77.5%	76.8%	83.8%
Kosovo	65.6%	69.8%	67.5%	69.9%	66.3%	67.7%
Montenegro	51.1%	54.5%	48.6%	52.0%	46.8%	50.8%
North Macedonia	85.9%	93.5%	90.8%	90.8%	88.1%	89.7%
Serbia	70.5%	70.4%	64.6%	64.8%	58.5%	65.6%
EaP						
Georgia	49.1%	41.6%	35.0%	27.8%	26.6%	35.2%
Moldova	82.0%	81.5%	81.5%	83.4%	82.3%	82.1%
Ukraine	55.2%	49.1%	58.4%	51.8%	60.8%	54.7%

Source: Bruegel based on Damgaard et al (2019) and Darvas et al (2023).

Annex 4. Logistics and trade-related infrastructure for EaP



Source: Bruegel based on World Bank Logistical Performance Index.

Annex 5. Non-tariff barriers

Figure 5.1. Difference in compliance costs of international trade between the Western Balkans and OECD high income countries (left) and the EU (right), $

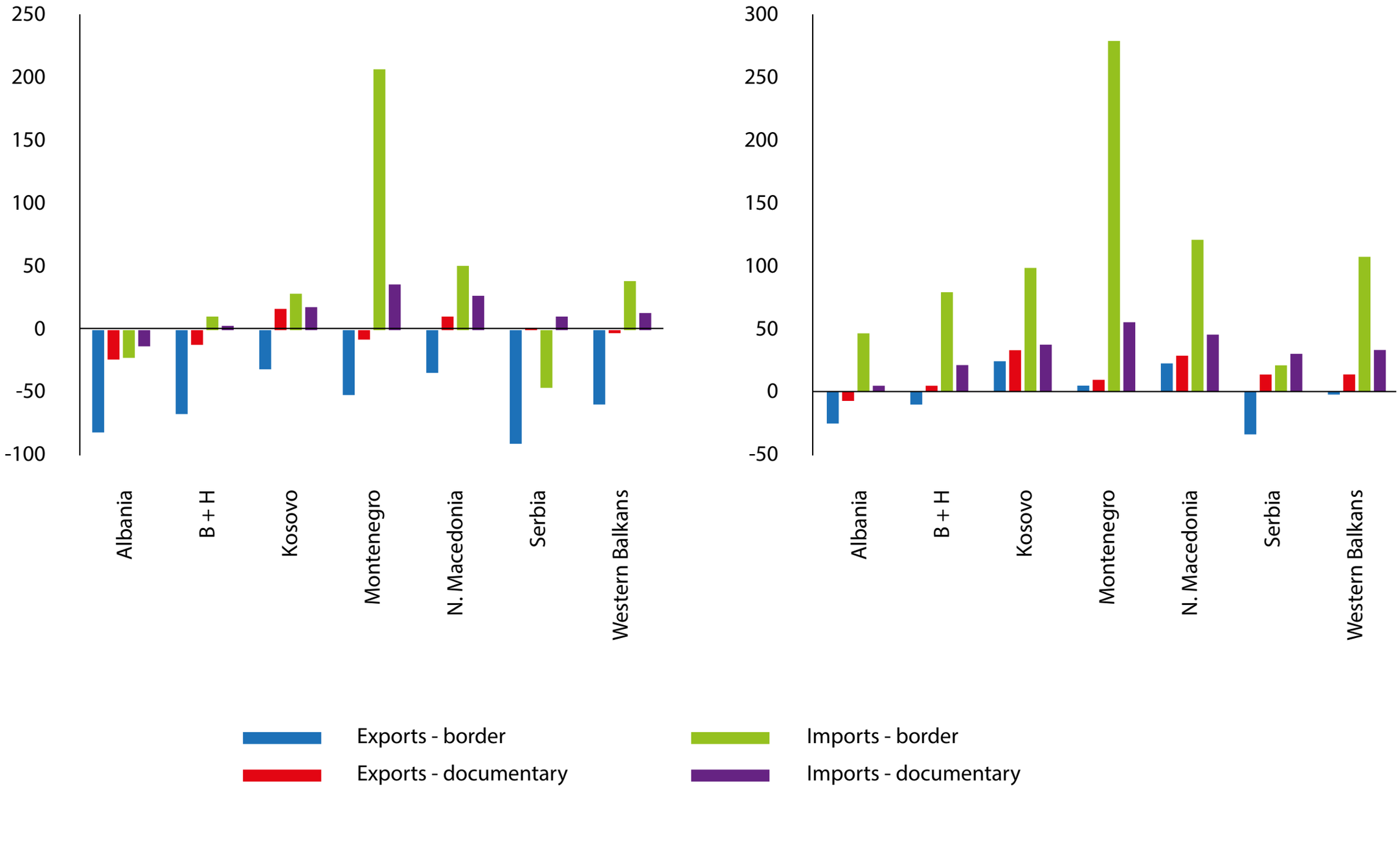
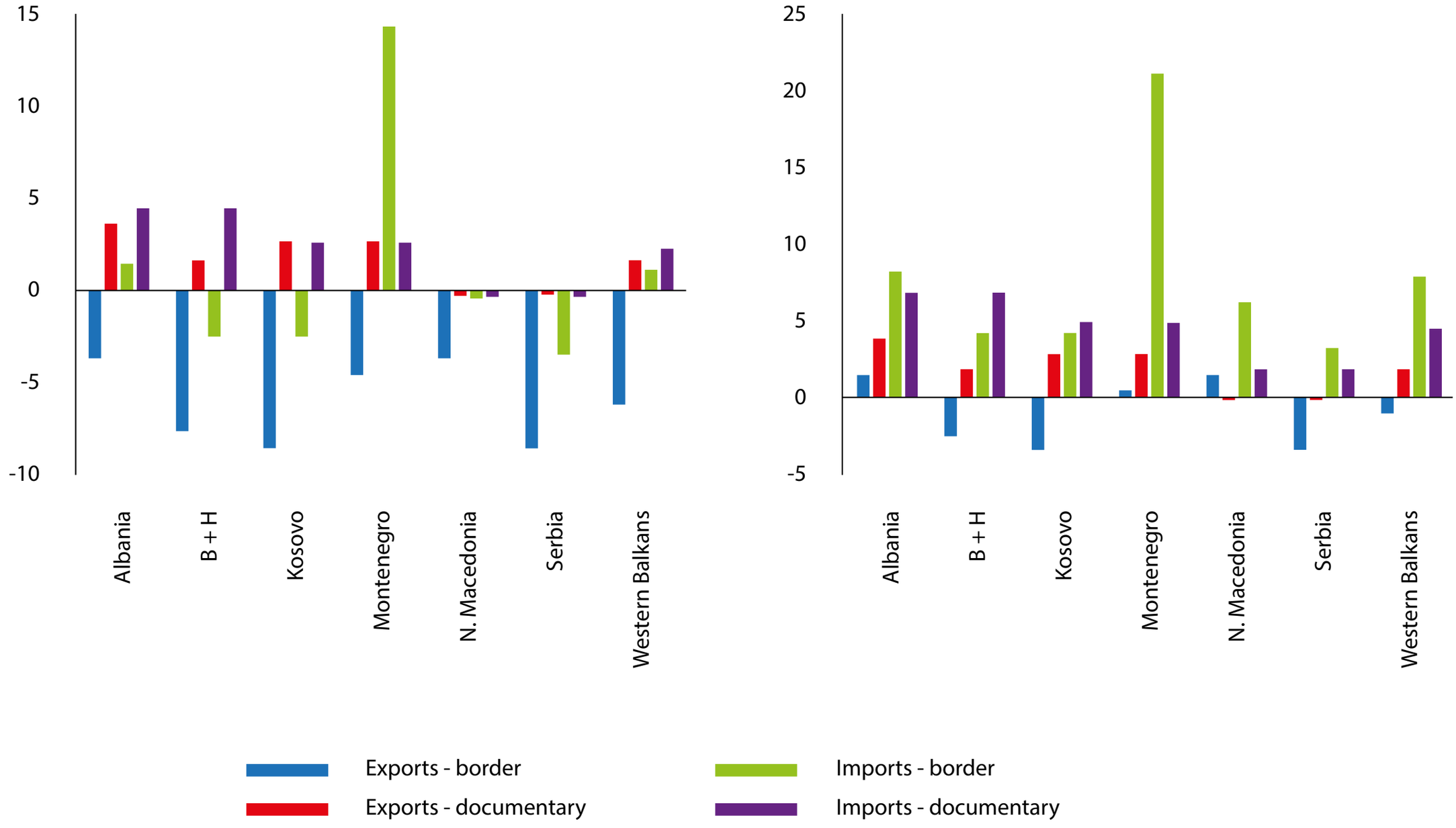
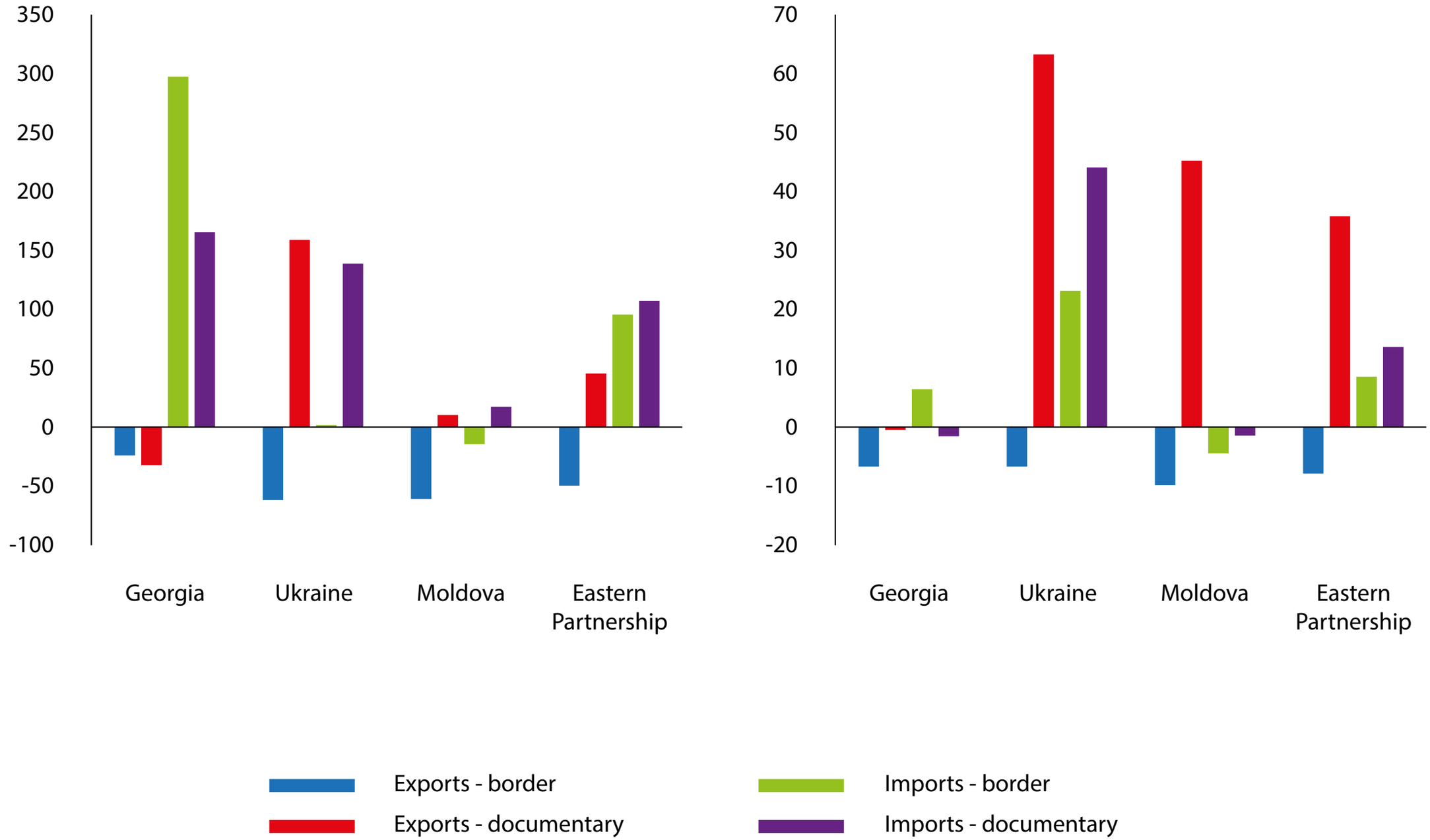


Figure 5.2. Difference in time compliance of international trade between the Western Balkans and OECD high-income countries (left) and the EU (right), hours



Note: WBs refers to a simple average of the six WB countries. EU refers to the simple average of the EU27 countries.
Source: The World Bank 'Trading across Borders'.

Figure 5.3. Difference in compliance costs of international trade between the EaP and OECD high-income countries in $ (left) and hours (right)



Note: WBs refers to a simple average of the six WB countries. EU refers to the simple average of the EU27 countries.
Source: The World Bank 'Trading across Borders'.

Instruments of economic security

Geopolitical and economic developments have raised concerns about the EU's exposure to hostile countries. Conor McCaffrey and Niclas Poitiers assess the nature of this threat and outline lessons that can be drawn

1 Introduction

Recent years have seen rising concern over the 'weaponisation of interdependence', ie. the exploitation of economic links for geopolitical purposes (Farrell and Newman, 2019). There has been a significant shift in the prevailing narrative on both sides of the Atlantic, from seeing economic interdependence as leverage to achieve political liberalisation, to a geopolitical view that sees it as a liability that exposes Western economies to foreign influence[1]. The relationship between the United States and China has soured and China's accession to the World Trade Organisation may now be seen as a mistake[2].

Meanwhile, Russia's invasion of Ukraine is portrayed as glaring example of a failed Western strategy of *Wandel durch Handel* (change through trade). Rather than reducing tensions, economic interdependence instead left some parts of Europe significantly dependent on Russia at the time of the invasion, arguably strengthening Russia's hand.

However, a strategy of economic decoupling, undoing decades of globalisation and therefore vastly reducing the gains from trade, seems neither feasible nor desirable (Aiyar *et al* 2023). There is a new consensus among the G7 countries that the 'de-risking' of economic relationships with revisionist countries is a more feasible strategy[3].

The central idea is to diversify supply chains and build a 'high fence' around a 'small yard'[4], to protect vital economic sectors from foreign interference without jeopardising the economic benefits of globalisation. Put simply, the aim of this strategy is to reduce risks without starting all-out trade wars and undermining the rules-based economic order.

Many of the solutions put forward as part of this strategy include significant government intervention. While additional state support in certain areas, in particular for green industries, could have positive outcomes, this approach is not without risks. State support can backfire unless accompanied by strong governance.

This risk is exacerbated in the case of the European Union because the cohesion of its single market is threatened when discipline on state aid given by member states is eroded (Kleimann *et al* 2023). Therefore, it is important to have a thorough understanding of the problems that 'economic security' measures aim to solve, in order to judge the trade-offs involved in the proposed solutions.

To support the development of such an understanding, we attempt to derive a nuanced view of the economic risks that arise from economic interdependence with China in particular[5]. Based on this view, we assess the

The rise in global geopolitical tensions has coincided with deeper economic integration of EU and non-democratic countries, and an increase in the market concentration of EU imports

appropriateness of EU instruments aimed at improving economic security. We conclude that the EU has made significant steps forward in terms of ex-ante instruments, though many of them need more European coordination to avoid risks for the single market.

However, credible ex-post instruments are lacking. We see the need for a new ex-post instrument that shares the costs from economic coercion and helps countries and firms respond. Such instruments have to be underwritten by member states, and therefore the credibility of any European economic-security instrument depends crucially on a closely coordinated foreign policy.

2 What is economic security?

Despite its prominence in recent debates, the term 'economic security' is only vaguely, if at all, defined. The term has been used in varying contexts, and at times has been employed as a catch-all for policies aimed at mitigating all kinds of economic shocks, as well as a wide range of 'national/physical security' measures. This conflation of different types of risk can unsurprisingly lead to poorly targeted government interventions.

We employ a narrow definition that is centred around the notion of economic 'de-risking' from shocks, and not the use of economic measures to pursue national security objectives. We focus in particular on risks surrounding 'economic coercion' – the politically motivated disruption of supply chains and targeting of economic interdependencies.

Examples of such coercion include sanctions and trade embargoes, the weaponisation of energy markets following the Russian invasion of Ukraine, and Chinese economic coercion against Japan, South Korea, Lithuania and Australia.

In these cases, a hostile government targeted vulnerable economic sectors with the aim of inflicting economic and political damage. We assess instruments and strategies that are aimed at mitigating and limiting the risks from such deliberate and targeted economic shocks. It is noteworthy that these types of shock are not only a concern for strategic imports. Recent cases of economic coercion have actually targeted exports more than imports.

While threats to economic security can come from a range of sources, such as climate-related shocks or natural disasters, we focus on improving resilience against economic coercion for two reasons. First, the policy lessons are equally applicable to other supply-chain disruptions. Second, economic coercion includes an additional factor (the behaviour of hostile governments) not present in 'accidental' shocks.

This additional factor necessitates additional policy responses to affect other countries' incentives. As such, policies designed to address threats arising from economic coercion should also address wider risks to economic security.

We also focus on foreign-trade shocks and not domestic shocks, which can have similar implications and are part of some broader definitions of economic security. We are concerned with the interaction between economic outcomes and foreign policy, which is less of a concern with shocks of domestic origin and so the relevant policy instruments differ.

We deliberately abstract from policies that are framed as part of 'economic security' (eg. in the European Commission's Economic Security Strategy; European Commission, 2023a), but are not 'economic' in either nature or objective. With the exception of the very rare cases in which technical complexity creates monopolistic power and therefore the potential for future economic coercion[6], measures aimed at preventing technology transfers are hard to justify on economic security grounds alone.

While maintaining European technological leadership in certain cutting-edge sectors is clearly desirable, it fails to meet the definition of economic security as articulated here. Other justifications – such as maintaining an edge in dual-use technologies for defence reasons – are thus generally necessary to justify measures that restrict technology transfers.

The distinction between 'economic' security risks and national security is important for two reasons. First, economic-efficiency arguments become less important when considering policies with direct national security implications. Economic analysis can help identify the most efficient way to achieve a desired outcome, but cannot ascertain whether a policy is necessary for defence purposes.

Second, separating economic security from national security has legal implications. WTO rules give countries the ability to react to policies that harm their economic interests (eg. with countervailing duties and rebalancing of tariffs) and to call panels to adjudicate on whether rules were broken.

The WTO framework also includes exemptions for measures pertaining national security[7]. The principle that states can intervene in markets to ensure their national security in ways that would be otherwise prohibited is generally recognised. However, there has been considerable debate about the wide-ranging usage of these exemptions by the United States (see Maruyama and Wolff, 2023).

In several cases, the US has justified policies that arguably primarily have protectionist aims with such national-security exemptions (for a discussion of the role of transatlantic relations see Box 1).

The EU and the US have converged on a shared paradigm of 'de-risking', a notion that was first embraced by European Commission President Ursula von der Leyen in March 2023[13]. It is noteworthy that the EU and US have come from opposite directions to arrive at similar strategies.

Box 1. Economic security and the transatlantic relationship

While there have been regular trade conflicts between the EU and the US (such as a long running dispute on subsidies for Airbus and Boeing), these were concerned primarily with protectionist measures and support for national champions.

However, during the Trump Administration, new conflicts arose that were framed explicitly around security. While not directly comparable to the current economic security debate relating to Russia and China, certain aspects of the European discourse can be traced back to these origins.

The retreat of the United States from the Iran nuclear deal (the Joint Comprehensive Plan of Action, or JCPOA) was a leading cause of the European desire for a more autonomous foreign policy.

Even though the EU believed it to be in its interest to keep trading with Iran, the US threatened European companies with secondary sanctions if they did so (see Leonard *et al* 2019). This did not affect European 'economic security' per se, but it did advance a discourse on how to harden European trade flows against foreign interference.

In 2018, the Trump Administration put tariffs on EU steel and aluminium exports, justified by national security concerns (Department of Commerce, 2018), launching a transatlantic trade conflict with a vague notion of national security at its centre.

Since President Biden took office, the EU and the US have managed to resolve major trade conflicts. The Airbus-Boeing trade dispute was suspended[8], an agreement on transfers of personal data found[9] and the trade and technology council established[10] with the aim of preventing future conflicts through intergovernmental consultations.

The US tariffs on European steel and aluminum justified by 'national security' have been put under a moratorium, though a permanent solution has not yet been reached (Dadush, 2021). There are ongoing efforts to enhance economic security in the G7[11] and to cooperate on common concerns, such as those surrounding critical raw materials[12].

However, should political dynamics change again after the 2024 US presidential election, transatlantic relations could be tested once again and new EU-US trade disputes could arise.

In the US, the emphasis in 'economic security' has primarily been on security, representing a 'securitisation' of economic policy. Major economic policies have been announced by National Security Advisor Jake Sullivan, rather than by economic policymakers.

Many actions considered to fall under the umbrella of economic security, such as the US CHIPS and Science Act[14] or outbound investment screening[15], have been explicitly justified on national-security grounds. This stands in contrast to the European context, with the European Commission hitherto primarily concerned with economic policies and without a strong national-security mandate.

The 'Geopolitical Commission' of President von der Leyen[16] is trying to use its economic powers to assert itself as a player in foreign policy. Yet its economic-security strategy includes many measures that are not directly related to economic considerations and mirror US policies (European Commission, 2023a).

3 A brave new world of economic interdependence

The idea of using economic linkages to achieve political goals is by no means new (see Mulder, 2022). Since the end of the Second World War, outright economic sanctions have mostly been used by the US and its allies against emerging-market developing countries (Hufbauer, 2007). Even before the Russian invasion of Ukraine, there was a surge in the number of sanctions imposed by Western countries (Felbermayr *et al* 2020).

However, while sanctions have historically been mostly used by Western countries, economic coercion is by no means an exclusive to the West. The examples of such measures targeting Western countries range from the oil embargo during the Yom Kippur War in 1973 (Hansen, 2023) to import restrictions on Norwegian salmon by China after the 2014 Nobel Peace Prize for Liu Xiaobo (Harrell *et al* 2018).

Given the dominance of Western economies in finance and technology, the types of economic linkage targeted by non-Western economies have historically often been access to raw materials. However, recent decades have seen a remarkable shift in the goods that are available for use in economic coercion against the West.

Figure 1 shows the breakdown of the main categories of EU imports by the political systems of source countries, as defined by Freedom House. While raw materials were long primarily imported from non-free countries, as recently as 2001 only 10 percent of imports of intermediate inputs came from such countries. By 2019 this share had increased to almost 40 percent.

As a result, EU industry imports many more intermediate goods from countries with authoritarian political systems. Intermediate imports are often more specialised and differentiated, limiting their substitutability compared to commodities. This thus represents a new type of risk. Meanwhile, advanced technologies are increasingly dependent on specialised materials as critical inputs, meaning raw materials have also become more susceptible to economic coercion (Le Mouel and Poitiers, 2023).

One additional and often overlooked source of European vulnerability is export dependency. China in particular has become an increasingly important market for Western exports (Figure 6), with approximately 10 percent of German passenger car exports in 2022 going there, for example[17].

As will be shown, this means that import bans are also available as a means for China to put political pressure on Western governments. As Baqaee *et al* (2024) showed, the potential economic costs of sudden trade disruptions with China for a country like Germany are significant (they assess that the effect of a total cessation of trade with China for Germany would be 'severe but not devastating').

Figure 1. EU import sources by political system

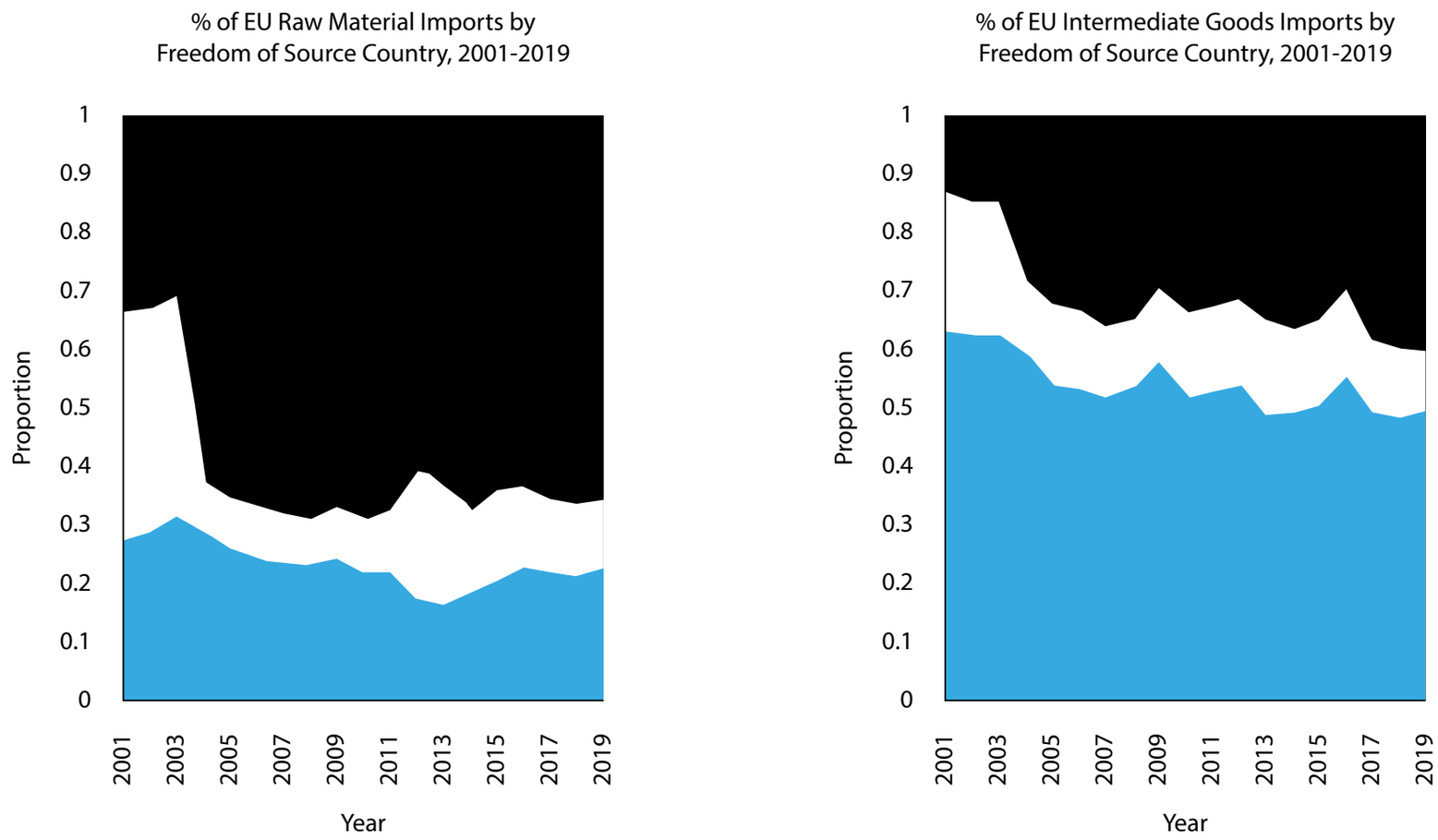
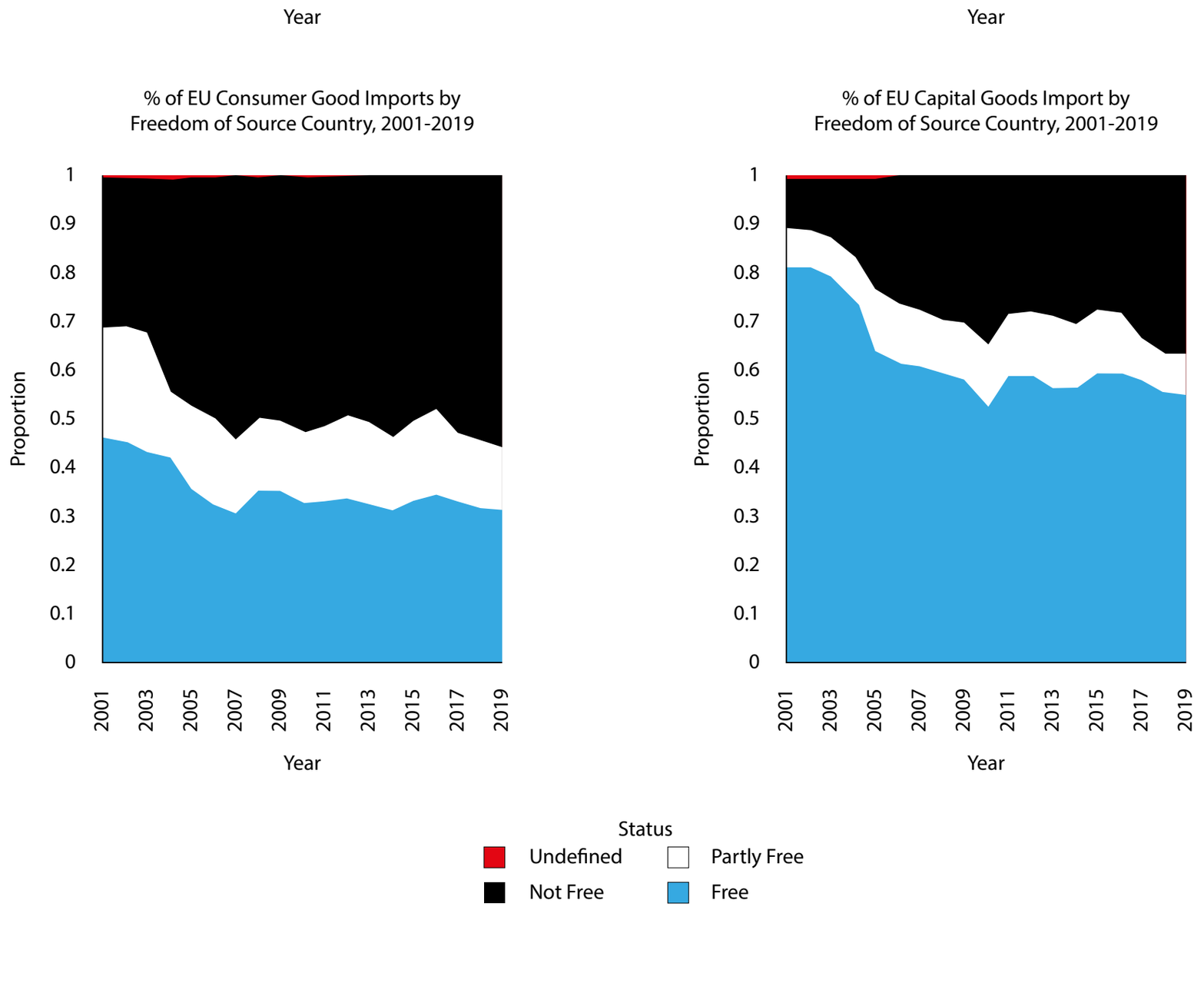


Figure 1. EU import sources by political system cont.



Source: Bruegel basted on Eurostat, UNCTAD & Freedom House.

4 The threat of economic coercion

Economic coercion comes in many shapes and forms. Adachi et al (2022) tallied Chinese coercive methods since 2012 (Figure 2). Many measures targeted individual firms, while trade restrictions have been the most common form used to target countries. Within these trade restrictions, import restrictions (China blocking the imports of goods from foreign markets) have been used more often than export restrictions[18].

Unlike Western sanctions that follow formal legal procedures and can be challenged in courts, measures taken by China are often informal. Documentation detailing measures can be difficult to find, and targeted entities might thus find it difficult to challenge measures even when avenues to do so might exist (Hackenbroich *et al* 2022).

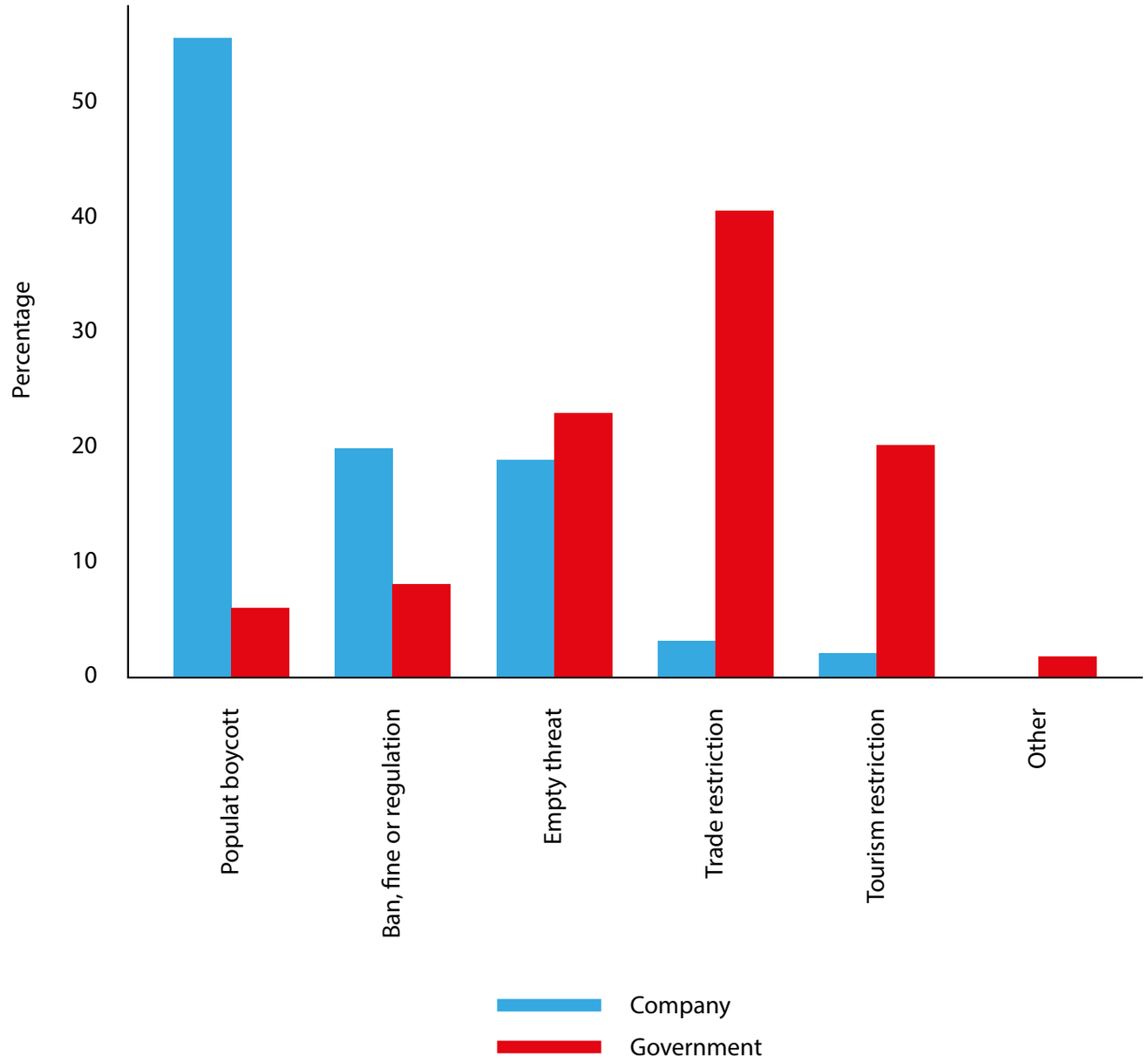
A particularly problematic example is popular boycotts against certain foreign brands, individuals or firms. While sometimes genuine, these movements to encourage firms and consumers to punish certain firms are often stoked by state-controlled media and on social media[19]. They represent the most common form of economic coercion used by China against firms, and are particularly difficult to attribute to undue state intervention.

The experiences of trade wars and Western sanctions against Russia provide some insights into what types of goods are vulnerable to economic coercion. In episodes such as the China-US trade war that began in 2018, trade diversion has been a major feature, limiting the effects of trade restrictive measures (Fajgelbaum *et al* 2023).

Similarly, sanction circumvention and alternative sourcing pose major challenges for the effectiveness of Western sanctions against Russia (Babina *et al* 2023).

The effectiveness of any type of coercive economic measure depends on the market power of a country or coalition. If alternatives are widely available, a targeted economy can easily switch its sources of imports for a product.

Figure 2. Forms of Chinese economic coercion



Source: Adachi et al (2022).

Similarly, if alternative export markets exist, a bilateral trade relationship cannot easily be weaponised. This rules out most commodities from being used or targeted effectively for economic coercion, as they have many sources and markets. Even where a high degree of market concentration is found, this does not necessarily imply high monopolistic power.

The contestability of a market also depends on barriers to entry for newcomers. Many of the products for which there is a high degree of market concentration are low-tech products, such as artificial flowers and electric blankets (Mejean and Rousseaux, 2024).

If the dominant producers would limit exports of these products, it would be rather easy for new companies to enter the market. This was the case for rare gases (neon, krypton and xenon), the supply of which was disrupted by the Russian invasion of Ukraine (Georgitzikis and D'Elia, 2022). Their prices spiked after the outbreak of the war, but came down rather quickly as new producers entered the market[20].

Furthermore, there might exist close subsitutes that might not be employed presently but could become commercially viable if the supply chain of the incumbent technology is disrupted. Examples of this dynamic have been documented during trade embargoes (Mulder, 2022). However, it can be difficult to assess the feasibility of such substitution before an actual disruption occurs.

An economy can have monopolistic power for several reasons. First, a natural resource might only exist in a few countries, giving them effective control over where the supply goes. Second, infrastructure bottlenecks might create monopolistic power in segregated markets.

This was the case for Russian pipeline gas in the wake of the invasion of Ukraine: a lack of liquified natural gas (LNG) capacity in central Europe allowed Russia to hike prices in European gas markets.

Third, economies of scale or industrial policy can lead to dominance on certain markets, as is the case of China in the solar panel industry (García Herrero *et al* 2023). Fourth, advanced technological capacities might give monopoly power. An example for this would be ASML in the chip industry (Poitiers and Weil, 2021; Kleinhans and Baisakova, 2020).

The 'contestability' of a market is also important. Only if a monopoly market can be maintained over time can it be exploited over extended periods without the risk of losing future markets.

In 2022, there was considerable concern over the supply of certain gases that were primarily produced via a Russian-Ukrainian supply chain. However, alternative sources were brought online relatively quickly, preventing lasting shortages (Darvas *et al* 2023).

To induce harm that is macroeconomically significant, the impact of a bilateral flow needs to have a material impact on the overall export or import performance of the targeted economy. For certain goods, in the fields of health, defence or clean energy, for instance, disruptions to imports may be highly damaging on have some non-economic outcome, with prominent examples being personal protective equipment and vaccines during the COVID-19 pandemic.

In highly diversified advanced economies such as the EU, the capacity to induce truly significant shocks, either macroeconomic or otherwise, is limited to a very small number of strategic goods. However, in many cases of economic coercion, the harm is market- and industry-specific rather than macroeconomic.

Though few individual products are of such importance that they can affect the economy as a whole, targeted measures can easily harm politically important constituencies, and thus exert political pressure on policymakers.

In the following, we consider two recent cases of economic coercion that illustrate how economic interdependence can be weaponised: the measures taken by China against Australia and Lithuania since 2020.

4.1 Australia: a tale of two sectors

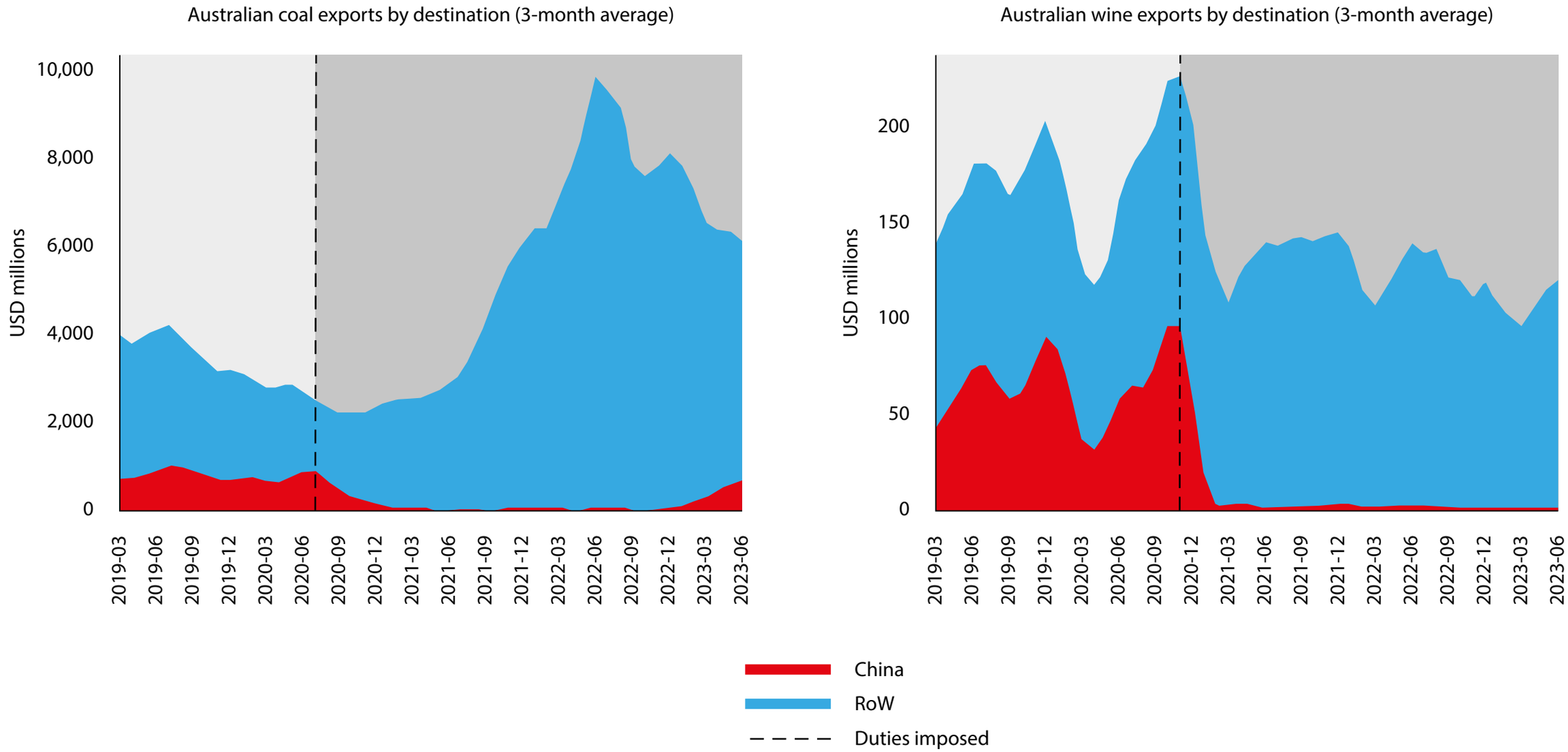
In mid-2020, following then-Australian Prime Minister Scott Morrison's calls to open an investigation into the origins of the COVID-19 pandemic[21], China began a campaign of economic coercion against Australia that only began to be eased in early 2023.

It targeted Australian exporters and introduced *"discriminatory tariffs on wine and barley"* and *"informal and WTO-illegal bans on coal, beef, lobster, cotton, wood, nickel and copper concentrates"* (Urden 2023a)[22]. As a result, China's share of Australian exports fell from its mid-2021 peak of almost 45 percent to less than 30 percent by the end of 2022[23].

The Australian economy as a whole successfully navigated the coercive measures introduced by China. The value of Australian exports rose between 2020 and the end of 2022, largely driven by energy exports to Asian markets other than China.

There was however important variation in the impacts on the various targeted sectors. For the coal sector, the decline in exports to China was more than offset by higher exports to the rest of the world, in particular to Asian countries that were also indirectly affected by China's actions (Figure 3, Panel A)[24].

Figure 3. Chinese economic coercion against Australia



Source: Bruegel based on Australian Bureau of Statistics (left) and UN COMTRADE (right).

Significant export diversification, coupled with high coal prices following the Russian invasion of Ukraine, meant that Australian coal exporters enjoyed surging import revenues over the period of the unofficial embargo.

This makes for a stark contrast with Australian wine exporters. Because of a 2015 free trade agreement[25], Australian wine exporters had been at an advantage in China compared to many other wine-exporting countries, making China an important export destination.

However, following the imposition of countervailing duties as high as 218 percent in late 2020, wine exports to China collapsed from approximately 38 percent of total Australian wine exports in 2019 to zero since 2022. Unlike coal, the industry failed to expand into other markets.

Consequently, monthly Australian wine exports in June 2023 were down over 40 percent from their October 2020 peak. Chinese duties, coupled with a strong harvest, led to a significant oversupply of Australian wine[26], depressing prices and adversely impacting the industry[27].

The two industries detailed here are representative of the broader range of targeted industries. Some, such as barley, succeeded in diversifying away from Chinese buyers (to Saudi Arabia) and saw their exports grow over the period in question. Lobster and wood exporters on the other hand failed to move into new markets and suffered the same fate as their counterparts in the wine industry (Buckland *et al* 2023).

4.2 Lithuania: much ado about nothing?

The trade restrictions introduced by China against Lithuania in 2021 marked the most serious incident of Chinese economic coercion against an EU member.

The relationship between the two countries had been particularly fraught since the formation of a new Lithuanian government in 2020[28], but broke down entirely in mid-2021 when the Lithuanian authorities announced that they would allow a Taiwanese representative office to be opened in Vilnius[29].

After two years of an essential trade ban (detailed below), the Lithuanian government reported in November 2023 that 'most' Chinese trade measures had been lifted[30].

Given the opacity of China's actions, it is difficult to disentangle precisely which measures were implemented and when. However, the European Commission (2022) detailed that the original measures enacted included disruption to logistics networks (leading to more expensive and delayed freight deliveries), difficulty obtaining trade credit insurance for imports, and general disruption to supply chains containing Chinese components.

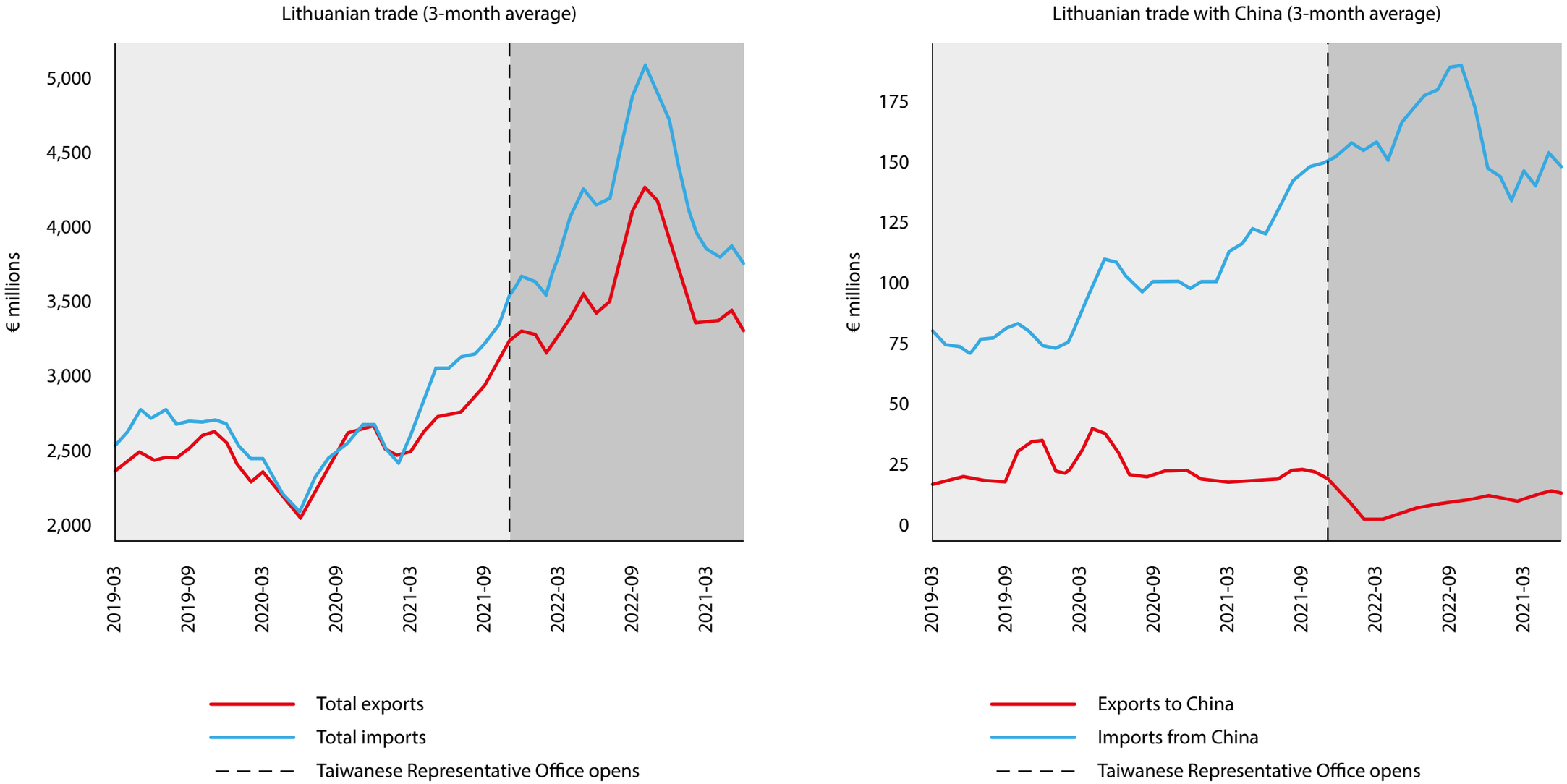
These measures were escalated following the actual opening of the Taiwanese office in November 2021, to go beyond direct trade between the two nations. They also targeted Lithuanian participation in global supply chains, with products from other European countries containing Lithuanian components being threatened with rejection by Chinese customs authorities.

Official import bans on certain products were introduced in 2022, with China relying once again on spurious justifications, such as a 'lack of documentation'[31].

Lithuanian exports to China fell by two-thirds between 2020 and 2022, but imports from China grew by the same amount over the period in question, which reinforces the idea that China most often targets countries' exports.

Neither Lithuanian total exports nor total imports were significantly impacted, which is unsurprising given that China made up just 1% and 4% of Lithuania's 2020 exports and imports respectively[32].

Figure 4. Lithuanian exports and imports to the world (left) and to China (right), 3-month average in € millions



Source: Bruegel based on Eurostat.

Box 2. Lithuanian support scheme

In April 2022, the European Commission approved under EU state aid rules a Lithuanian loan scheme designed *"to support and facilitate access to finance by companies affected by the exceptional circumstances resulting from China's discriminatory trade restrictions on Lithuania"* (European Commission, 2022). This was approved to last until the end of 2027 or the end of the trade restrictions, whichever came first. However, because of a lack of uptake, the scheme was wound down in 2023[35].

Administered by INVEGA, the Lithuanian national promotional institution, the scheme was capped at a maximum of €130 million overall, and at €5 million per firm. Access was limited to Lithuanian firms for which the *"proportion of either imports from or exports to China represents at least 25% of the beneficiary's total imports or exports in 2021"*, and that were unable to receive financing on the market (which had to be proven by rejections from three financial institutions). The loans had to be used: (i) to source inputs from different sources, (ii) to explore the possibility of entering new markets or (iii) to use *"the time to undertake such efforts."*

Estimates at the time of approval were that there were 130 potential beneficiaries, with this expected to increase to up to 500 as Chinese restrictions persisted and grew. However, only three firms, each an SME, made use of the support offered. The total amount of loans granted was €4.22 million, just 3 percent of the maximum amount permitted.

However, as in the case of Australia, certain sectors were negatively affected by the measures, with two of the three firms claiming assistance under a national support scheme (Box 2) operating in the solar PV industry[33].

Several observations can be made on the joint experiences of Australia and Lithuania of Chinese economic coercion[34]. First, exports to China were targeted more strongly than imports. Second, despite significant trade restrictions from one of the world's largest economies, neither country suffered macroeconomically. Third, targeted industries can emerge unscathed without government intervention, largely through successful diversification.

As Australian coal and barley exports showed, commodities are particularly poor targets for economic coercion as global markets provide alternative buyers. However, it also shows that even if the wider economy can withstand coercion, certain sectors can be strongly impacted.

The markets where Chinese coercion had the greatest effects (wine, lobsters and wood in Australia) are macroeconomically insignificant, yet their targeting affected some constituencies. In other words, the inflicted damage was political rather than macroeconomic.

5 Where is the EU exposed to economic coercion?

As monopolistic power is a necessity for economic coercion, potential vulnerabilities can be identified by looking at market concentration. The Herfindahl-Hirschman index (HHI) provides an index that measures market concentration. It is used widely not only for assessing competition cases, but also in defining economic security risks (European Commission, 2021a; Jaravel and Mejean, 2021; Welslau and Zachmann, 2023).

The HHI has a value between 0 and 1. The lower the value, the more competitive a market. In competition policy, any market with a value above 0.25 is considered indicative of a high degree of market concentration, and any market with a concentration above 0.6 is considered 'monopolistic' (US Department of Justice, 2010).

While these measures might not apply one-to-one to import vulnerabilities, they provide a yardstick of how concentrated import markets are.

Figure 5 plots the distribution of HHI values of EU imports by product category for 2001 and 2019[36]. For easier comparison, estimated distributions for both years are displayed in the right panel. We highlight goods with an HHI above 0.6 as monopolistic and thus problematic.

This is a conservative choice, compared to the threshold values used in other analyses (an HHI of 0.4 in the case of the European Commission). However, this analysis is meant to illustrate the evolution of EU import markets and we abstract from the second stage of import concentration analysis, justifying a more restrictive approach[37].

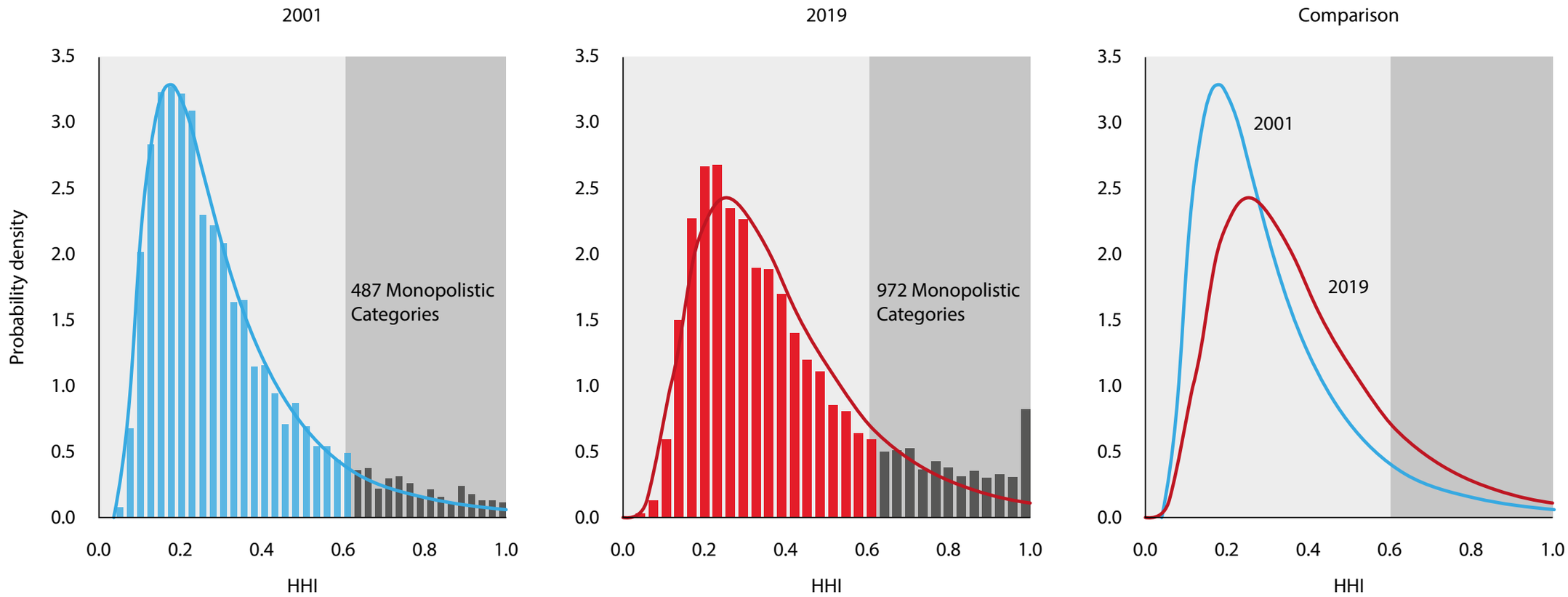
Between 2001 and 2019, the distribution of EU import market concentration shifted considerably to the right. While in 2001, 487 products had concentrations considered 'monopolistic', in 2019, 972 products fell into this category.

Table 1 provides for the EU a breakdown of the types of product that were in highly concentrated markets in both 2001 and 2019. In both periods, most of the products in highly concentrated markets were manufactured goods.

For instance, in 2019, 626 products were manufactured goods, but they accounted for only 11 percent of the value of manufactured goods imports into the EU. This was more than double the 5 percent of the import value of manufactured goods falling into the 'problematic' category in 2001.

For non-fuel raw materials, 22 percent of products were in monopolistic markets in 2019. While the share of value of non-fuel raw materials in monopolistic markets did not change significantly over the time period in question, many more of the highly concentrated goods categories were classified as 'critical raw materials' in 2019 than in 2001.

Figure 5. Evolution of concentration of EU imports



Source: Bruegel based on Eurostat.

Similarly, many more of the highly concentrated manufactured goods imports are 'high tech' goods, with the share increasing from 25 percent to 43 percent. A significant part of the increase can be directly attributed to China. It was the main source country for 20 percent of the highly concentrated import categories in 2001, with this share more than doubling to 49 percent in 2019.

Meanwhile, the share of the US in concentrated EU imports roughly halved in almost all categories (for an analysis of the trends, see Welslau and Zachmann, 2023).

Overall, EU imports of both raw materials and manufactured goods were much more concentrated in 2019 than in 2001. This shows that a high degree of market concentration is not merely a feature of a few goods categories that might have been supported through strategic Chinese industrial policy, but rather the effect of an increase in market concentration across the entire spectrum of imports.

Therefore, a strategy to limit import concentration cannot be focused only on strategic imports, as potential targets for import bans are plentiful and new ones are likely to arise in an overall concentrated market environment. An effective diversification strategy should therefore aim to lower the degree of market concentration more generally.

It is also important to note that import dependencies alone are not necessarily concerning. Among the categories of goods for which Mejean and Rousseaux (2024) found the EU to be reliant on highly concentrated import markets are, for instance, artificial flowers and camping flasks.

While shocks in the countries of origin would likely lead to EU import disruptions in these sectors, it seems implausible that these shocks would cause social welfare losses significant enough to warrant government intervention.

Table 1. Breakdown of highly concentrated import markets

	Year	# Products	Products	Value	Products HT/CRM	Value HT/CRM	Products China	Products US
Raw materials								
Total	2001	71	15%	4%	7%	4%	13%	21%
	2019	110	22%	2%	6%	18%	16%	11%
Non fuels	2001	66	15%	7%	8%	8%	14%	20%
	2019	101	22%	9%	7%	21%	17%	11%
Manufactured goods								
Total	2001	348	9%	5%	11%	25%	120%	37%
	2019	626	15%	11%	10%	43%	49%	19%

Note: HT = high tech goods according to classification by the United States Census Bureau.; CRM = critical raw materials as defined by the European Commission.
Source: Bruegel based on Eurostat.

There are important precedents for the weaponisation of import vulnerabilities. These include the Chinese threat to ban exports of certain critical raw materials during a 2010 trade dispute with Japan[38], and recent export restrictions on critical minerals[39].

However, most cases of economic coercion by China have either directly targeted companies operating on its markets or exports to China. This stands in contrast to the almost exclusive focus of economic security on risks stemming from Western imports from China.

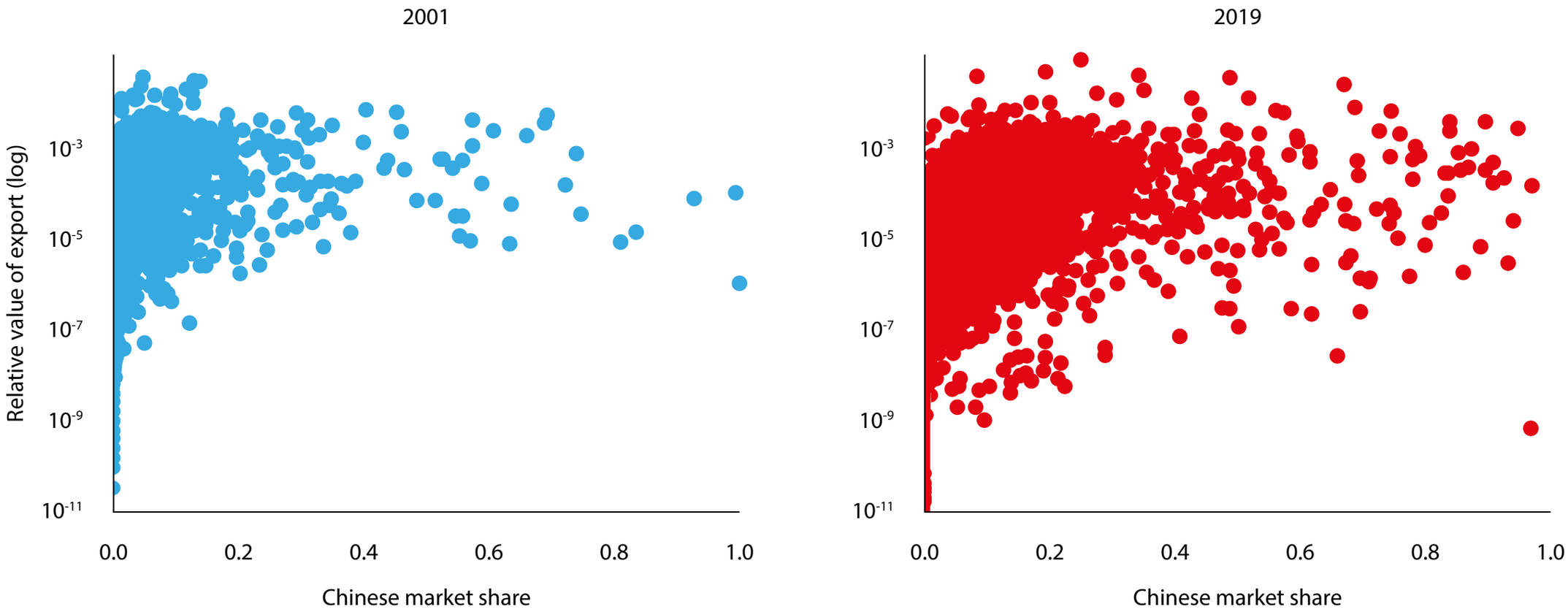
As Adachi *et al* (2022) showed and the Australian and Lithuanian cases illustrate, imports from China are not typically the primary vulnerability for economic coercion. Instead, these past experiences have shown that China tends to weaponise access to its domestic market for foreign exporters.

Given that a market must be sufficiently large to have monopolistic power as an export destination, China is virtually the only country of concern to the EU for this type of risk[40]. While other countries can also harm EU export interests, they are unlikely to be sufficiently large to inflict significant damage.

Therefore, we use in Figure 6 Chinese market shares as proxy for export vulnerabilities instead of the HHI index. The economic importance of an export is measured by its relative value (it's share of total exports to China). A product in the lower left corner is of relatively low value and is not exported a lot to China, whereas a product in the upper right corner is of high value with most of it being exported to China.

Overall, a large shift to the right is evident. In other words, there is now a much larger number of products where a Chinese embargo on EU exports would inflict significant harm, increasing the number of potential targets for Chinese restrictions.

Figure 6. Concentration of EU export markets



Source: Bruegel based on Eurostat.

As in the case of the increasing import concentration, the increase in Chinese market shares in exports represents a structural shift rather than something that is product-specific. A focused strategy on the most exposed exports might limit some potential harm in the short term, but the number of potential targets is so high that broader diversification is necessary and overarching policy instruments are required.

6 Instruments of economic security

The increased exposure of the EU to economic security risks has rightly drawn the attention of policymakers. Various initiatives have been proposed with the aim of increasing the resilience of the European economy against such risks. Given the different types of threat, these initiatives rightly include a wide range of instruments[41].

Table 2 provides an overview of the policy instruments relevant to the economic-security debate, including both those announced under the auspices of economic security but that are in fact more pertinent to national security, and policies relevant to addressing economic security risks that have not yet been put forward.

We distinguish them depending on the nature of the threat (eg. whether it targets exports or imports)[42] and the intended timing of implementation (pre-emptive, ex-post or both, which we term 'overarching'). It is noteworthy that many of these policies have the potential to improve the resilience of the European economy in areas beyond responding to economic security threats.

As mentioned, Table 2 includes a number of policies mentioned in the Commission's Economic Security Strategy but that are arguably more concerned with non-economic risks. The downsides to many cyber-attacks or research interference are not primarily economic in nature.

Table 2. Instruments of economic security

Vulnerability	Threat	Ex-ante instruments	Ex-post instruments	Overarching instrument
Economic security				
High export concentration	**Targeted trade embargoes**	**Diversification** - Free/Preferential Trade Agreements (FTAs/PTAs) - Secondary instruments, eg. export credit agencies, development policies, 'clubs', TTCs, Global Gateway	**Bespoke national support** eg. state aid-sanctioned scheme in Lithuania	**Anti-coercion instrument** - Introduction of proportionate retaliatory measures
High import concentration	**Disruption of supply of crtical components**	**Diversification** - FTAs/PTAs - Secondary instruments (see above) **Increase domestic production** - Industrial policy - Strenthening the single market **Ensure a level playing field** - Foreign Subsidy Regulation - Trade defence instruments	**Missing:** EU support	**Internal Market Emergency and Resiliance Act** - Monitoring, stockpiling, joint procurement and potential 'priority rated orders'.

National security	**'Critical' competitive advantages** ('Narrow yard with high fences')*	**Loss of technological advantage in dual use goods and the control over their production**	**Boost industry** - Industrial policy
			Prevent leakages - Export controls* - Outbound investment screening* - Toolkit on Tackling Foreign R&I Interference*
			Ensure a level playing field - Foreign Subsidy Regulation - Trade defence instruments
	Not otherwise classified		**Protecting infrastructure** - Cyber Resilience Act* - NIS2 Directive*
			FDI Screening*
			EU Standardisation Strategy*

*Note: Includes current/proposed EU policy measures, as well as those we believe are missing. * denotes policies or ambitions put forward under the umbrella of economic security that generally fall outside of our definition[43].*
Source: Bruegel.

The Commission has declared certain technologies to be of particular concern because of *"the enabling and transformative nature of the technology; the risk of civil and military fusion; and the risk of misuse of the technology for human rights violations"* (European Commission, 2023d).

The latter two criteria are not relevant in terms of our narrow definition of economic security. The former, which the Commission defines as assessing the technology's *"potential and relevance for driving significant increases of performance and efficiency and/or radical changes for sectors, capabilities, etc"*, could fall under the remit of economic security only in sectors where high degrees of technological complexity create monopolies, as described earlier.

In the following, we discuss the role of some instruments in more details, as part of four complementary strategies to enhance economic security: mapping of vulnerabilities; diversification of imports and exports; industrial policy and technology security in strategic sectors; and ex-post policies to help redress political damage.

6.1 Mapping vulnerabilities

The first step of responding to economic security concerns is to identify risks. Global value chains are enormously complex and not all dependencies are direct (Qiu *et al* 2023). Coercive measures can go beyond direct bilateral trade, as was the case with China's actions against Lithuania.

As such, a detailed understanding of the EU's dependencies on other countries for both exports and imports is necessary. This would allow authorities to identify potential vulnerabilities ahead of shocks, and assist affected firms, in particular SMEs, to diversify their supply chains and mitigate the risk in question.

Hackenbroich *et al* (2022) argued that there may be scope for an EU body to carry out detailed data analysis for this purpose.

Monitoring supply chains by requesting, and in some instances requiring, firms in strategic sectors to disclose information on their suppliers, stocks and productive capacities is a key, and controversial[44], component of the proposed EU Internal Market Emergency and Resilience Act[45].

Similarly, the European Chips Act entails mapping and monitoring the semiconductor supply chain to assess ex-ante risks of potential import disruptions[46]. Depending on the importance of a sector, a balance has to be found between the administrative burden on firms and the benefits from further insights. For instance, informational requirements should be higher on those sectors flagged by Mejean and Rousseaux's method (2024) as being at risk.

However, awareness of risks alone does not directly lead to mitigation measures; economic incentives have to align as well. While over 95 percent of firms surveyed in the EIB Investment Survey (European Investment Bank, 2023) had experienced some form of disruption to international trade, less than half of them had changed or were planning to change their sourcing strategies.

Even when potential downsides are large enough to warrant a change in sourcing, there might not be readily available alternatives. This leads us to the next strategy.

6.2 Diversification

Since monopolistic power is a necessary condition for effective economic coercion, trade diversification is the most effective strategy to reduce vulnerabilities, as it can lead to more competition across a wide range of imports and exports.

While precise results change depending on the criteria used to determine dependence, there has been significant churning in the products in which the EU has been overly import dependent (Vicard and Wibaux, 2023).

Failing to further comprehensively diversify both imports and exports will likely lead to more goods falling into the concerning range of high export or import concentration. Otherwise, in focusing on individual goods in structurally concentrated markets, policymakers will be constantly racing to address different areas of concern.

To achieve greater diversification, a combination of policy tools offers the most promising avenue. First and foremost, free and preferential trade agreements (FTAs/PTAs) open new markets for both exporters and importers.

The EU has made progress in broadening its level of trade covered under PTAs. As of 2020, half of extra-EU exports were covered by reciprocal PTAs, up 8 percentage points from 2010 as trade agreements with Canada, Japan and Korea came into force (Dadush and Dominguez Prost, 2023)[47].

The December 2023 agreements[48] between the EU and Chile, an important exporter of some CRMs, to enhance and modernise their existing FTA, also shows how these agreements are not static, and should be updated if needed to reflect the increased focus on economic security.

However, mainly because of domestic political pressure, the EU has struggled to conclude trade agreements with major trading partners such as the Mercosur countries, while even negotiations with close allies like Australia have proven difficult[49].

Besides the difficulty of ratifying FTAs, there are other limits to relying on FTAs for diversification. Many of the products for which the EU has problematic import dependencies do not have significant tariffs precisely because there is no European industry that would justify protective measures.

Where Most Favoured Nation (MFN) tariffs offered to all WTO members are already very low, the EU cannot offer significantly better market access through an FTA compared to the access that, for instance, China has. This is the case for CRMs, many of which have no tariffs at all applied to them (Le Mouel and Poitiers, 2023).

Therefore, a diversification strategy must complement FTAs with external financial instruments[50]. The European Commission aims to harmonise and streamline European development assistance under the umbrella of the Global Gateway.

Beyond its primary objective of promoting economic development globally, this initiative has as a stated goal to support the EU by *"strengthening the resilience of its supply chains, and to opening up more trade opportunities for the EU economy"* (European Commission, 2021b, p.2).

To an extent, this is indeed already happening. In October 2023 the EU signed Memoranda of Understanding under the Global Gateway framework with both the Democratic Republic of Congo and Zambia to deepen cooperation around the development of resilient value chains for critical raw materials, which could help to improve import diversification[51].

More should be done in this area, such as potentially investing in infrastructure in northern Africa to further diversify European energy imports (as argued by Rizzi and Varvelli, 2023).

Export credit agencies (ECAs) should play an important role in this strategy, including the potential creation of a European export credit agency. ECAs are state-owned or publicly financed bodies that are used to support exports by providing a range of financing instruments (primarily insurance and guarantees, but also loans) at below market rates to de-risk trade.

Going beyond facilitating direct exports, they can also be used to support investments in third countries which, if targeted appropriately, can ultimately improve diversification of supply. A European ECA could compliment the 24 national ECAs (European Commission, 2023c)[52]. The support in question is significant, with EU ECAs in 2021 insuring projects amounting to approximately €90 billion (Schlögl *et al* 2023).

The ECAs' funding could be boosted and applied strategically to support the objective of economic security. It will not be commercially viable in a high-wage economy to produce many of the products for which the EU is reliant on imports from China. Some raw materials do not exist in Europe, or local resistance to their extraction could be too high.

In such cases, ECAs can play a critical role in promoting investments in alternative sourcing in partner countries (Le Mouel and Poitiers, 2023). Export-promotion offices could also be useful to help firms identified as being overly reliant on a particular export market to identify and access new markets.

The Enterprise Europe Network (EEN), a Single Market Programme-funded umbrella of national SME support organisations (including chambers of commerce and government agencies) already offers assistance to SMEs in the areas of 'resilience' and 'internationalisation'. This role, however, could be boosted, with awareness of the network at just 9 percent among SMEs[53].

6.3 Targeted industrial policy and interventions

For sectors that combine a high degree of dependency with a high degree of economic importance, diversification might not be enough to safeguard economic security. There are very few sectors from which macroeconomically significant impacts might arise because of supply chain shocks.

As noted, concerns beyond economic outcomes, such as defence and health, may justify such policies in other areas, but this group should also be limited. Three types of strategies are possible: (i) maintaining strategic reserves; (ii) growing domestic production; or (iii) improving productive capacities in third countries.

In some cases, stockpiling a certain buffer level will often be the most cost-effective option, but it is not always feasible. Certain goods (like medicines) might spoil, and in certain fast-moving sectors (for instance PVs), technology quickly becomes obsolete. As such, this should play only a limited role.

The global trend thus far has been to prioritise boosting domestic supply via industrial policy. Examples include the European Chips Act and the Net Zero Industry Act in the EU, the Inflation Reduction Act and CHIPS and Science Act in the US, and the K-Chips Act in Korea.

However, competing policies have led to costly subsidy races even among likeminded partners, and heavy-handed reshoring policies can have unintended consequences. Javorcik *et al* (2022) estimated that friend-shoring could generate global real GDP losses as high as 4.6 percent.

Reshoring drug production to avoid shortages could lead to prices increasing by up to 30 percent (Galdin, 2023). Import restrictions have likely contributed to shortages of infant formula in the US[54].

Meanwhile, producing green technology in Europe would lead to much higher decarbonisation costs, slowing the green transition and Europe's attempts to diversify away from Russian hydrocarbons. In the EU, the emphasis on national state aid also poses risks to the single market (see Kleimann *et al* 2023; Tagliapietra *et al* 2023)[55].

In the instances in which increasing domestic production is justified, a bespoke strategy should be designed for the sector in question that aims to minimise distortions and leverage the comparative advantages of the EU in that area.

For instance, McWilliams *et al* (2024) argued that an EU industrial policy for the solar panel industry should focus on recycling and innovation, not import substitution. Given the different abilities of EU countries to support their domestic industries, a 'Europeanisation' of state-aid tools such as the Important Projects of Common European Interests (IPCEIs) will be indispensable if single market fragmentation is to be avoided.

Currently, IPCEIs and similar policies, such as the European Chips Act and funding for clean tech through the Temporary Crisis and Transition Framework, rely on national funding. While they have to be part of a common European framework, individual projects are chosen via opaque processes by EU countries based on (sometimes competing) national interests. Project selection should rather be based on more thorough, transparent methodologies (Poitiers and Weil, 2022).

Internationalising industrial policy provides a very promising avenue to increasing the security of supply while simultaneously minimising protectionism, though international policy coordination will be challenging. Variations of this approach include critical raw materials (CRM) 'clubs' and the establishment of clean-tech partnerships to leverage different countries' relative comparative advantages, as proposed by García-Herrero *et al* (2023).

Beyond growing domestic production, technology security measures (such as export controls or outbound investment screening) to prevent diffusion in the aforementioned key sectors at risk of complexity-driven monopolisation, must also be complemented by policies that reinforce and strengthen existing advantages, through support for R&D, skilled immigration and via bespoke industrial policies.

In addition, policymakers must be aware of the risk of reciprocity in these measures (as was the case with China in 2023[56]) and should therefore be judicious in their application.

In sum, there may be cases in which the risks associated with supply disruption warrant application of industrial policy to promote alternative supply chains, either in the EU or in other countries, or the imposition of technology security measures.

However, policymakers should not pretend that this is a cost-free approach, and need to weigh losing the gains from trade against the potential welfare losses from supply chain disruptions. If they opt for industrial policy, how exactly they choose to design this approach, in particular to minimise any protectionist elements, is critically important.

6.4 Ex-post instruments

While some goods and industries are of such strategic importance that they warrant state intervention, as discussed above, it would be prohibitively expensive to do so for all smaller industries that are exposed to economic security risks (think for instance again of the artificial flower industry identified by Mejean and Rousseaux, 2024).

Therefore, ex-ante policies alone will not suffice. Ex-post policies can help deter targeted attacks against such industries and can soften their impact when they do occur. The first instrument in this regard is the Internal Market Emergency and Resilience Act.

In cases of severe supply chain disruptions or the risk thereof, this law allows the EU to impose reporting obligations and build-up strategic stockpiles and, in case of crisis, it lists the potential ways in which the EU can intervene in supply chains (Ragonnaud, 2024).

However, the primary ex-post EU instrument to this end is the new Anti-Coercion Instrument (ACI, Regulation (EU) 2023/2675), a wide- ranging trade defence instrument intended to be applied in retaliation in cases of economic coercion against an EU country. To quote the Commission, *"the primary objective of the ACI is deterrence"*[57], and it will therefore be considered a success if it is never used. However, if triggered, the retaliatory measure could apply in virtually all areas of economic policy.

This instrument should be complemented with another instrument that helps share the burden of economic coercion. This would entail providing affected firms with financial and perhaps logistical support to enable them to find new markets for their exports or imports.

The logic behind supporting firms is twofold: it removes the ability of adversaries to target groups and inflict political damage on European countries, which they could try to leverage to change policy, and it supports firms that will likely have suffered a serious shock to their business model through no fault of their own.

While in most cases the economic damage from economic coercion will be small enough that national government could finance support for affected workers and firms, there would be several benefits from setting up an EU-wide tool.

EU solidarity assistance would reinforce the signal that an attack against one country is an attack against all and would disincentivise divide-and-rule strategies on the part of third countries[58].

It would also potentially allow firms in other countries that are indirectly affected by the coercive measures (eg. German firms that export to China but use Lithuanian components, in the case of sanctions against Lithuania) to be supported without the need for new state-aid schemes to be approved in each country.

Such a measure to fortify the joint EU response will become more important as other European countries, such as Czechia, pursue foreign policy akin to that of Lithuania (McVicar, 2023).

The challenge of this proposed instrument is that it introduces the potential for moral hazard. If firms believe that the EU will bail them out in the event of supply chain disruption, they may choose to deepen their exposure to geopolitical risks, rather than diversifying, increasing their potential exposure to economic coercion.

Similarly, countries themselves could feel emboldened to pursue foreign policy beyond the EU consensus, safe in the knowledge that their firms will be supported by other member states[59].

Therefore, any new ex-post instrument should be accompanied by new incentives for companies to diversify their supply chains and customer bases to limit potential abuse through moral hazard, as well as further progress on common foreign policy.

Part of this could be accomplished through the nature of the support itself. For instance, limiting support to capped, concessional loans with strict terms of use would reduce any perverse incentives to double down on critical imports from China.

Eligibility requirements should also be used to minimise these risks: receiving state aid could be made conditional on previously having fulfilled certain reporting obligations, having conducted risks assessments ('supply chain stress testing') or on companies insuring themselves against certain economic security risks in private markets[60].

There could be some symbiosis with the supply chain monitoring detailed previously, with firms operating in dependent sectors required to demonstrate diversification efforts before being deemed eligible for support, for example.

Overall, there is a need to strike a balance in both the nature of the instrument and the eligibility: too generous and lenient and there is the risk of moral hazard; too frugal and restrictive and the instrument could become pointless, unable to adequately support those negatively impacted and therefore failing to negate the political pressure points[61].

For the success of both the deterrence value of the ACI and any EU-wide support scheme, a common or at least strongly coordinated foreign policy is a prerequisite. All EU countries should have to underwrite the potential backlash against a forceful application of the ACI and be willing to pay for EU assistance for affected companies, even if they did not necessarily agree with the action that provoked the coercion in question.

As detailed in Hackenbroich *et al* (2022), when considering their responses, countries must weigh up both the underlying policy and the value of preserving EU solidarity and unity against coercion, which will likely be successful if it succeeds in dividing member states.

With Lithuania, this was not the case, as other EU countries appeared unwilling to pay a price for a foreign policy action taken by Lithuania alone. Despite public proclamations of outrage by other EU countries, there was neither material support nor immediate retaliation against China for what even the Commission described as 'discriminatory trade measures'[62].

In contrast to the US, which promised a $600 million export credit agreement to Lithuania[63], and Taiwan, which established both a loan and investment fund focused on Central and Eastern Europe of approximately €190 million and €1 billion respectively[64], the only response from the EU was to allow Lithuania to provide state aid from its own finances (Box 2) and to file a complaint to the WTO[65].

This failed to send a message of European unity, nor did it create a precedent that could serve as deterrence against future economic coercion.

Therefore, it is unlikely that any additional support scheme could be introduced in the absence of further progress on aligning foreign policy.

7 Conclusion

The rise in global geopolitical tensions has coincided with deeper economic integration of EU and non-democratic countries, and an increase in the market concentration of EU imports. While the EU benefits from this trade in many ways, the links have also created economic security risks beyond traditional trade wars.

To counter these risks, the EU should invest in a deeper understanding of its supply chains and pursue targeted industrial policies in a small number of carefully selected industries of strategic importance.

However, the depth of exposure to economic coercion and other shocks stems from structurally more concentrated imports and exports. Unless the EU manages to diversify its trading relationships, many products will remain exposed.

While it is difficult to inflict macroeconomically-relevant harm through economic coercion alone, there are many products over which pressure could be applied on politically important constituencies.

Therefore, the EU should invest in ex-post policies that mitigate economic harm where it occurs. Such policies, taken together with deterrence through the threat of defensive measures under the ACI, would disincentivise the use of economic coercion against the EU.

However, for ex-post policies to be effective, a more common foreign policy is necessary, as otherwise common burden-sharing and unified responses are not credible. ■

Niclas Poitiers is a Research Fellow, and Conor McCaffrey is a Research Analyst, at Bruegel

Endnotes

1. Jean Pisani-Ferry, 'The Geopolitical Conquest of Economics', Project Syndicate, 30 September 2021.

2. For a discussion see Hillman (2022).

3. The G7 wants "coordinate our approach to economic resilience and economic security that is based on diversifying and deepening partnerships and de-risking, not de-coupling". See G7 Hiroshima Leaders' Communiqué of 20 May 2023.

4. "Many of you have heard the term 'small yard, high fence' when it comes to protecting critical technologies. The concept has been citied at think tanks and universities and conferences for years. We are now implementing it." Remarks by National Security Advisor Jake Sullivan on the Biden-Harris Administration's National Security Strategy on 13 October 2022.

5. We focus predominantly on China because of the documented potential exposure of EU firms to Chinese shocks; see for instance the survey results reported in Attinasi et al (2023). See Box 1 for a discussion of the US.

6. Given the potential for technological advantages to give monopolistic powers to semiconductor firms, coupled with the immense capacity for economic coercion in this sector, we believe that the 2023 export controls introduced by the Netherlands on advanced semiconductor manufacturing equipment are one of the very few instances in which technological defence measures can be justified by economic security arguments; for an English translation of the justification given by the Dutch government, see: https://csis-website-prod.s3.amazonaws.com/s3fs-public/2023-07/230721_CSISTranslations_Dutch_Export.pdf.

7. Article XXI of the General Agreement on Tariffs and Trade.

8. J Brunsden, S Fleming, A Williams and J Politi, 'EU and US end Airbus-Boeing trade dispute after 17 years', Financial Times, 15 June 2021.

9. See European Commission Questions and Answers of 10 July 2023, 'EU-US Data Privacy Framework'.

10. See https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/stronger-europe-world/eu-us-trade-and-technology-council_en.

11. See The White House, 'G7 Leaders' Statement on Economic Resilience and Economic Security', 20 May 2023.

12. See The White House, 'Joint Statement by President Biden and President von der Leyen', 10 March 2023.
13. See European Commission, 'Speech by President von der Leyen on EU-China relations to the Mercator Institute for China Studies and the European Policy Centre', 30 March 2023.
14. See The White House, 'FACT SHEET: CHIPS and Science Act Will Lower Costs, Create Jobs, Strengthen Supply Chains, and Counter China', 9 August 2022.
15. "It's important to recognize this is a national security action, not an economic one … This executive order protects our national security interests … Again, I want to be clear: This is a national security action, not an economic one." The White House, 'Background Press Call by Senior Administration Officials Previewing Executive Order on Addressing U.S. Investments in Certain National Security Technologies and Products in Countries of Concern', 10 August 2023.
16. See European Commission press release of 10 September 2019, 'The von der Leyen Commission: for a Union that strives for more'.
17. See German Association of the Automotive Industry, https://www.vda.de/en/news/facts-and-figures/annual-figures/exports.
18. "Beijing frequently restricts trade by targeting imports of agricultural goods or commodities. Only on rare occasions has it employed or threatened to employ export restrictions, as was the case with rare earths to Japan in 2010" (Adachi et al 2022).
19. See Lim and Ferguson (2019) for a discussion of the use of boycotts by China during the dispute with South Korea regarding the THADD missile defence programme.
20. The Economist, 'How rare-gas supply adapted to Russia's war', 30 March 2023.
21. Some analysis has also pointed to Australia deciding to exclude Huawei from 5G infrastructure as a cause for the Chinese response; see Hackenbroich et al (2022).
22. The justifications given for these different de-facto import embargoes were both imaginative and spurious. For instance, mandatory testing for traces of heavy metal was introduced for the import of crustaceans, with the testing period long enough that live lobster exports could not survive the process (Buckland et al 2023).

23. The value of Australian exports to China did grow slightly over this period, because of an increase in the price of iron ores, a key input into the Chinese economy and overwhelmingly the largest component of Australian exports to China- averaging over 50 percent of monthly bilateral exports in 2019.
24. As detailed by Urden (2023b): "China started buying coal from Indonesia, which then cut its sales to India and elsewhere. India boosted its purchases of Australian coal that had previously gone to China". Japan and Korea also massively increased their purchases of Australian coal over this period. This also coincided with energy shortages following the Russian invasion of Ukraine, which meant that coal prices increased significantly.
25. See Casey Hall and Xiaoyu Yin, 'China's wine market ready to welcome likely return of Aussie wine as ties improve', Reuters, 3 November 2023.
26. Reports estimate it at two billion litres, see for example Rabobank news release, '"Swimming in wine" – navigating oversupply in Australia's wine industry'.
27. UN Comtrade data shows that Australian wine imports actually increased steadily each year between 2019 and 2022, which seems to suggest limits on the wine industries' ability to diversify into the domestic market.
28. For instance, in May 2021, Lithuania became the first country to withdraw from the China-CEEC initiative.
29. The standard practice to avoid Chinese disapproval has been to allow institutions that represent the city of Taipei, not Taiwan. For more details on the actions undertaken by Lithuania, see Andrijuaskas (2022).
30. See Foreign Minister Gabrielius Landsbergis' comments in Milda Seputyte and Natalia Drozdiak, 'Lithuania Says Businesses Remain Wary on China Trade', Bloomberg, 28 November 2023.
31. Dominique Patton and Andrias Sytas, 'China suspends Lithuanian beef, dairy, beer imports as Taiwan row grows', Reuters, 10 February 2022.
32. The decrease in Lithuanian imports and exports visible from late 2022 onwards was accounted for largely by the economic slowdown in trading partners and was unrelated to the Chinese actions.
33. This is unsurprising given the well-documented dominance of China in this supply chain.
34. The experiences also match those of South Korea during the THAAD dispute of 2016-17 (Lim and Ferguson, 2019).

35. Gabija Sabaliauskaitė, '„Invega“ stabdo paskolas nukentėjusiems nuo Kinijos veiksmų: iš 130 mln. Eur paskolų suteikta už 3 mln. Eur', Verslo žinios, 6 February 2022.
36. We focus on individual goods categories rather than market values, as harm to an individual industry might come even from a low value if an indispensable import is affected.
37. See Mejean and Rousseaux (2024) for both a more detailed discussion of how to identify dependencies and a more comprehensive data exercise.
38. For a discussion, see Le Mouel and Poitiers (2023).
39. Mai Nguyen and Eric Onstad, 'China's rare earths dominance in focus after it limits germanium and gallium exports', Reuters, 21 December 2023.
40. For a discussion of the role of security concerns with regards to the US, see Box 1.
41. Due to capacity constraints, we do not consider here general policy measures to improve the single market, even if these measures could improve the competitiveness of European firms, thus likely contributing to the economic security of the EU. For a discussion on these measures, see Kleimann et al (2023).
42. Some have attempted to argue that potentially losing current comparative advantages in critical technologies constitute a threat to economic security, given that it may result in future import dependencies. In our view, this is currently too many degrees removed to fall under economic security concerns.
43. As discussed previously, there are some rare instances involving technology-induced monopoly that legitimise the use of technology security tools to maintain economic security.
44. See Sultan et al (2023), for example.
45. Formerly called the Single Market Emergency Instrument (SMEI).
46. See European Commission, 'European Chips Act: Monitoring and crisis response', undated.
47. If intra-EU trade is also included, the average of EU countries' exports covered by reciprocal PTAs was 81 percent.
48. For more information on the Advanced Framework Agreement and Interim Trade Agreement, see European Commission press release of 13 December 2023, 'EU and Chile sign modern and ambitious trade and political agreements'.

49. Negotiations between the EU and the Mercosur states on a deal began in 2000 and only concluded with an agreement in June 2019. Five years later, EU ratification is still awaited. The October 2023 breakdown in EU-Australian trade agreement negotiations also fails to bode well for the prospect of new deals on the horizon.
50. Article 5 of Regulation (EU) 2021/947 establishing the Neighbourhood, Development and International Cooperation Instrument (NDICI, the EU's primary international development tool) states that the EU should "seek to promote increased synergies and complementarities" between trade policy and sustainable development". See https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A32021R0947.
51. European Commission press release of 26 October 2023, 'Global Gateway: EU signs strategic partnerships on critical raw materials value chains with DRC and Zambia and advances cooperation with US and other key partners to develop the "Lobito Corridor"'.
52. The Commission has raised concerns that national ECAs "do not follow overarching EU interest and policies… and can be also in competition with one another" (European Commission 2023b, p.7). It also argued that better coordination between national ECAs and EU and national development finance agencies could lead to better outcomes across a range of policy areas, including the sourcing of CRMs and "the trade aspects of EU geopolitical strategies" (European Commission, 2023b, p. 39).
53. Source: Flash Eurobarometer 537 (2023); Firms would likely be more aware of their local branches of the EEN, such as national export promotion offices.
54. Gabriella Beaumont-Smith, 'Rock-a-Bye Trade Restrictions on Baby Formula', Cato at Liberty, 10 May 2022.
55. This already at a time when concerns are growing over single market fragmentation due to the relaxing of state aid rules following Russia's invasion of Ukraine; see Théo Bourgery-Gonse, 'EU subsidy race is on – and Germany is winning it', Euractiv, 12 September 2023.
56. See Reuters, 'China export curbs choke off shipments of gallium, germanium for second month', 20 October 2023.
57. European Commission, 'Questions & Answers regarding the Anti-Coercion Instrument', undated.

58. This was a feature of Chinese measures against Lithuania, as it sought to pressure German industry to intervene. See for instance Andrius Sytas and John O'Donnell, 'German big business piles pressure on Lithuania in China row', Reuters, 21 January 2022.
59. This same moral hazard applies to the ACI, as discussed in Hackenbroich et al (2022).
60. To reduce the administrative burden, we would propose limiting these additional requirements to larger firms, with SMEs covered regardless.
61. The lack of uptake of the Lithuanian support scheme warrants consideration.
62. See European Commission press release of 26 April 2022, 'State aid: Commission approves €130 million Lithuanian scheme to support companies affected by discriminatory trade restrictions'.
63. Andrius Sytas, 'Lithuania to get U.S. trade support as it faces China fury over Taiwan', Reuters, 19 November 2021.
64. Giedre Peseckyte, 'Taiwan encourages companies to invest in Lithuania to deepen bilateral cooperation', Euractiv, 3 October 2023.
65. European Commission press release of 7 December 2022, 'EU requests two WTO panels against China: trade restrictions on Lithuania and high-tech patents'.

References

Adachi, A, A Brown and M Zenglein (2022) 'Fasten your seatbelts: How to manage China's economic coercion', MERICS China Monitor, 25 August, Mercator Institute for China Studies.

Aiyar, S, J Chen, CH Ebeke, R Garcia-Saltos, T Gudmundsson, A Ilyina … JP Trevino (2023) 'Geoeconomic Fragmentation and the Future of Multilateralism', Staff Discussion Note SDN/2023/001, International Monetary Fund.

Andrijauskas, K (2022) 'An Analysis of China's Economic Coercion Against Lithuania', mimeo, Council on Foreign Relations.

Attinasi, MG, D Ioannou, L Lebastard and R Morris (2023) 'Global production and supply chain risks: insights from a survey of leading companies', in ECB Economic Bulletin 7/2023, European Central Bank.

Baqaee, DJ Hinz, B Moll, M Schularick, FA Teti, J Wanner and S Yang (2024) 'What if? The Effects of a Hard Decoupling from China on the German Economy', chapter 4 in J Pisani-Ferry, B Weder Di Mauro and J Zettelmeyer (eds) Paris Report 2: Europe's Economic Security, CEPR Press.
Babina, T, B Hilgenstock, O Itskhoki, M Mironov and E Ribakova, (2023) 'Assessing the Impact of International Sanctions on Russian Oil Exports', mimeo.
Buckland, A, A Duncan, M Farhad, A Hailemariam, D Kiely, V Sanchez Arenas and P Sotiraakopoulos (2023) 'Trading Up: International trade features and the Western Australian economy', Focus on Industry Report Series 8, Bankwest Curtin Economics Centre.
Dadush, U (2021) 'What to make of the EU-US deal on steel and aluminium?' Bruegel Blog, 4 November.
Dadush, U and E Dominguez Prost (2023) 'The problem with preferential trade agreements', Analysis, Bruegel.
Darvas, Z, C Martins, L Moffat and N Poitiers (2023) 'EU trade and investment relations with Ukraine and with Russia before Russia's illegal invasion of Ukraine in 2022', Briefing requested by the INTA Committee, European Parliament.
Department of Commerce (2018) The Effect of Imports of Steel on the National Security, US Department of Commerce.
European Commission (2021a) 'Strategic dependencies and capacities', Commission Staff Working Document, European Commission May 2021.
European Commission (2021b) 'The Global Gateway', JOIN(2021) 30 final.
European Commission (2022) 'State Aid SA.102002 (2022/N) – Lithuania; The Incentive Financial Instrument: Direct loans to business entities affected by third party actions', C(2022) 2846 final.
European Commission (2023a) 'European Economic Security Strategy', JOIN(2023) 20 final.
European Commission (2023b) 'Main outcomes of the mapping of external financial tools of the EU', Joint Staff Working Document, SWD(2023) 96.
European Commission (2023c) 'Feasibility Study on an EU Strategy on Export Credits', Final Report, DG TRADE.
European Commission (2023d) 'Commission recommendation on critical technology areas for the EU's economic security for further risk assessment with Member States', C(2023) 6689 final.

European Investment Bank (2023) EIB Investment Survey 2023.
Fajgelbaum, P, PK Goldberg, PJ Kennedy, A Khandelwal and D Taglioni (2023) 'The US-China Trade War and Global Reallocations', NBER Working Paper No. 29562, National Bureau of Economic Research.
Farrell, H and AL Newman (2019) 'Weaponized Interdependence: How Global Economic Networks Shape State Coercion', International Security 44 (1): 42–79.
Felbermayr, G, A Kirilakha, C Syropoulos, E Yalcin and YV Yotov (2020) 'The global sanctions data base', European Economic Review 129: 103561.
Galdin, A (2023) 'Resilience of Global Supply Chains and Generic Drug Shortages', mimeo.
García-Herrero, A, H Grabbe and A Kaellenius (2023) 'De-risking and decarbonising: a green tech partnership to reduce reliance on China', Policy Brief 19/2023, Bruegel.
Georgitzikis, K and E D'Elia (2022) 'Rare Gases (Krypton, Neon, Xenon): Impact assessment for supply security', Science for Policy Brief, Joint Research Centre, European Commission.
Hackenbroich, J, F Medunic and P Zerka (2022) 'Tough trade: The hidden costs of economic coercion', Policy Brief, February, European Council on Foreign Relations.
Hansen, R (2023) 'Date with history: How the oil crisis changed the world', The World Today, Chatham House, 29 September.
Harrell, P, E Rosenberg and E Saravalle (2018) China's use of coercive economic measures, Center for a New American Security.
Hillman, J (2022) 'China's Entry into the WTO - A Mistake by the United States?' mimeo.
Hufbauer, G, J Schott, K Elliott and B Oegg (2007) Economic sanctions reconsidered, Peterson Institute for International Economics.
Jaravel, X and I Mejean (2021) 'A data-driven resilience strategy in a globalized world,' Notes du CAE 64: 1-12, Conseil d'Analyse Economique.
Javorcik, B, L Kitzmueller, H Schweiger and MA Yıldırım (2022) 'Economic Costs of Friend-Shoring', Working Paper 274, European Bank of Reconstruction and Development.

Kleimann, D, N Poitiers, A Sapir, S Tagliapietra, N Véron, R Veugelers and J Zettelmeyer (2023) 'How Europe should answer the US Inflation Reduction Act', Policy Contribution 04/2023, Bruegel.
Kleinhans, J and N Baisakova (2020) 'The Global Semiconductor Value Chain: A Technology Primer for Policy Makers', Policy Brief, Stiftung Neue Verantwortung.
Le Mouel, M and N Poitiers (2023) 'Why Europe's critical raw materials strategy has to be international', Analysis, Bruegel.
Leonard, MJ Pisani-Ferry, E Ribakova, J Shapiro and G Wolff (2019) 'Redefining Europe's economic sovereignty', Policy Brief 09/2019, Bruegel.
Lim, DJ and V Ferguson (2019) 'Chinese Economic Coercion during the THAAD Dispute', The ASAN Forum, 28 December.
Maruyama, W and A Wolff (2023) 'Saving the WTO from the national security exception', PIIE Working Paper 23-2, Peterson Institute for International Economics.
McVicar, D (2023) 'How the Czech Republic Became One of Taiwan's Closest European Partners and What It Means for EU-China Relations', Blog Post, 24 April, Council on Foreign Relations.
McWilliams, B, S Tagliapietra and C Trasi (2024) 'Smarter European Union industrial policy for solar panels', Policy Brief 02/2024, Bruegel.
Mejean, I and P Rousseaux (2024) 'Identifying European trade dependencies', chapter 3 in J Pisani-Ferry, B Weder Di Mauro and J Zettelmeyer (eds) Paris Report 2: Europe's Economic Security, CEPR Press.
Mulder, N (2022) The economic weapon: The rise of sanctions as a tool of modern war, Yale University Press.
Poitiers, N and P Weil (2021) 'A new direction for the European Union's half-hearted semiconductor strategy', Policy Contribution 17/2021, Bruegel.
Qiu, H, H Song Shin and L Zhang (2023) 'Mapping the realignment of global value chains', Bank of International Settlements Bulletin 78.
Ragonnaud, G (2024) 'Internal Market Emergency and Resilience Act (IMERA)', Briefing – EU Legislation in Progress, European Parliamentary Research Service, PE 739.338.
Rizzi, A and A Varvelli (2023) 'Opening the Global Gateway: why the EU should invest more in the Southern Neighbourhood', Policy Brief 03/2023, European Council on Foreign Relations.

Schlögl, L, D Pfaffenbichler and W Raza (2023) 'Aligning European export credit agencies with EU policy goals', In-Depth Analysis for the European Parliament Directorate-General for External Policies.
Sultan, S, C Rusche, H Bardt and B Busch (2023) 'Single Market Emergency Instrument – A Tool with Pitfalls', Intereconomics 58(3): 160-166
Tagliapietra, S, R Veugelers and J Zettelmeyer (2023) 'Rebooting the European Union's Net Zero Industry Act', Policy Brief 06/2023, Bruegel.
Uren, D (2023a) 'Why China's coercion of Australia failed', The Strategist, 27 April, The Australian Strategic Policy Institute.
Uren, D (2023b) 'Despite the risks, Australian exports to China are booming again', The Strategist, 22 August, The Australian Strategic Policy Institute.
US Department of Justice (2010) Horizontal Merger Guidelines, US Department of Justice and the Federal Trade Commission.
Vicard, V and P Wibaux (2023) 'EU Strategic Dependencies: A Long View', Policy Brief 2023-41, CEPII.
Welslau, L and G Zachmann (2023) 'Is Europe failing on import diversification?' Bruegel Blog, 20 February.

This article is based on a Bruegel Working Paper 12/2024, that is based on McCaffrey, C and N Poitiers (2024) 'Instruments of economic security', chapter 6 in Europe's Economic Security, Paris Report 2, CEPR and Bruegel. We would like to thank Alicia García-Herrero, André Sapir, Beatrice Weder di Mauro, Chad Bown, Heather Grabbe, Guntram Wolff, Jean Pisani-Ferry and Jeromin Zettelmeyer for their comments and discussion and Luca Léry Moffat for excellent research assistance.

Advancing China's economic growth

China is approaching a fork in the road: rely on the policies that have worked in the past or update its policies for a new era of high-quality growth. Kristalina Georgieva discusses

The global economy

For the world, the year ahead will require careful calibration of monetary and fiscal policies to secure a soft landing — bringing inflation down while maintaining growth firmly in positive territory. Many central banks have the difficult task of deciding when to cut interest rates and by how much, based on data. They can no longer take cues from others as both the pace of disinflation and growth are diverging across countries.

It will be also a challenging year for fiscal authorities in most countries — they need to embrace consolidation to reduce debt and rebuild buffers, and at the same time finance the digital and green transformations of their economies.

Fortunately, the global economy has proven to be remarkably resilient to the shocks of the last years. This resilience is mostly due to strong macroeconomic fundamentals in most of the advanced and emerging market economies and robust consumer and government spending. Labour markets have held up and supply chains have normalized.

Therefore, despite the higher global interest rates, we project over 3 percent growth this year and next. While inflation remains above target in many countries, we see it continuing to fall. The picture here in Asia is more nuanced, because inflation did not rise as much as elsewhere, and it is coming down faster. As a result, interest rates have not risen as much.

But looking to the medium term, we expect global growth to be around 3 percent, which is weak by historical standards – during the pre-COVID decade the annual average was 3.8 percent. Low productivity growth and high debt levels are posing challenges to all, but especially to emerging and developing economies. And geopolitical tensions affect trade and capital flows, which have been essential engines of growth over the last decades.

The good news is that the digital and green transformations present opportunities to boost productivity growth and improve living standards. Deep structural reforms can enhance the conditions for entrepreneurship, innovation and economic performance.

Our analysis shows that decisive steps to reduce the stock of unfinished housing and giving more space for market-based corrections in the property sector could both accelerate the solution to the current property sector problems and lift up consumer and investor confidence

China – a new era of high-quality growth

Zooming in on China, we saw a strong post-COVID rebound in 2023, with growth exceeding five percent. In the medium-term, China will continue to be a key contributor to global economic growth. While low productivity growth and an aging population are factors affecting growth, there are also tremendous opportunities.

China is poised to face a fork in the road — rely on the policies that have worked in the past or update its policies for a new era of high-quality growth.

According to our analysis, with a comprehensive package of pro-market reforms, China could grow considerably faster than a status quo scenario. This additional growth would amount to a 20 percent expansion of the real economy over the next 15 years—in today's terms, that is like adding US$ 3.5 trillion to the Chinese economy.

What are the ingredients of such a package of reforms? It all starts with sound macroeconomic fundamentals. I was very encouraged to hear the commitment to sound fundamentals and strong institutions here in China.

Decades of impressive growth in China have significantly improved living standards and provided ample policy buffers to address its most-pressing near-term challenges. These include transitioning the property sector to a more sustainable footing and reducing local government debt risks. Tackling these challenges is essential for a smooth transition to a new era of high-quality growth.

Our analysis shows that decisive steps to reduce the stock of unfinished housing and giving more space for market-based corrections in the property sector could both accelerate the solution to the current property sector problems and lift up consumer and investor confidence.

A key feature of high-quality growth will need to be higher reliance on domestic consumption. Doing so depends on boosting the spending power of individuals and families. China's social security system covers more people than any other on the planet. But there is room to expand its reach further and increase benefits—think of strengthening the pension system in a fiscally responsible way.

Domestic consumption also depends on income growth, which in turn relies on the productivity of capital and labour. Reforms such as strengthening the business environment and ensuring a level playing field between private and state-owned enterprises will improve the allocation of capital. Investments in human capital — in education, life-long training and reskilling – and quality health care will deliver higher labour productivity and higher incomes.

This is particularly important as China seeks to seize the opportunities of the AI 'big bang'. Countries' preparedness for the world of artificial intelligence is no longer a goal for the future — it is already an issue for today. The IMF has identified four areas that are critical for countries' AI preparedness — digital infrastructure, human capital and labour markets, innovation, and regulation and ethics.

Our analysis finds that China is at the forefront of emerging economies in terms of AI preparedness, with well-developed digital infrastructure providing a head start. Establishing a robust AI regulatory framework and strengthening economic ties with other innovative countries will help China power ahead.

Similarly, China has enormous potential in advancing the green economy. It is already the global leader in deploying renewable energy, and is making rapid progress in green mobility. Its continued leadership is vital to addressing the global climate crisis. Building on progress in recent years to sell a greater share of electricity at market prices would make China's decarbonization even more efficient. So, too, would extending the coverage of the emissions trading system to the industrial sector.

The transformation ahead is not easy. China's remarkable development success has delivered tremendous benefits to hundreds of millions of people. The younger generations, who have lived their whole lives in an environment of exceptionally high growth rates, are experiencing what many countries have experienced before as economies mature and growth moderates.

But this transition from high rates to high quality of growth is the right fork in the road to take and China is determined to do so. As the government recognizes, high-quality development ultimately depends on reforms. In this endeavour, the IMF is committed to being a partner, including through our ongoing policy dialogue and mutual learning. And also to working together to tackle global challenges such as fragmentation, climate change, and debt.

Working together delivers for all

International cooperation in our interconnected world is essential to solving these challenges—which we know have an outsized impact on the most vulnerable members of our global community. The world comes together at the IMF to tackle problems, and we appreciate China's continued support for our efforts.

China has helped strengthen the IMF's financial capacity through contributions to our concessional lending instrument for low-income countries, our recently created Resilience and Sustainability Facility, and our capacity development initiatives. China has shown remarkable leadership in helping forge the agreement to increase the IMF's permanent resources by 50 percent.

We also recognize China's important role in addressing debt distress in emerging and developing economies. With many countries at or near debt distress much work is needed among creditors to speed up debt relief and we look forward to China's continued strong engagement.

With the dynamism, confidence, and luck of the dragon—and a renewed spirit of international cooperation—China and the world can rise together to the challenges we face today for a more prosperous future for everyone. ■

Kristalina Georgieva is Managing Director of the International Monetary Fund

This article is based on remarks at the China Development Forum, Beijing, March 24, 2024.

Planning a New Career Challenge?

Unique Access to Confidential Opportunities

InterExec is the global leader in assisting Top Executives
to access £200k to £2m+ unadvertised vacancies worldwide.

We act discreetly through our 15,000 strong Headhunter network.

InterExec
UNIQUE NETWORK ♦ OUTSTANDING TALENT

london@interexec.net www.interexec.net +44 (0)20 7256 5085

Global economic fracturing and shifting investment patterns

There is a risk of economic deglobalisation. Bruno Casella, Richard Bolwijn and Francesco Casalena highlight ten FDI trends and their development implications

The trends stemming from technological advances, policy shifts, and the sustainability imperative, along with the effects of the trade tensions starting in late 2017 and the shock of the COVID pandemic, have sparked a debate on the risk of a reversal of economic globalisation (Fajgelbaum *et al* 2020, Antràs 2020, Zhan *et al* 2020, Kukharskyy *et al* 2021, Baldwin 2022a, 2022b, 2022c, 2022d).

The subsequent shocks of conflicts and political fragmentation have brought to the fore the heightened pressure towards global economic fracturing and the decoupling of global value chain (GVC) links between the US and other developed economies and the Chinese economy, with implications for many other countries and regions (Campos *et al* 2023, Aiyar *et al* 2023, Javorcik *et al* 2023).

So far, this debate has mainly focused on the trade perspective (Aiyar *et al* 2023 is an exception). The objective of this study is to explore the investment angle, offering a comprehensive reference for policymakers and analysts on the main trends reshaping the global FDI landscape amidst global economic fracturing. Given the intertwined nature of trade and FDI in the global production landscape dominated by GVCs, it also aims to build a much-needed bridge between connected narratives in the FDI and trade areas.

The underlying analysis owes credit to, and is directionally consistent with, previous studies investigating specific aspects of the FDI trends, particularly various recent editions of UNCTAD's World Investment Reports (eg. UNCTAD 2017, 2020, 2021). However, to date, a fully integrated diagnostic covering both short- and long-term perspectives, as well as the sectoral, geographical, and bilateral dimensions of FDI patterns, has been lacking.

This column highlights ten empirical FDI trends, grouped into three overarching themes: the triple divergence, the rise of economic fracturing and the implications for sustainability and development (UNCTAD 2024). These trends

fundamentally alter the development paradigm based on promoting investment in manufacturing and export-led growth, as will be discussed in the concluding part.

Triple divergence

Over the past two decades, FDI patterns have adapted to the transformative shifts reshaping economic globalisation in three key aspects.

Since the escalation of the trade war – with an acceleration after the outbreak of the pandemic and the recent geopolitical crisis – escalating international tensions are turning divergence into fracturing

1. Divergence between trends in FDI and GVCs and trends in GDP and trade. Historically closely intertwined under the common shaping force of GVCs, global trends in FDI and GVCs and in GDP and trade have been growing apart since the 2010s. While global GDP and trade have continued to grow steadily, crossborder investment and GVCs are coping with a long-term stagnation (FDI trend #1).

2. Divergence in FDI trends between services and manufacturing. FDI's long-term stagnation is characterised by starkly divergent trajectories between rapidly growing investment in services (FDI trend #2) and shrinking investment in manufacturing activities (FDI trend #3) (Figure 1).

The transition from manufacturing to services is part of a broader change in the role of FDI in global value creation, whereas crossborder investment is moving from the centre to the two ends of the smile curve (FDI trend #4). This major shift is involving developed and developing economies alike, blurring the traditional boundaries in terms of their FDI sectoral footprints (FDI trend #5).

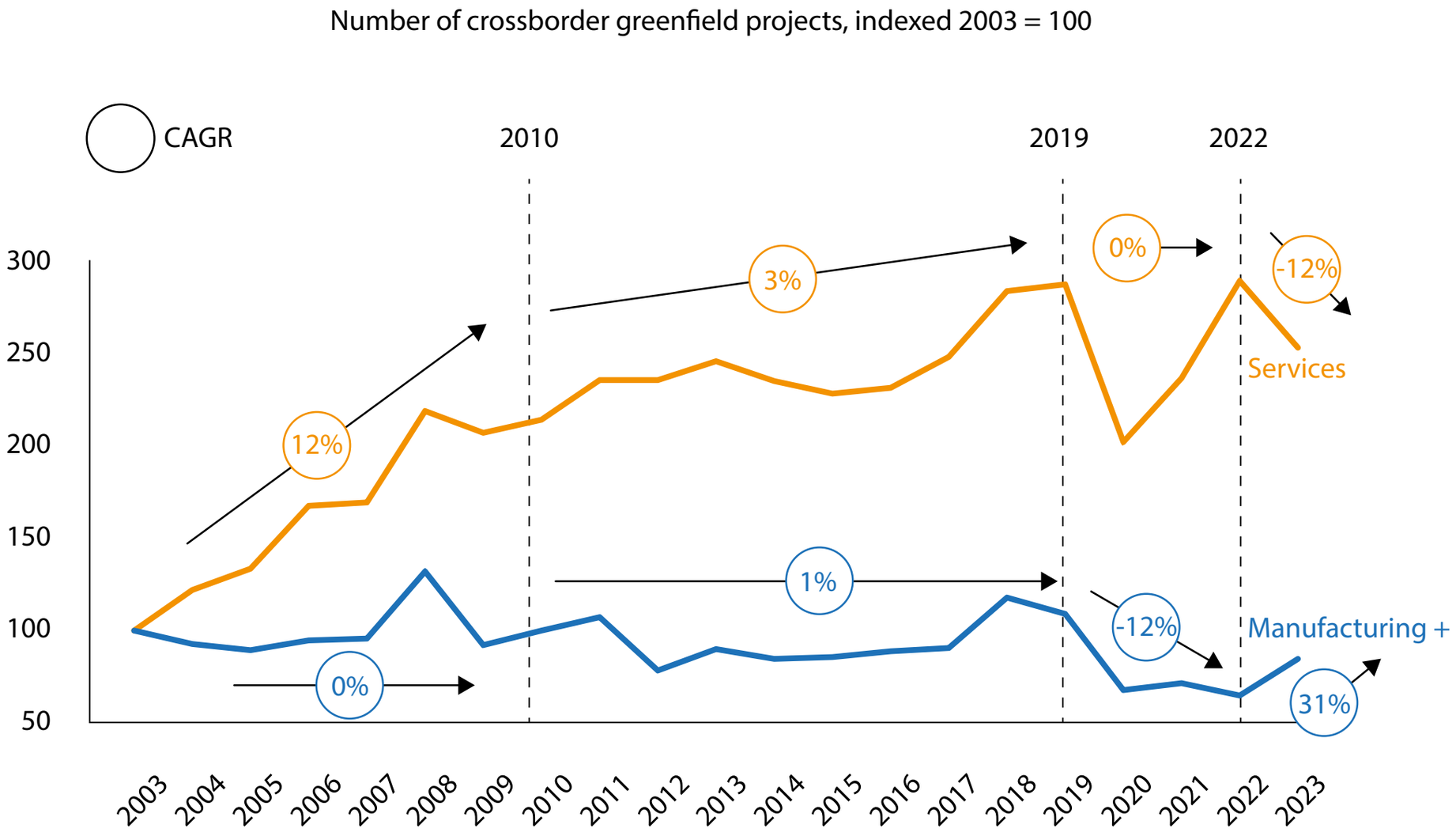
3. Divergence in FDI trends between China and the rest of the world. Chinese share in crossborder greenfield projects has been consistently declining for two decades, with an acceleration after the pandemic.

Despite a waning interest from multinational corporations in initiating new investment projects in China, the country continues to maintain a dominant position in global manufacturing and trade. Far from downsizing, 'Global Factory China' is changing its operational model from globally integrated to more domestically focused production networks, while still maintaining its leadership in global trade (FDI trend #6).

From divergence to fracturing

Since the escalation of the trade war – with an acceleration after the outbreak of the pandemic and the recent geopolitical crisis – escalating international tensions are turning divergence into fracturing, leading

Figure 1. Diverging FDI trends in manufacturing and services



Note: CAGR: Compound Annual Growth Rate. The sectoral analysis is based on the variable 'Business Activity' from fDi Markets. 'Manufacturing+' includes 'Manufacturing' and 'Other non-services' activities. The latter group comprises the following categories: construction, electricity, extraction and infrastructure.
Source: UNCTAD, based on information from the Financial Times Ltd, fDi Markets (www.fDimarkets.com).

to the disruption of historical investment patterns. Fracturing is associated with heightened uncertainty and unpredictability in the FDI landscape, and limited possibilities for countries to strategically benefit from diversification (FDI trend #7).

The fracturing process is characterised by the rising importance of geopolitics. Overall, between 2013 and 2022, the share of FDI projects between geopolitically distant countries decreased by 10 percentage points, from 23% to 13% (figure 2). Geopolitical motivations are thus emerging as primary drivers of investment decisions, at times overriding traditional FDI determinants (FDI trend #8).

Sustainability push, but marginalisation of developing countries

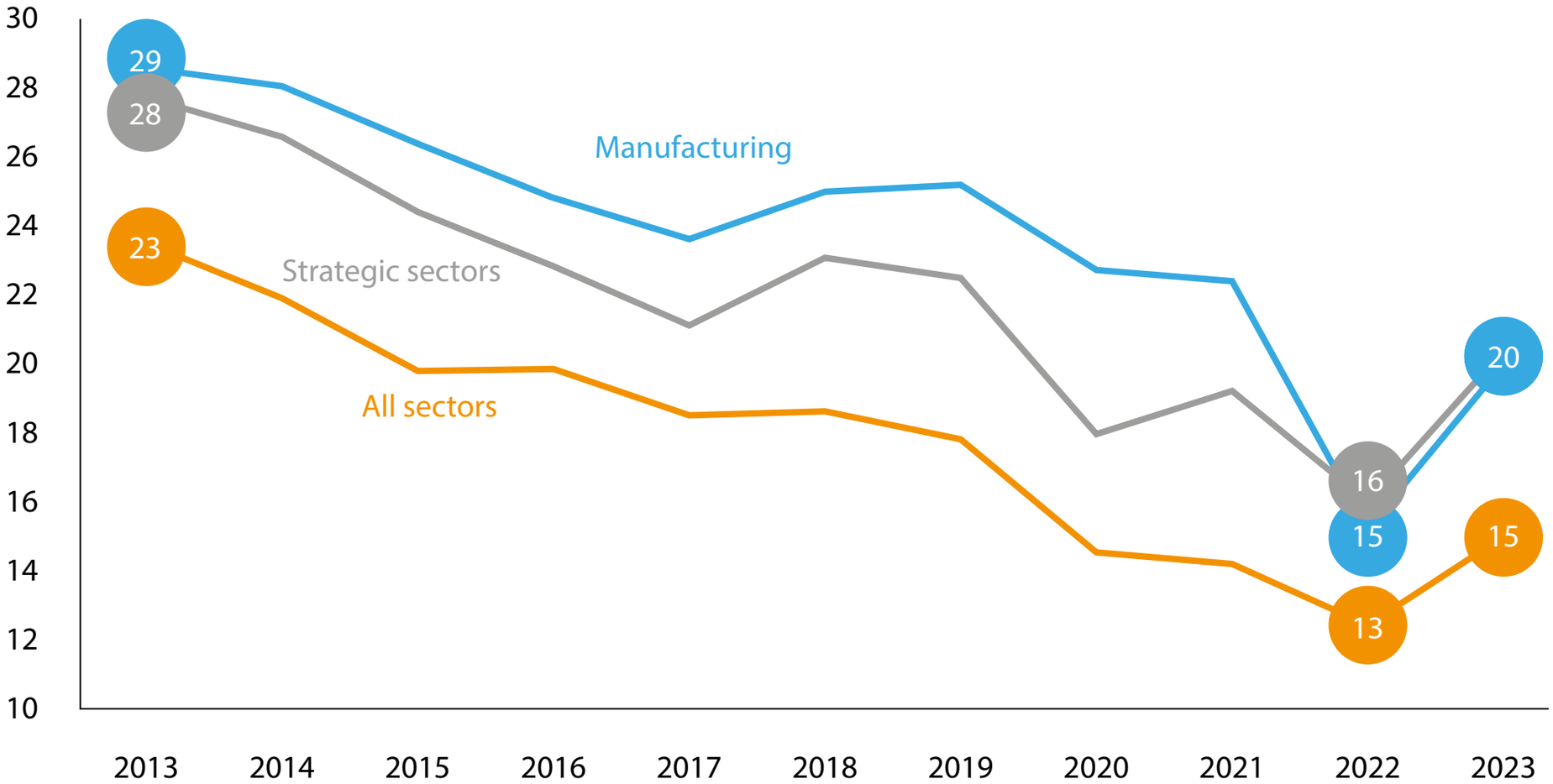
Amid long-term stagnation of manufacturing investment across all industries, the number of crossborder greenfield projects in renewable energy generation (environmental technologies) as well as in the manufacturing of batteries and electric vehicles (EVs) has steadily increased (Figure 3).

The sustainability imperative and the drive to stimulate investment in the Sustainable Development Goals (SDGs) have opened new opportunities for investment in industrial development (FDI trend #9). However, these new opportunities can only compensate in part for the lack of FDI growth in other industrial sectors that are critical for GVC development strategies.

Historical shifts and economic fracturing are leading to a decrease in the share of FDI in smaller developing countries and least developed countries. This trend exacerbates their marginalisation and vulnerability, as FDI becomes increasingly concentrated in developed and emerging economies (FDI trend #10).

Figure 2. Declining share of FDI between geopolitically distant countries

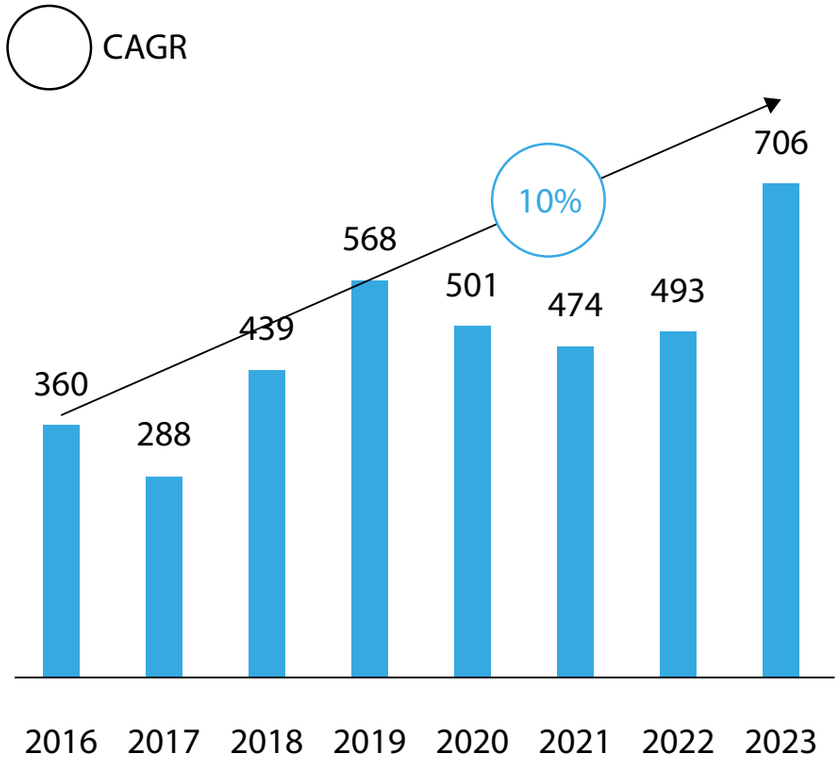
Crossborder greenfield projects between geopolitically distant countries as a share of total, per cent



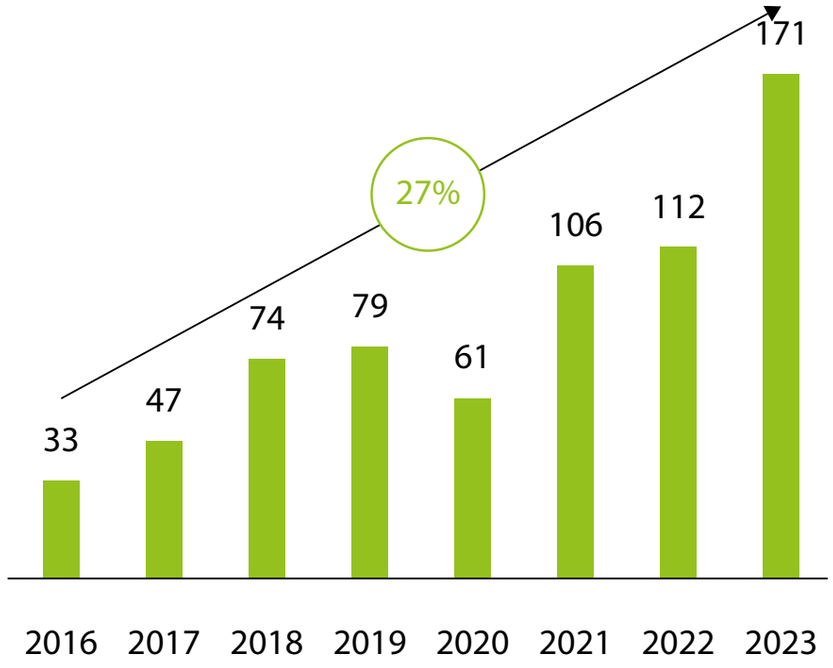
Note: The classification 'Strategic sectors' follows IMF (2023). Assessment of geopolitical alignment is based on United Nations voting patterns (Bailey et al 2017). The findings remain robust under alternative definitions of geopolitical groupings.
Source: UNCTAD, based on information from the Financial Times Ltd, fDi Markets (www.fDimarkets.com).

Figure 3. Growth of green FDI

Number of crossborder greenfield projects in environmental technologies



Number of crossborder greenfield projects in manufacturing of batteries and electric vehicles



Note: CAGR: Compound Annual Growth Rate.
Source: UNCTAD, based on information from the Financial Times Ltd, fDi Markets (www.fDimarkets.com).

Conclusions: rethinking the FDI-GVC-development nexus

Based on a diagnostic of ten trends in foreign direct investment, in this column we put forward three major implications for developing countries and their development and industrialisation strategies.

First, the long-term stagnation of investment in GVCs and the sectoral shifts in investment patterns fundamentally alter the development paradigm based on promoting investment in manufacturing and export-led growth. These shifts affect the prospects for developing countries to increase their GVC participation and to gradually upgrade to higher value-added industrial activities.

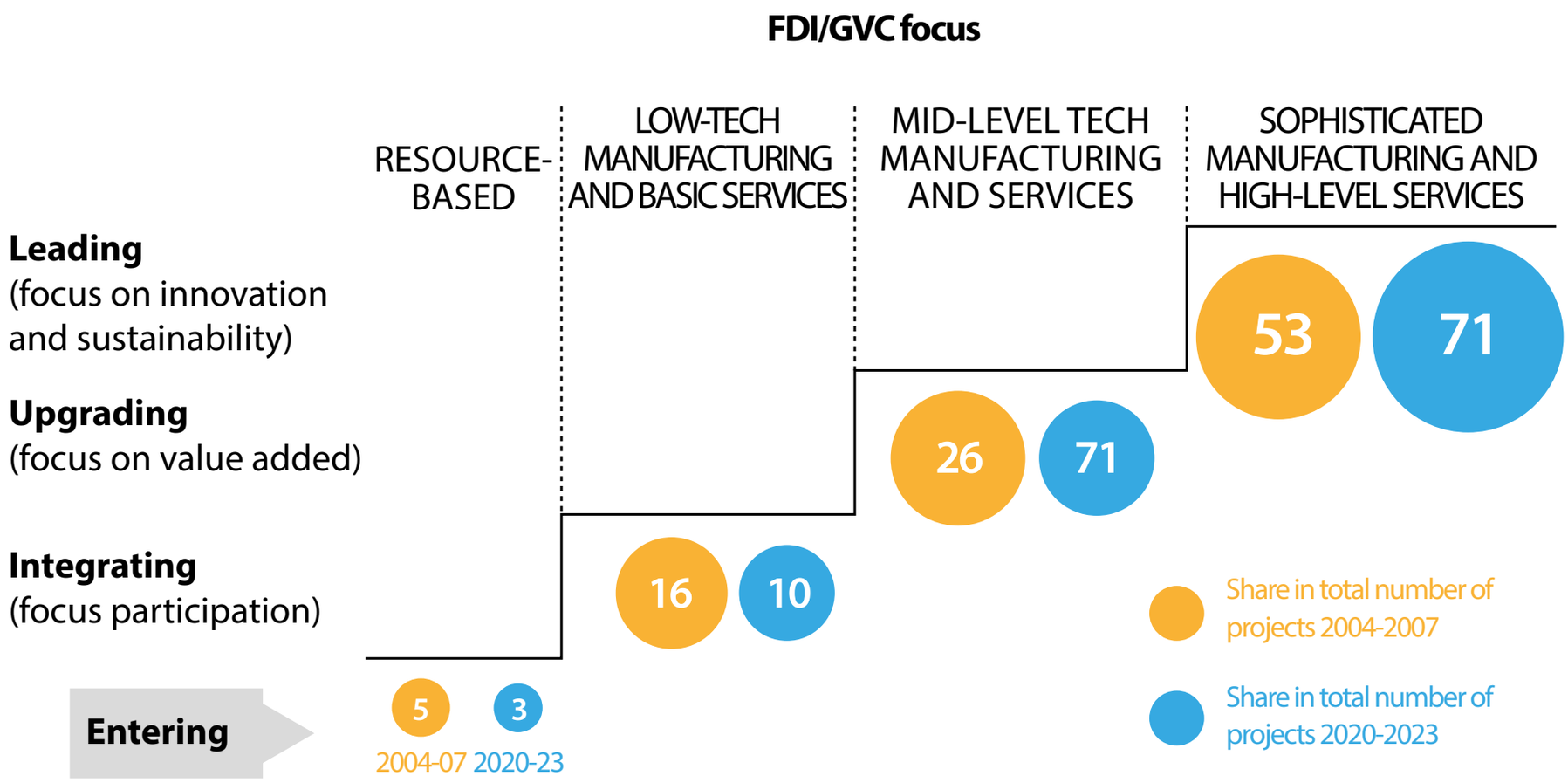
The GVC development ladder – a concept developed in UNCTAD's *World Investment Report 2013* (UNCTAD 2013) – is becoming harder to climb as the least developed countries face declining manufacturing investment and a shrinking pool of efficiency-seeking, lower value-added projects to leverage for GVC participation (Figure 4).

Second, changes in the patterns of sources and destinations of investment due to global economic fracturing, de-risking, and resilience trends can bring opportunities for some countries, but are a challenge for most. They not only reinforce the effects of the long-term trends but also introduce new complexity into international production and increased uncertainty for both investors and investment policymakers as geopolitical considerations become more important FDI determinants.

Third, the ongoing marginalisation of countries at the lower levels of the GVC development ladder, combined with diminishing opportunities in traditional GVC-intensive industries, requires investment policymakers in these countries to intensify their search for investment promotion opportunities in sectors that are less reliant on GVCs.

Figure 4. The GVC development ladder: Shifting FDI weights

Distribution of crossborder greenfield projects across stages of the GVC-development ladder, per cent



Note: The classification of projects along the ladder is based on fDi Markets variables 'Business Activity' and 'Cluster' and Lall's technological classification (Lall 2000, Sturgeon and Gereffi 2009).
Source: UNCTAD, building on the concept developed in UNCTAD World Investment Report 2013 (pages 179-181); project shares based on information from the Financial Times Ltd, fDi Markets (www.fDimarkets.com).

This includes industries where growth is driven by policy factors other than those influencing the general trend in GVCs. Notably, the promotion of investment in environmental technologies and sustainable energy serves as a notable example, albeit not the only one (UNCTAD 2023). ■

Bruno Casella is Senior Economist, Investment and Enterprise Division, Richard Bolwijn is Head of Investment Research, Division on Investment and Enterprise, both at the United Nations Conference on Trade and Development, and Francesco Casalena is a PhD student at the Geneva Graduate Institute (IHEID)

References

Aiyar, S, D Malacrino and A Presbitero (2023), "Investing in Friends: The Role of Geopolitical Alignment in FDI Flows", CEPR Discussion Paper No. 18434.

Antràs, P (2020), "De-Globalisation? Global Value Chains in the Post-COVID-19 Age," CEPR Discussion Paper No. 15462.

Baldwin, R (2022a), "The peak globalization myth: Part1", VoxEU.org, 31 August.

Baldwin, R (2022b), "The peak globalization myth: Part2 – Why the goods trade ratio declined," VoxEU.org, 1 September.

Baldwin, R (2022c), "The peak globalization myth: Part3 – How global supply chains are unwinding", VoxEU.org, 2 September.

Baldwin, R (2022d), "The peak globalization myth: Part4 –Services trade did not peak," VoxEU.org, 3 September 2022.

Bailey, M, A Strezhnev and E Voeten (2017), "Estimating dynamic state preferences from United Nations voting data," Journal of Conflict Resolution 61(2): 430-456.

Blanchard, E J, A Santos-Paulino, C Trentini and E Milet (2021), "Implications of rising trade tensions for FDI projects", Transnational Corporations 28(2): 161-183.

Campos, R, JE Flores, D Furceri and J Timini (2023), "Geopolitical fragmentation and trade," VoxEU.org, 31 July.

Fajgelbaum, PD, P Goldberg, PJ Kennedy and A Khandelwal (2020), "The return to protectionism," The Quarterly Journal of Economics 135(1): 1–55.

IMF (2023), "Chapter 4: Geoeconomic fragmentation and foreign direct investment," in World Economic Outlook: A rocky recovery.

Javorcik, B, L Kitzmueller, H Schweiger and A Yıldırım (2022), "Economic costs of friend-shoring," EBRD Working Paper No. 274.

Kukharskyy, B, G Felbermayr, O Krebs and P Eppinger (2021), "Decoupling from global value chains," VoxEU.org, 7 July.

Lall, S (2000), "The Technological structure and performance of developing country manufactured exports, 1985-98," Oxford Development Studies 28(3): 337-369.

Sturgeon, TJ and G Gereffi (2009), "Measuring success in the global economy: international trade, industrial upgrading, and business function outsourcing in global value chains," Transnational Corporations 18(2): 1-36.

UNCTAD (2013), World Investment Report 2013: Global Value Chains: Investment and Trade for Development.
UNCTAD (2017), World Investment Report 2017: Investment and the Digital Economy.
UNCTAD (2020), World Investment Report 2020: International Production Beyond the Pandemic.
UNCTAD (2021), World Investment Report 2021: Investing in Sustainable Recovery.
UNCTAD (2023), World Investment Report 2023: Investing in Sustainable Energy for All.
UNCTAD (2024), Global economic fracturing and shifting investment patterns.
Zhan, J, B Casella, A Santos-Paulino and R Bolwijn (2020), "Global value chain transformation to 2030: Overall direction and policy implications", VoxEU.org, 13 August.

This article was originally published on VoxEU.org.

How geopolitics is changing trade

There has been a rise in trade restrictions. Costanza Bosone, Ernest Dautović, Michael Fidora and Giovanni Stamato explore the impact of geopolitical tensions on trade flows

Since the global financial crisis, trade has been growing more slowly than GDP, ushering in an era of 'slowbalisation' (Antràs 2021). As suggested by Baldwin (2022) and Goldberg and Reed (2023), among others, such a slowdown could be read as a natural development in global trade following its earlier fast growth.

Yet, a surge in trade restriction measures has been evident since the tariff war between the US and China (see Fajgelbaum and Khandelwal 2022) and geopolitical concerns have been heightened in the wake of Russia's invasion of Ukraine, with growing debate about the need for protectionism, near-shoring, or friend-shoring.

The impact of geopolitical distance on international trade

Rising trade tensions amid heightened uncertainty have sparked a growing literature on the implications of fragmentation of trade across geopolitical lines (Aiyar *et al* 2023, Attinasi *et al* 2023, Campos *et al* 2023, Goes and Bekker 2022).

In Bosone *et al* (2024), we present new evidence and quantify the timing and impact of geopolitical tensions in shaping trade flows over the last decade. To do so, we use the latest developments in trade gravity models. We find that geopolitics starts to significantly affect global trade only after 2018, which, timewise, is in line with the tariff war between the US and China, followed by the Russian invasion of Ukraine.

Furthermore, the analysis sheds light on the heterogeneity of the effect of geopolitical distance by groups of countries: we find compelling evidence of friend-shoring, while our estimates do not reveal the presence of near-shoring. Finally, we show that geopolitical considerations are shaping European Union trade, with a particular focus on strategic goods.

In this study, geopolitics is proxied by the geopolitical distance between country pairs (Bailey *et al* 2017). As an illustration, Figure 1 (Panel A) plots the evolution over time of the geopolitical distance between four country pairs: US-China, US-France, Germany-China, and Germany-France. This chart shows a consistently higher distance from China for both the US and Germany, as well as a further increase in that distance over recent years.

Our findings point to a redistribution of global trade flows driven by geopolitical forces, reflected in the increasing importance of geopolitical distance as a barrier to trade

Geopolitical distance is then included in a standard gravity model with a full set of fixed effects, which allow us to control for unobservable factors affecting trade. We also control for international border effects and bilateral time-varying trade cost variables, such as tariffs and a trade agreement indicator.

This approach minimises the possibility that the index of geopolitical distance captures the role of other factors that could drive trade flows. We then estimate a set of time-varying elasticities of trade flows with respect to geopolitical distance to track the evolution of the role of geopolitics from 2012 to 2022.

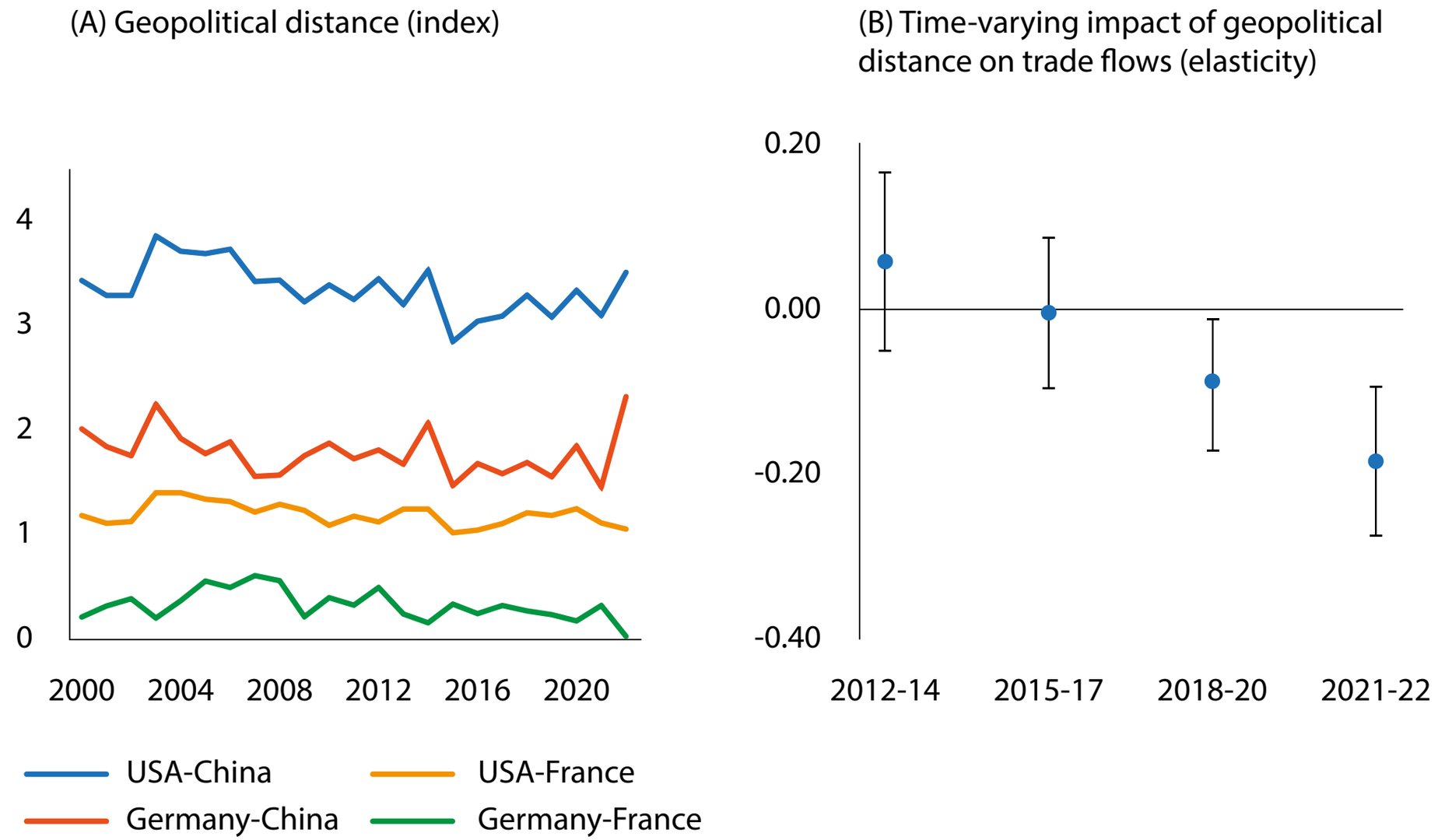
To the best of our knowledge, we cover the latest horizon on similar studies on geopolitical tensions and trade. To rule out the potential bias deriving from the use of energy flows as political leverage by opposing countries, we use manufacturing goods excluding energy as the dependent variable. We present our results based on three-year averages of data.

Our estimates reveal that geopolitical distance became a significant driver of trade flows only since 2018, and its impact has steadily increased over time (Figure 1, Panel B). The fall in the elasticity of geopolitical distance is mostly driven by deteriorating geopolitical relations, most notably between the US and China and more generally between the West and the East.

These reflect the effect of increased trade restrictions in key strategic sectors associated to the COVID-19 pandemic crisis, economic sanctions imposed to Russia, and the rise of import substituting industrial policies.

The impact of geopolitical distance is also economically significant: a 10% increase in geopolitical distance (like the observed increase in the USA-China distance since 2018, in Figure 1) is found to decrease bilateral trade flows by about 2%. In Bosone and Stamato (forthcoming), we show that these results are robust to several specifications and to an instrumental variable approach.

Figure 1. Evolution of geopolitical distance between selected country pairs and its estimated impact on bilateral trade flows



Notes: Panel A: geopolitical distance is based on the ideal point distance proposed by Bailey et al (2017), which measures countries' disagreements in their voting behaviour in the UN General Assembly. Higher values mean higher geopolitical distance. Panel B: Dots are the coefficient of geopolitical distance, represented by the logarithm of the ideal point distance interacted with a time dummy, using 3-year averages of data and based on a gravity model estimated for 67 countries from 2012 to 2022. Whiskers represent 95% confidence bands. The dependent variable is nominal trade in manufacturing goods, excluding energy. Estimation performed using the PPML estimator. The estimation accounts for bilateral time-varying controls, exporter/importer-year fixed effects, and pair fixed effects.
Sources: TDM, IMF, Bailey et al (2017), Egger and Larch (2008), WITS, Eurostat, and ECB calculations.

Friend-shoring or near-shoring?

Recent narratives surrounding trade and economic interdependence increasingly argue for localisation of supply chains through near-shoring and strengthening production networks with like-minded countries through friend-shoring (Yellen 2022).

To offer quantitative evidence on these trends, we first regress bilateral trade flows on a set of four dummy variables that identify the four quartiles of the distribution of geopolitical distance across country pairs. To capture the effect of growing geopolitical tensions on trade, each dummy is equal to 1 for trade within the same quartile from 2018 and zero otherwise.

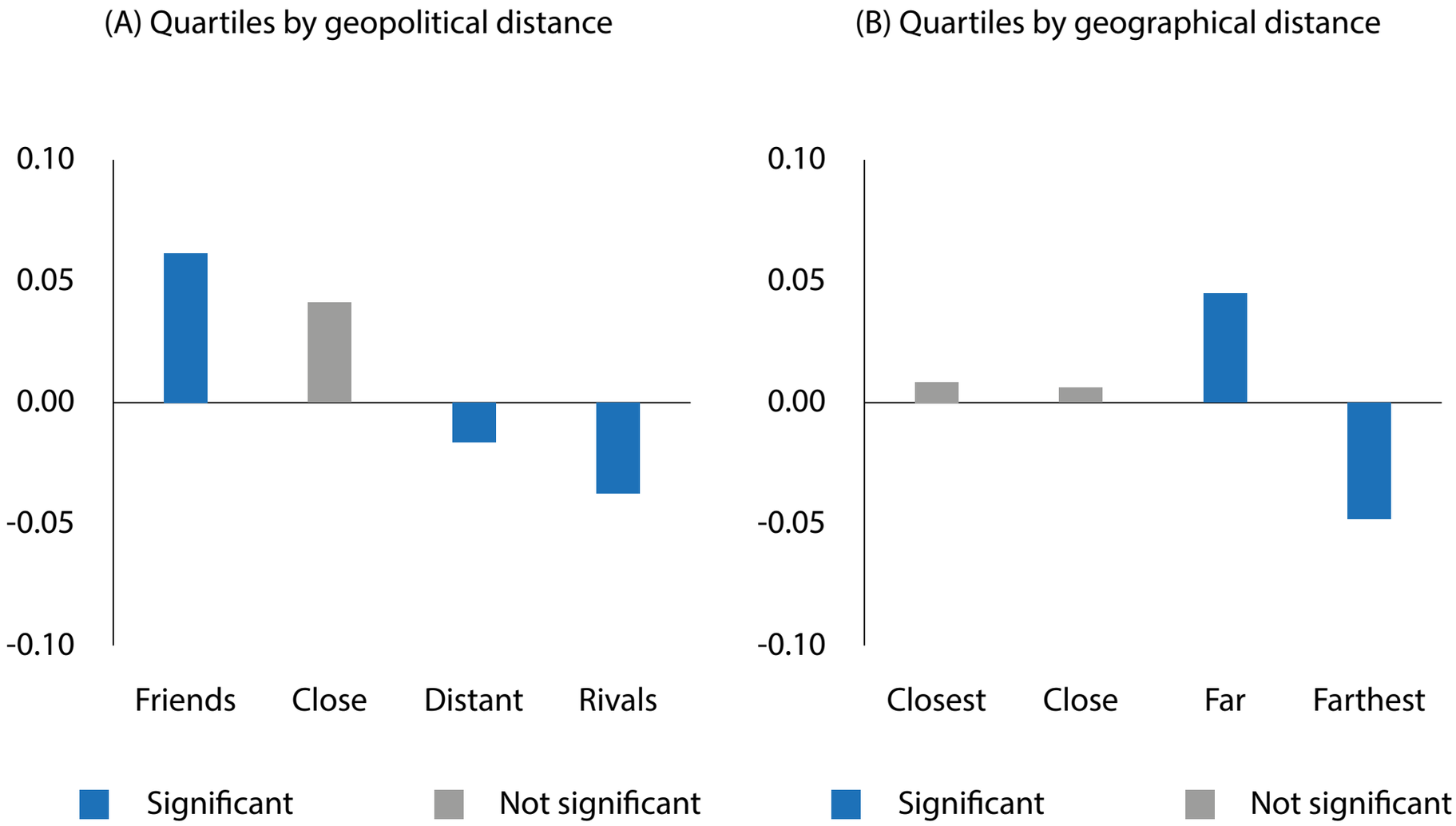
We find compelling evidence of friend-shoring. Trade between geopolitically aligned countries increased by 6% since 2018 compared to the 2012–2017 period. Meanwhile, trade between rivals decreased by 4% (Figure 2, Panel A). In contrast, our estimates do not reveal the presence of near-shoring trends (Figure 2, Panel B).

Instead, we find a significant increase in trade between *far*-country pairs, offset by a relatively similar decline in trade between the *farthest*-country pairs. Overall, shifts toward geographically close partners are less pronounced than toward geopolitically aligned partners.

Evidence of de-risking in EU trade

The trade impact of geopolitical distance on the EU is isolated by interacting geopolitical distance with a dummy for EU imports. We find that EU aggregate imports are not significantly affected by geopolitical considerations (Figure 3, Panel A).

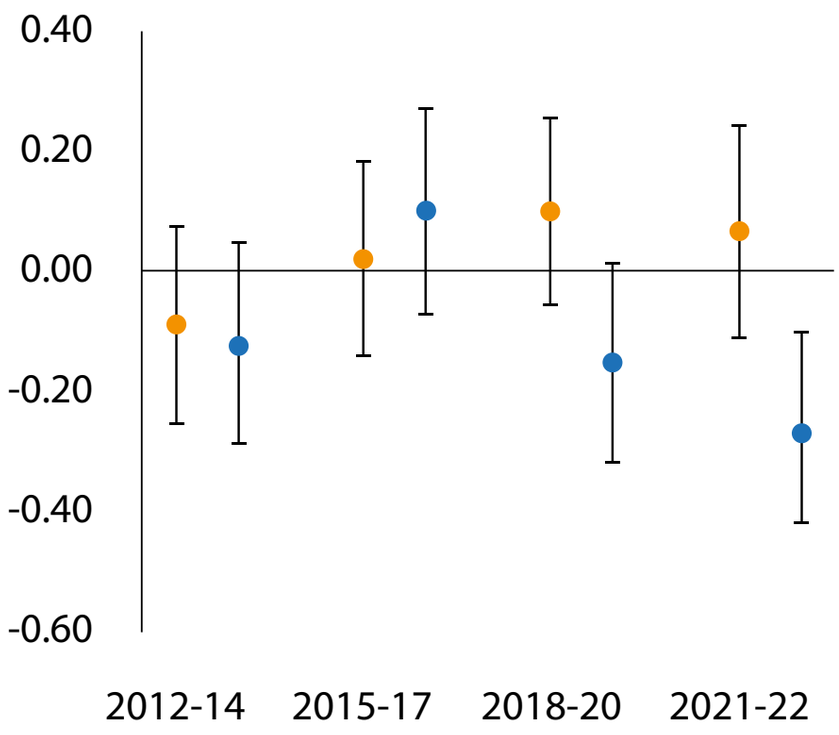
Figure 2. Impact of trading within groups since 2018 (semi-elasticities)



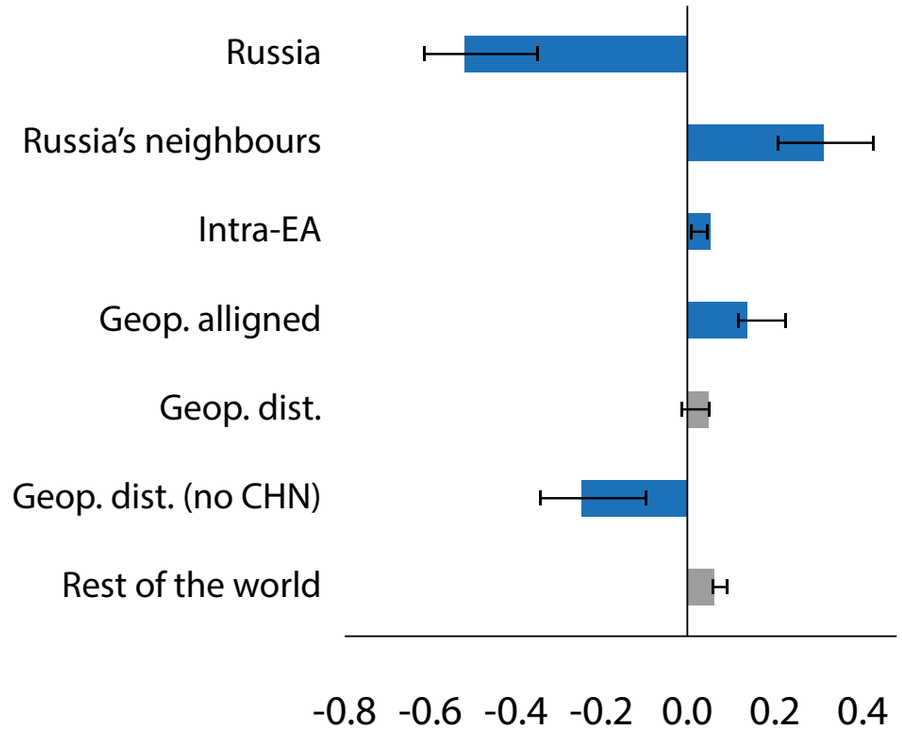
Notes: Estimates in both panels are obtained by PPML on the sample period 2012–2022 using consecutive years. Please refer to Figure 1 for details on estimation. The effects on each group are identified based on a dummy for quartiles of the distribution of geopolitical distance (panel A) and on a dummy for quartiles of the distribution of geographic distance (panel B) across country pairs. The dummy becomes 1 in case of trade between country pairs belonging to the same quartile since 2018.
Sources: TDM, IMF, Bailey et al (2017), Egger and Larch (2008), WITS, Eurostat, CEPII, and ECB calculations.

Figure 3. Impact of geopolitical distance on EU imports and of the Ukraine war on euro area exports

(A) Time-varying impact of geopolitical distance on EU imports (elasticity)



(B) Impact of the Ukraine war on euro area exports (semi-elasticity)



Strategic products
Aggregate imports

Significant
Not significant

Notes: Estimates in both panels are obtained by PPML on the sample period 2012–2022. Panel A: Dots represent the coefficient of geopolitical distance interacted with a time dummy and with a dummy for EU imports, using 3-year averages of data. Lines represent 95% confidence bands. Panel B: The sample includes quarterly data over 2012–2022 for 67 exporters and 118 importers. Effects on the level of euro area exports are identified by a dummy variable for dates after Russia's invasion of Ukraine. Trading partners are Russia; Russia's neighbours Armenia, Kazakhstan, the Kyrgyz Republic, and Georgia; geopolitical friends, distant, and neutral countries are respectively those countries that voted against or in favour of Russia or abstained on both fundamental UN resolutions on 7 April and 11 October 2022. The whiskers represent minimum and maximum coefficients estimated across several robustness checks.
Sources: TDM, IMF, Bailey et al (2017), Egger and Larch (2008), WITS, Eurostat, European Commission, and ECB calculations.

This result is robust to alternative specifications and may reflect the EU's high degree of global supply chain integration, the fact that production structures are highly inflexible to changes in prices, at least in the short term, and that such rigidities increase when countries are deeply integrated into global supply chains (Bayoumi *et al* 2019).

Nonetheless, we find evidence of de-risking in strategic sectors[1]. When we use trade in strategic products as the dependent variable, we find that geopolitical distance significantly reduces EU imports (Figure 3, Panel A).

We conduct an event analysis to explore the implications of Russia's invasion of Ukraine on euro area exports. We find that the war has reduced euro area exports to Russia by more than half (Figure 3, Panel B), but trade flows to Russia's neighbours have picked up, possibly due to a reordering of the supply chain.

Euro area exports with geopolitically aligned countries are estimated to have been about 13% higher following the war, compared with the counterfactual scenario of no war. We find no signs of euro area trade reorientation away from China, possibly reflecting China's market power in key industries.

However, when China is excluded from the geopolitically distant countries, the impact of Russia's invasion of Ukraine on euro area exports becomes strongly significant and negative.

Concluding remarks

Our findings point to a redistribution of global trade flows driven by geopolitical forces, reflected in the increasing importance of geopolitical distance as a barrier to trade.

In this column we review recent findings on geopolitics in trade and their impact since 2018, the emergence of friend-shoring rather than near-shoring, and the interactions of strategic sectors with geopolitics in Europe.

In sum, we bring evidence of new forces that now drive global trade – forces that are no longer guided by profit-oriented strategies alone but also by geopolitical alignment. ■

Costanza Bosone is a PhD candidate at the University of Pavia, and Ernest Dautović is a Supervisor, Michael Fidora a Senior Lead Economist Giovanni Stamato a Consultant, all at the European Central Bank

Endnote

1. We follow the definition given by the European Commission and define strategic products as military equipment, raw materials, battery packs, high-tech, medical goods, and all those goods which are particularly relevant for security, public health, and the green and digital transitions.

References

Aiyar, S, PO Gourinchas, A Presbitero, and M Ruta (2023), "Geoeconomic fragmentation: A new eBook", VoxEU.org, 2 October.

Antràs, P (2021), "De-globalisation? Global value chains in the post-COVID-19 age", Central Banks in a Shifting World: Conference Proceedings – ECB forum.

Attinasi, MG, L Boeckelmann, and B Meunier (2023), "The economic costs of supply chain decoupling", ECB Working Paper 2023/2389.

Bailey, MA, A Strezhnev, and E Voeten (2017), "Estimating dynamic state preferences from United Nations voting data", Journal of Conflict Resolution 61(2): 430–56.

Baldwin, R (2022), "The peak globalisation myth: Part 3 – How global supply chains are unwinding", VoxEU.org, 2 September.

Bayoumi, T, J Barkema, and D Cerdeiro (2019), "The inflexible structure of global supply chains", IMF Working Paper 2019/193.

Bosone, C, E Dautović, M Fidora, and G Stamato (2024), "How geopolitics is changing trade", Economic Bulletin Boxes 2/2024.

Bosone, C, and G Stamato (forthcoming), "Beyond borders: How geopolitics is changing trade", European Central Bank Working Paper Series.

Campos, RG, J Estefania-Flores, D Furceri, and J Timini (2023), "Geopolitical fragmentation and trade", VoxEU.org, 31 July.

Fajgelbaum, PD, and AK Khandelwal (2022), "The economic impacts of the US-China trade war", Annual Review of Economics 14: 205–28.
Góes, C, and E Bekkers (2022), "The impact of geopolitical conflicts on trade, growth, and innovation: An illustrative simulation study", VoxEU.org, 29 March.
Goldberg, PK, and T Reed (2023), "Is the global economy deglobalizing? And if so, why? And what is next?", Brookings Papers on Economic Activity, Spring.
Yellen, JL (2022), "Remarks on the way forward for the global economy", speech at the Atlantic Council, Washington DC, April.

This article was originally published on VoxEU.org.

Understanding the the WTO e-commerce moratorium

The WTO e-commerce moratorium has been renewed. Andrea Andrenelli and Javier López González explore the different issues around the moratorium debate

The last two decades have seen momentous shifts in globalisation as a result of digital transformation (Baldwin 2019, Winters and Borchert 2021, Savona 2020). During this time, the WTO moratorium on applying customs duties on electronic transmissions, the only WTO provision that applies explicitly to digital trade, has underpinned a stable, predictable, and duty-free environment for digital trade to thrive (IMF *et al* 2023).

At the last WTO Ministerial Conference, after difficult negotiations, the moratorium was renewed, and WTO members agreed to continue discussions on its *scope*, *definition*, and *impact*.

What is the e-commerce moratorium and why is it controversial?

The WTO e-commerce moratorium is a commitment to continue the current practice of not applying customs duties (ie. tariffs) on electronic transmissions. However, since 'electronic transmissions' were never defined, there is room for interpretation about the precise scope of the commitment.

Recently, several WTO members have raised questions about the opportunity costs of the moratorium[1]. Chief among their concerns is the potential loss of 'policy space' in the context of rapid technological change and potential losses in customs revenue due to the 'dematerialisation' of goods trade. For these WTO members, the lack of clarity on issues of scope and definition makes it difficult to understand the potential value, or opportunity cost of the moratorium.

In a recent paper (Andrenelli and López-González 2023), we review regional trade agreement provisions related to the electronic transmissions, provide new estimates of the potential foregone revenue implications of the moratorium, and explore some of the potential impacts of not renewing the moratorium on trade and competitiveness.

What can we learn from regional trade agreements about the scope and definition of the moratorium?
Some WTO members question whether the moratorium applies to the 'content' of the transmission (that is, the actual movies or e-books downloaded) or its 'carrier medium' (the bits and bytes that carry the content)[2]. Questions have also been raised about whether the Moratorium affects the ability of countries to apply other, internal, taxes beyond customs duties, or if the Moratorium erodes other commitments made in the WTO.

Our analysis suggests that the potential foregone revenue costs of the Moratorium are small and that its lapse would come at the expense of wider gains in the economy

Much can be learnt about the potential scope of the moratorium by looking at how countries have approached customs duties on electronic transmissions in their trade agreements. Analysis using the Trade Agreement Provisions on Electronic Commerce and Data (TAPED) database (Burri *et al* 2022) shows that, of the 105 regional trade agreements (RTAs) with an e-commerce chapter (by end of 2022), 100 included a provision on the non-imposition of customs duties on electronic transmissions (NICDET provision for short). More detailed analysis of these provisions reveals that (Figure 1):

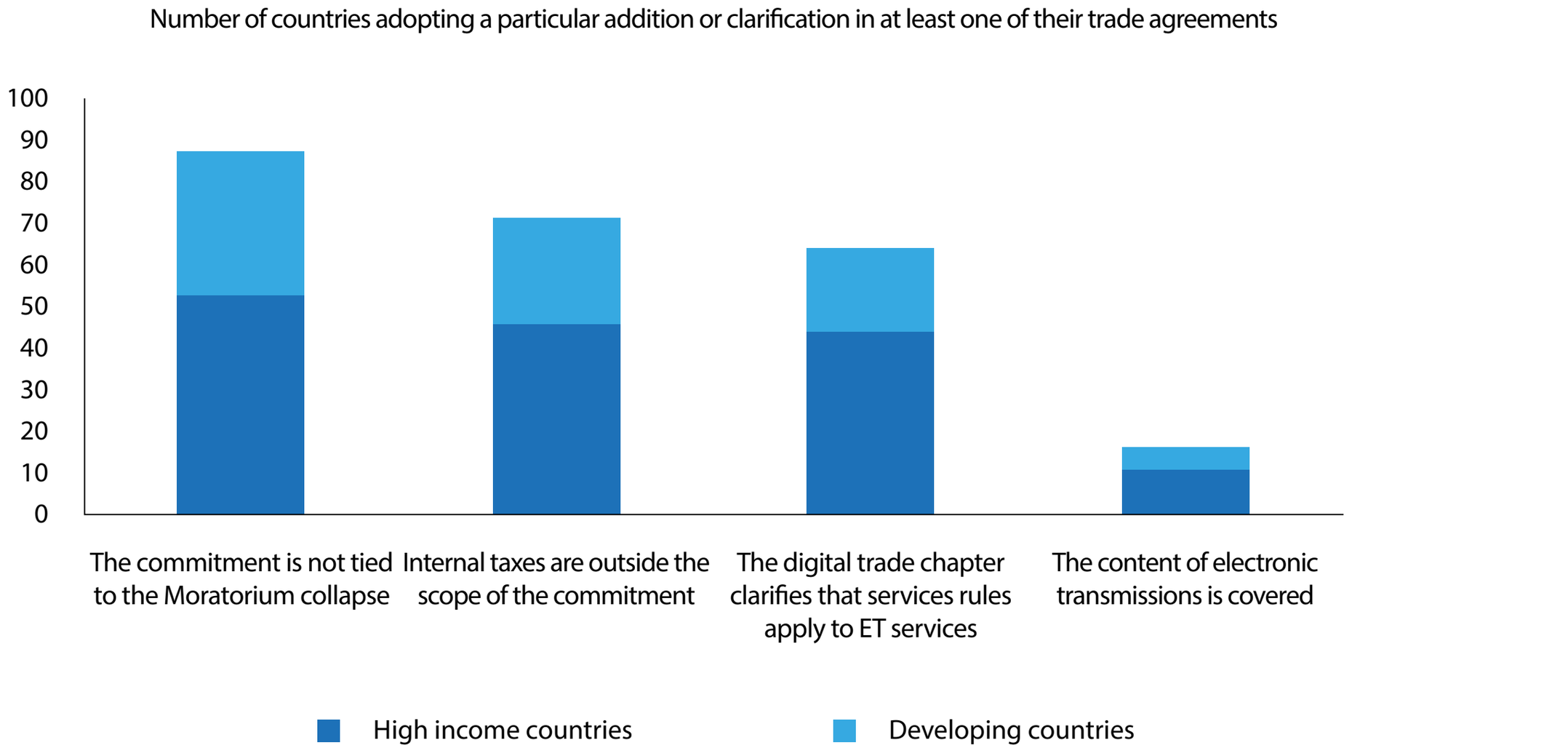
- The majority of NICDET commitments, 88%, are not tied to the e-commerce moratorium. Specifically, 54 high income and 33 developing countries would continue not imposing customs duties on electronic transmissions, at least on a reciprocal basis, even if the moratorium were to lapse.

- The majority of NICDET provisions clarify that internal taxation is outside the scope of commitments. Most countries do not see the commitment as having implications for applying other forms of taxation, including value added taxes (VAT) or goods and services taxes (GST).

- Digital trade chapters generally reaffirm that measures related to electronic delivery fall within the scope of obligations and exceptions related to services (eg. the General Agreement on Trade in Services (GATS) or regional trade agreement commitments and flexibilities remain). This suggests that the moratorium is unlikely to restrict 'policy space' beyond the non-imposition of tariffs.

- Since 2015, members have started to clarify that NICDET commitments apply to the content of electronic transmissions. There are no trade agreements clarifying that NICDET provisions apply to the 'carrier medium'.

Figure 1. Non-imposition of customs duties on electronic transmissions (NICDET) commitments in regional trade agreements can provide useful guidance on the interpretation of the potential scope and definition of the moratorium



Note: Income group classification based on the 2022-2023 World Bank classification, where developing countries refers to lower-middle-income and upper-middle-income countries.
Source: Andrenelli and López-González (2023).

Some countries define electronic transmissions as 'digital products' which include computer programmes, text, video, images, sound recordings, and other products that are digitally encoded. Others clarify that 'deliveries by electronic means shall be considered as the provision of services'. Others just use the term 'electronic transmissions', without any further clarifications.

However, differences in definitions have not prevented the conclusion of NICDET provisions between countries with different definitions[3]. While for some the lack of a precise definition might be considered a challenge, for others it is a way of enabling a variety of views to coexist.

What are the potential fiscal implications of the moratorium?

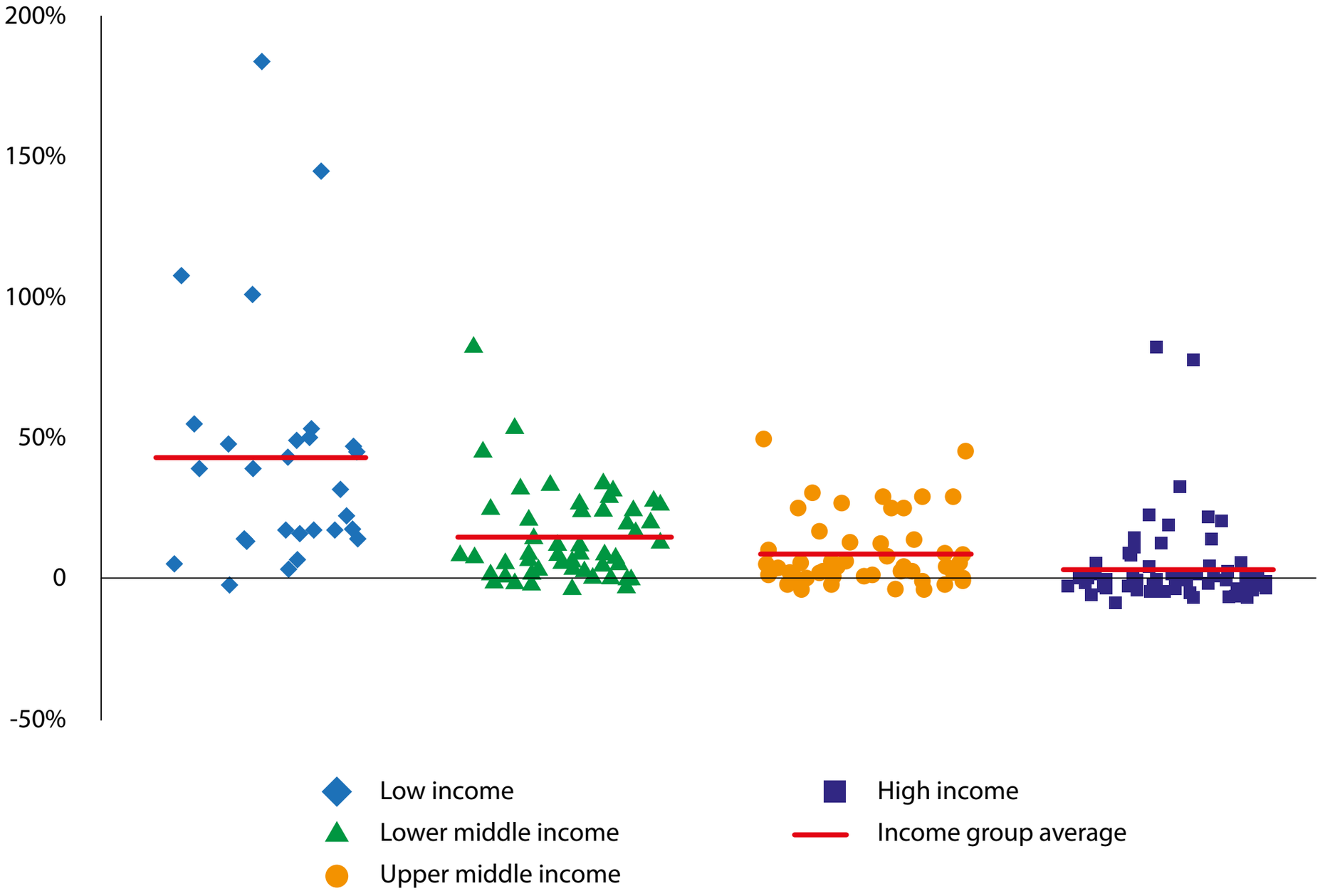
Some WTO members worry that not imposing customs duties on electronic transmissions may lead to foregone customs revenue. That is, a country importing a movie via an electronic transmission foregoes the tariff revenue associated with its import via a physical carrier medium, such as through a DVD. They argue that the rapid pace of digitalisation increases the scale of the problem, especially for developing countries, which tend to charge higher tariffs on these items.

However, imports of 'digitisable goods', which are physical goods that can be digitised and subsequently sent across borders digitally (e.g. CDs, books, calendars, videotapes), have generally been growing over the last decade, especially in developing countries (Figure 2), continuing to generate tariff revenue.

Accurately assessing the potential foregone revenue implications of the moratorium is not easy given uncertainties about scope and definition. However, we argue that existing empirical studies (Banga 2022, 2019) have not addressed three important issues that bias current estimates upwards.

Figure 2. Imports of digitisable goods have been growing, particularly in low-income countries

Average yearly change in physical imports of digitisable goods in 2008-2019, by income group



Note: Markers represent individual countries. Based on 206 countries and territories. Red lines show the income group average. The horizontal axis line indicates 0% average growth. Calculations based on BACI data.
Source: Andrenelli and López-González (2023).

The first is that existing commitments and practices, such as NICDET provisions or other preferences granted in regional trade agreements, limit the ability of countries to raise tariffs on digitisable goods and electronic transmissions, even in the absence of the e-commerce moratorium.

The second is that not all trade that can be electronically transmitted will be (as seen above, imports of digitisable goods have actually been increasing for many countries). The third is that assessments need to consider the potential offsetting effects of VATs/GSTs applied on growing digital imports.

We find that the foregone customs revenue that can be attributed to the moratorium is small – on average 0.68% of total customs revenue or 0.1% of overall government revenue. Given higher tariffs and lower levels of commitments, impacts are on average higher for low-income countries (0.33% of government revenue), and lower for high income country (0.01%).

That said, for 77 of 106 countries analysed, standard VAT/GST taxes applied on digital services imports which are 'born digital' completely offset the customs revenue effects of the moratorium[4].

These findings underscore the potential to find fiscal solutions, based on consumption taxes, to collect revenue on immaterial imports based on widely adopted and internationally accepted standards (OECD 2017). These taxes are efficient and have a demonstrated capacity of increasing tax revenues (Hanappi *et al* 2024).

In addition, since single rates tend to apply, there is no need to spend resources identifying how to classify products into detailed nomenclatures or to determine their origin. These taxes also target final, instead of intermediate consumption, which, as we will show below, is important.

What benefits are at stake with the potential lapse of the e-commerce moratorium?

We find that tariffs on electronic transmissions have the potential to hit low-income country trade most. If existing tariffs on digitisable goods were to be applied to digital services (which is where electronic transmissions are measured in existing trade statistics) imports of low-income countries would fall by 32% and exports would fall by 2.5%.

This is because more than 80% of digital services exports of low-income countries are to middle income countries which have more scope to increase tariffs. For middle-income countries, losses would be of 6% and 0.4% and for high-income countries of 0.04% and 0.5%, respectively.

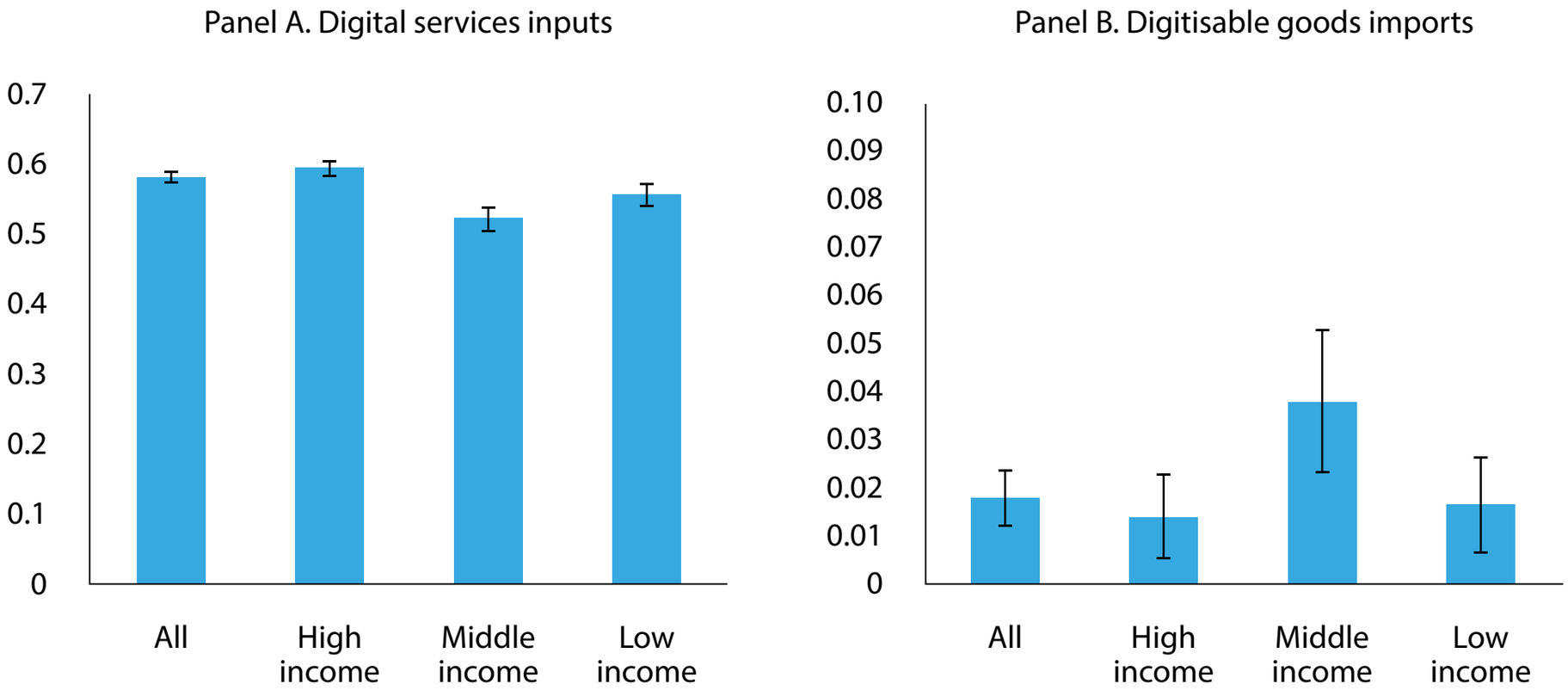
Evidence also shows that the use of foreign digital inputs and digitisable goods contributes to domestic competitiveness, measured as changes in the domestic value added in final consumption (Figure 3). This suggests that trade cost increases arising from the termination of the Moratorium would lead to losses in domestic competitiveness. Therefore, there is a self-interest argument for maintaining a duty-free environment for electronic transmissions.

The impact of greater barriers on electronic transmissions is also likely to be asymmetric, affecting small and medium-sized enterprises (SMEs) most. Analysis using the World Bank Enterprise Survey (WBES) suggests that being able to deliver trade digitally is associated with higher propensities to export of smaller firms and not larger ones.

Since SMEs generally have a lower propensity to export than larger firms, the ability to deliver products digitally may be an important mechanism to reach foreign markets, and this channel may be affected by the Moratorium lapse.

Figure 3. Digital inputs are key determinants of domestic competitiveness

Figures show the impact of increasing use of foreign digital inputs by one standard deviation on domestic value added



Note: Standardised regression coefficients capturing impact of increasing digital services inputs and digitisable goods imports on domestic value added with confidence intervals (95%). Calculations based on data from TRAINS and ITPDE.
Source: Andrenelli and López-González (2023).

There is a strong economic case for keeping electronic transmissions free from tariffs

Overall, our analysis suggests that the potential foregone revenue costs of the Moratorium are small and that its lapse would come at the expense of wider gains in the economy. ■

Andrea Andrenelli is a Trade Policy Analyst and Javier López González a Senior Economist, at the Organisation for Economic Co-Operation and Development (OECD)

Endnotes

1. See WTO Communications WT/GC/W/747, WT/GC/W/798 and WT/GC/W/833.

2. See WTO Communication WT/GC/W/859

3. For example, the EU-Canada agreement relies on flexible language, calling these "a delivery transmitted by electronic means", to bridge existing differences.

4. 'Born digital' trade is proxied using data on trade in computer, audio-visual, and information services imports. The intuition is that this captures growth in trade that might not have been previously delivered via physical carrier media. For instance, the is no physical goods equivalent of cloud computing services.

References

Andrenelli, A and J López-González (2023), "Understanding the scope, definition and impact of the e-commerce Moratorium", OECD Trade Policy Papers, No. 275, OECD Publishing, Paris.

Baldwin, R (2019), The globotics upheaval: Globalisation, robotics and the future of work, Oxford University Press.

Banga, R (2019), "Growing Trade in Electronic Transmissions: Implications for the South", UNCTAD Research Paper, No. 29, UNCTAD, Geneva.

Banga, R (2022), "WTO Moratorium on Customs Duties on Electronic Transmissions: How much tariff revenue have developing countries lost?", South Centre Research Paper 157.

Burri, M, M Vasquez Callo-Müller and K Kugler (2022), "TAPED: Trade Agreement Provisions on Electronic Commerce and Data", University of Lucerne.

Hanappi, T, A Jakubik and M Ruta (2023), "Fiscal Revenue Mobilization and Digitally Traded Products: Taxing at the Border or Behind It?", IMF Notes No. 2023/005.

Hanappi, T, A Jakubik and M Ruta (2024), "Why the WTO moratorium on digital tariffs should be extended", VoxEU.org, 26 February.

IMF, OECD, UNCTAD, WBG, WTO (2023), Digital trade for development.
OECD (2017), International VAT/GST Guidelines, OECD Publishing, Paris.
Savona, M (2020), "Governance models for redistribution of data value", VoxEU.org, 17 January.
Winters, A and I Borchert (2021), "Addressing impediments to digital trade: A new eBook", VoxEU.org, 27 April.

Authors' note: The opinions expressed and arguments employed are those of the authors and do not represent the official views of the OECD or of its member countries. This article was originally published on VoxEU.org.

Geoeconomic fragmentation and firms' financial performance

The threats of geoeconomic fragmentation have accelerated in recent years. Alessandro D'Orazio, Fabrizio Ferriani and Andrea Gazzani introduce a novel firm-level revenue-weighted geopolitical risk index

How do economic and financial interdependencies among countries and firms respond when seismic geopolitical shifts disrupt the rule-based international order? This question has gained prominence in policy debates as the advantages accrued over decades of economic integration are threatened by escalating tensions, leading to a reversal of international relations.

This phenomenon, labelled as *geoeconomic fragmentation* (Aiyar and Ilyina 2023), has been accelerating in recent years. Events such as Brexit, trade disputes between the US and China, trade flow restrictions associated with the COVID-19 pandemic, and, more dramatically, Russia's invasion of Ukraine and the Israeli-Palestinian conflict, have all contributed to this trend.

As of now, the analysis on the impacts of geoeconomic fragmentation has mainly focused on how the deterioration of international relations can impact international trade and highly interconnected global value chains (eg. Campos *et al* 2023, Attinasi *et al* 2023). Concerns of geoeconomic fragmentation in commodity markets have intensified since the start of Russia's invasion of Ukraine (eg. IMF 2023a, Emiliozzi *et al* 2024, Albrizio *et al* 2023).

Conversely, the study of the financial impacts of geopolitical tensions has been more limited so far, with most analyses focusing on cross-border capital flows (especially foreign direct investment), asset prices and investors' risk aversion at the *aggregate level* (IMF 2023b, Feng *et al* 2023, Aiyar *et al* 2024).

In a recent study (D'Orazio *et al* 2024), we present evidence on the financial impacts of geoeconomic fragmentation from a *firm-level* perspective. Our study covers the period 2010-2022 and relies on a large sample of non-financial corporations included in the Eurostoxx 600 and the S&P500 equity indexes.

A novel firm-level measure of exposure to geopolitical risk

We propose a novel *firm-level* measure of exposure to geopolitical risk combining detailed information on the geographic distribution of corporate revenues with country-specific assessments of geopolitical risk to create a revenue-weighted geopolitical risk indicator.

Data on the geographical breakdown of corporate revenues are retrieved from explanatory notes to the official financial statements and is used to identify the ultimate origin of firms' business risk, specifically the location where firms generate their revenues. Figure 1 displays the geographical breakdown of corporate revenues (using macro-aggregates for readability).

As global tensions continue unabated, the financial consequences of fragmentation for firms may intensify, amplifying macro-financial turbulence

Not surprisingly, the largest share of revenues originates from the geographical area where firms are listed: approximately 72% of revenues are generated in the US and Canada for S&P500 firms, compared to an average of 64% of revenues generated in Europe for Eurostoxx companies. In both regions, revenue generated in China hovers around 3%.

Data on country geopolitical risk relies on the yearly assessment of political risk developed by the International Country Risk Guide (ICRG). The political risk rating ranges from 0 to 100, with higher scores associated with lower risk levels, and it covers the assessment of geopolitical risk across twelve dimensions: *government stability, socioeconomic conditions, investment profile, internal conflict, external conflict, corruption, military in politics, religious tensions, law and order, ethnic tensions, democratic accountability,* and *bureaucracy quality*.

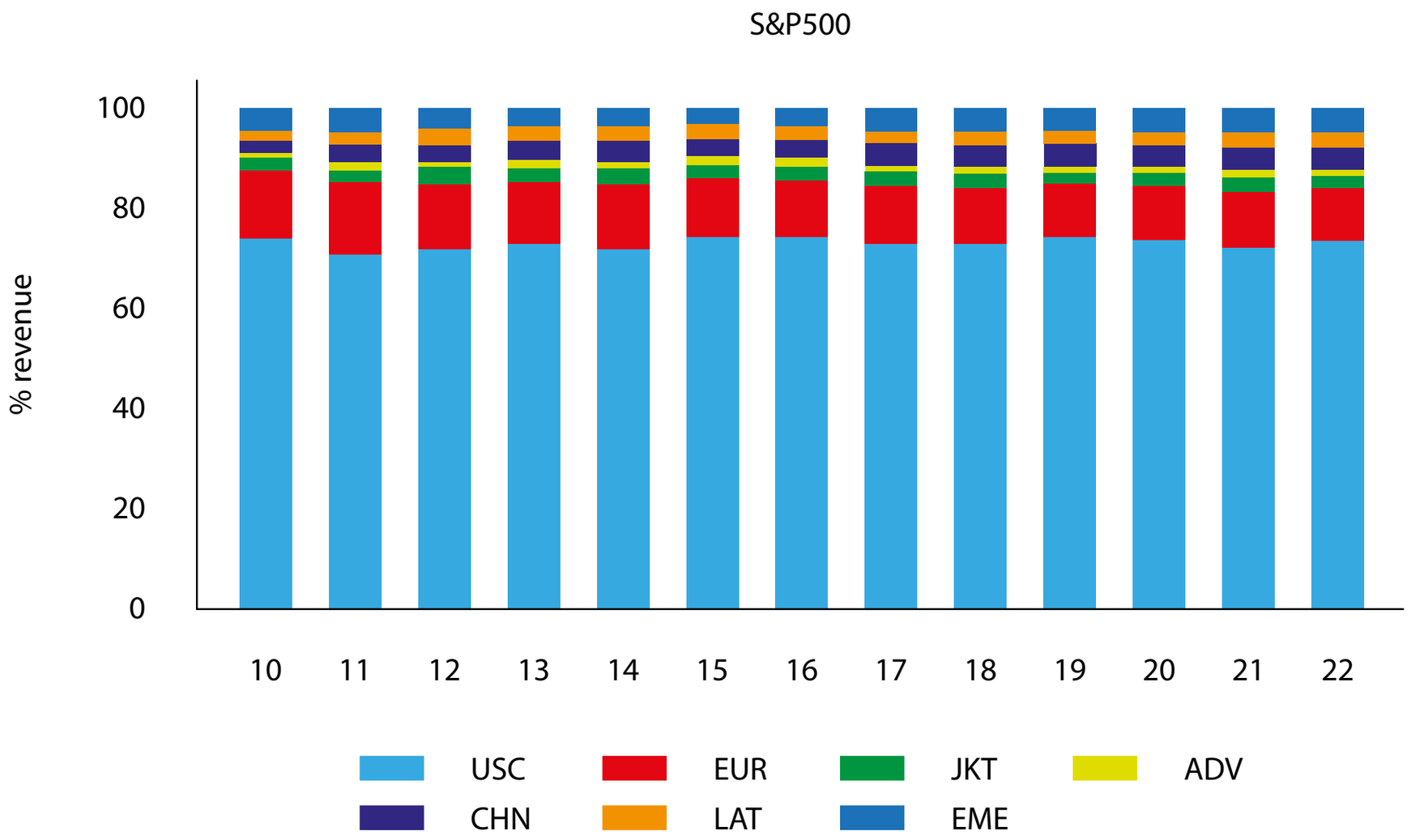
Global average risk score is moderately declining over time (ie. geopolitical risk increases) and exhibits high variability across countries, with the political risk score ranging between approximately 30 and 90 points out of 100.

Figure 2 presents the risk ranking based on the ICRG 2022 assessment: lower geopolitical scores are generally associated with advanced economies (Western countries, Japan, Australia, South Korea), while most emerging economies exhibit higher geopolitical risk.

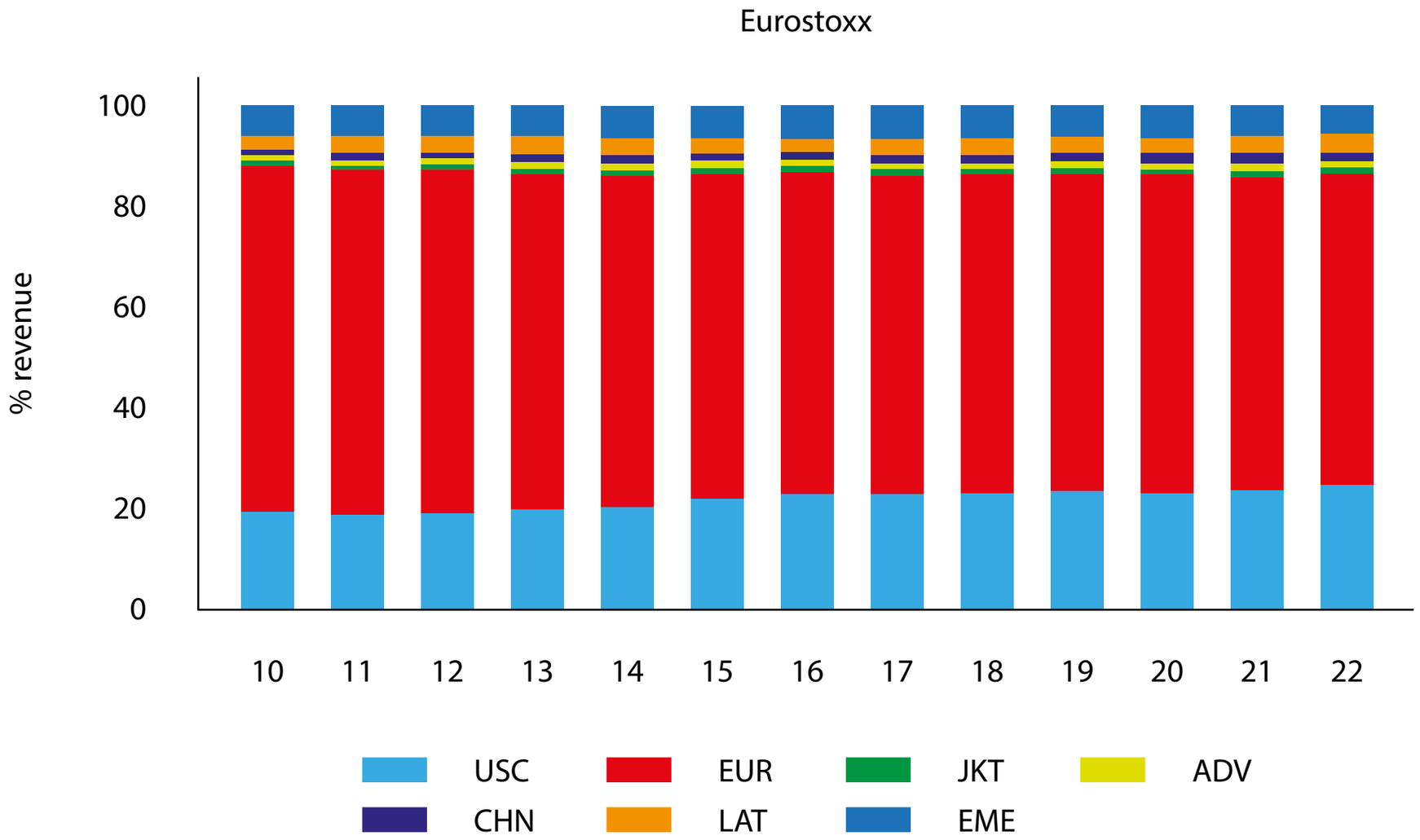
We multiply the shares of firm revenues originating in each national market by the corresponding value of the country-specific ICRG index to obtain a revenue-weighted measure of firms' exposure to geopolitical risk (*Gprisk revenue weighted*).

We analyse the impact of this measure on several indicators of corporate financial performance, namely the *Altman Z-score*, which constitutes an inverse proxy of firms' default probability based on accounting variables, and the

Figure 1. Geographical breakdown of corporate revenues: S&P500 vs. Eurostoxx

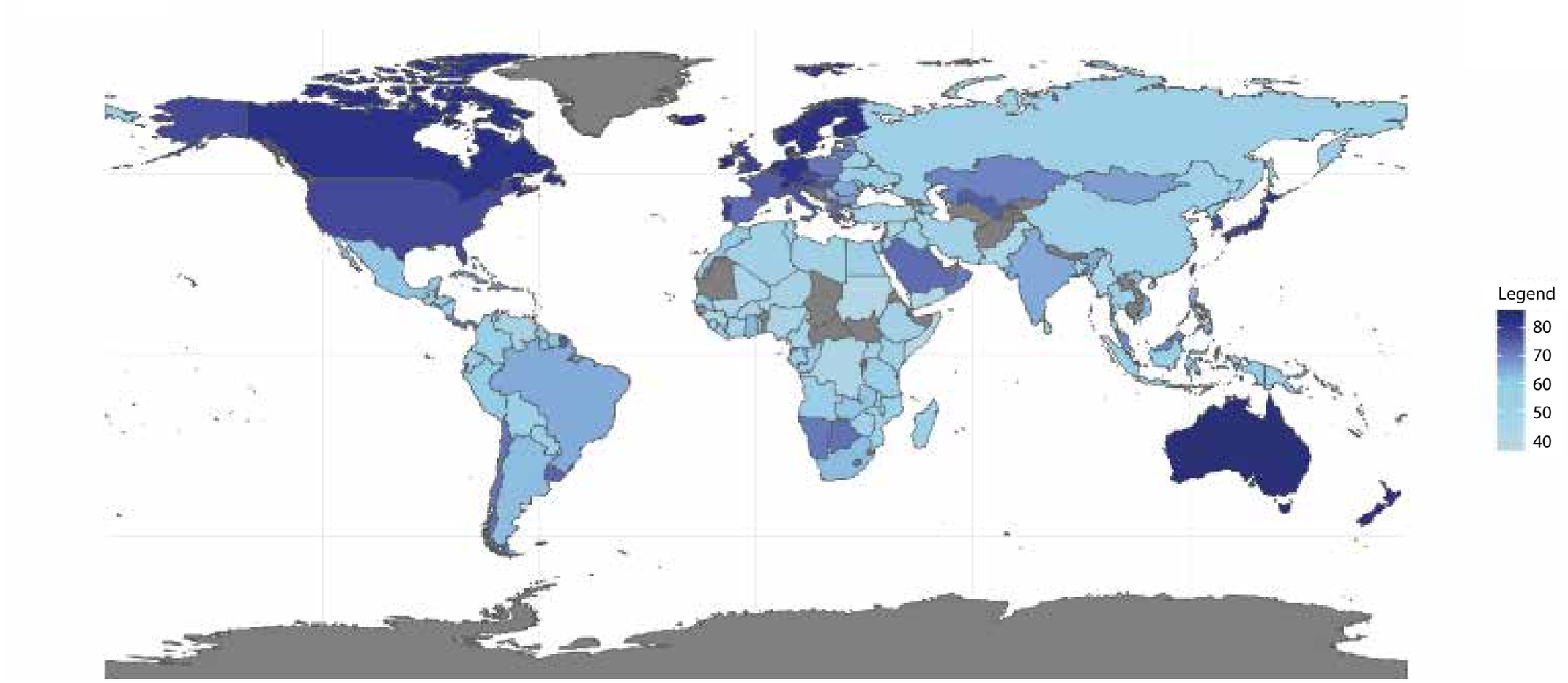


Note: Acronyms are as follows: United States and Canada (USC), Europe (EUR), Japan, South Korea and Taiwan (JKT), other advanced economies (ADV), China (CHN), Latin America (LAT), other emerging markets (EME). Our elaborations from Orbis-Bureau van Dijk database.



Note: Acronyms are as follows: United States and Canada (USC), Europe (EUR), Japan, South Korea and Taiwan (JKT), other advanced economies (ADV), China (CHN), Latin America (LAT), other emerging markets (EME). Our elaborations from Orbis-Bureau van Dijk database.

Figure 2. Country-level measure of geopolitical risk



Note: The plot displays countries' geopolitical risk in 2022. Data are from ICRG, higher values correspond to lower risk levels; grey countries have no available score.

price-to-earnings (P/E) ratio, and the *Tobin's Q ratio*, reflecting investors' assessment of firms' profitability and market value.

The impact of geopolitical risk exposure on firms' financial performance

Firms' revenue exposure to markets characterised by high geopolitical risk impacts corporate viability and this is also reflected in lower investors' valuations. Figure 3 graphically presents our results, including, for comparative purposes, the effects of a naïve and less sophisticated measure of firm geopolitical risk based on firms' headquarters (*Gprisk* HQ).

The two measures can imply very distinct assessments of corporate exposure to geopolitical risk. For instance, consider two firms headquartered in the US – one generating all revenues from the local market and another with half revenues from the US and half from China.

The geopolitical risk based on the headquarters' exposure is identical for both firms, amounting to 79, according to the ICRG scores in 2022. In contrast, the assessment based on revenue exposure is 79, for the former firm with no foreign revenues, but only 68 for the latter firm with more diversified revenues.

In economic terms, a one standard deviation increase in our revenue-weighted geopolitical risk measure, ie. an improvement in terms of risk exposure, results in a roughly 0.5 standard deviations increase in corporate viability (Z-score), 4.7% increase for the P/E ratio, and 3.3% increase in the case of the Tobin Q ratio.

Conversely our estimates show that, this relationship is muted when examining geopolitical risk based on firms' headquarters. This result squares with the graphical evidence reported in Figure 2 and stems from S&P 500 and Eurostoxx firms being headquartered in countries with generally lower geopolitical risk.

However, it is noteworthy that even relatively modest shares of revenue exposure to markets with higher geopolitical risks (roughly 12-15% on average across time) have substantial financial effects.

These results should be interpreted as conservative estimates of the actual effect, as our revenue geographical breakdown pertains to revenues originating from the sale of final goods and services and does not account for other forms of cross-country linkages (e.g. intermediate output trades) arising from a firm exposure to sourcing from different countries.

In an additional exercise, we investigate whether the recent upswing in geopolitical tensions has led to more significant repercussions on the viability and valuations of firms (see Figure 4 for a fragmentation index obtained from corporate earnings calls). We find that the financial impact of geopolitical risk has increased since 2017.

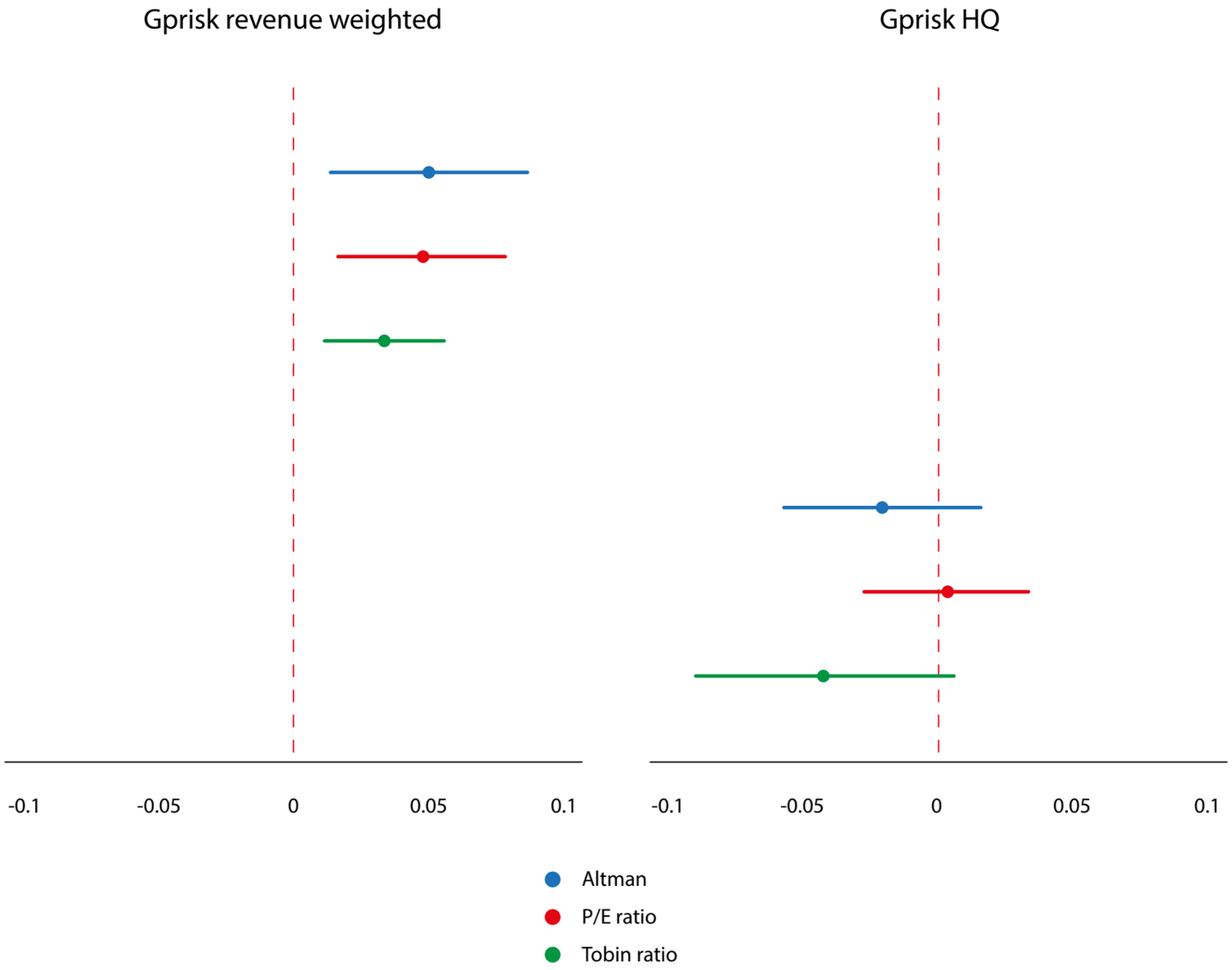
During this period, concerns regarding geoeconomic fragmentation began to be more prominently reflected in firms' risk assessments, amid escalating trade tensions and heightened protectionist rhetoric.

Conclusions

We introduce a novel revenue-weighted geopolitical risk index at the firm level and observe that geopolitical risk substantially affects firms' default probability and market valuations, with a notable escalation in the impacts since 2017.

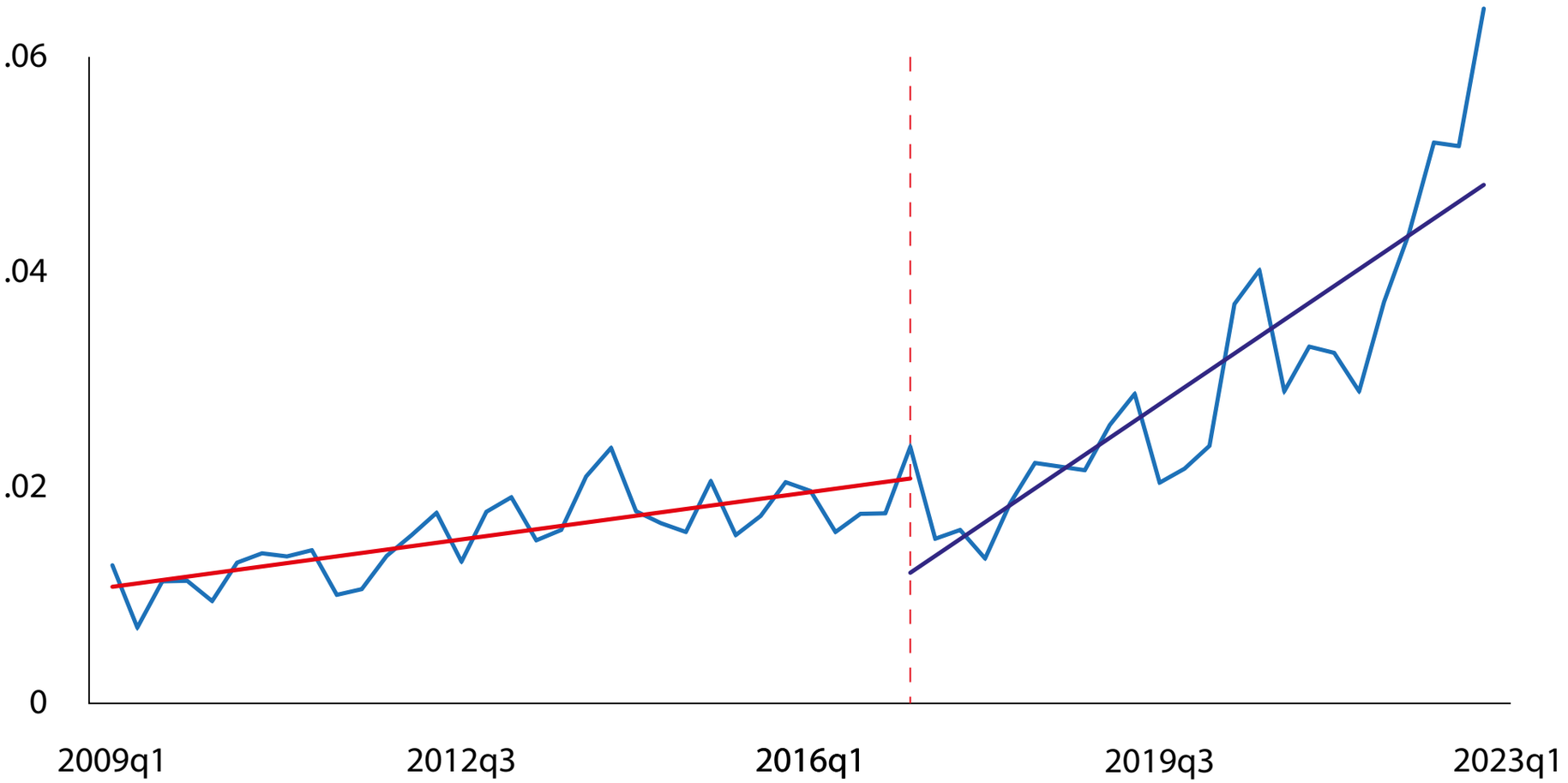
The absence of statistical significance regarding geopolitical risks associated with firms' headquarters emphasises the importance of accessing accurate microdata to precisely measure the real-financial interdependencies of geoeconomic fragmentation (Borin *et al* 2024).

Figure 3. Impact of exposure to geopolitical risk on firms' financial performance



Note: Gprisk revenue weighted (left panel) combines corporate revenue distribution with geopolitical risk across countries, Gprisk HQ (right panel) is based on geopolitical risk of firms' headquarters.

Figure 4. Dynamics of fragmentation index



Note: Fragmentation index measures the average number of sentences, per thousand earnings calls, that mention at least one of the following keywords: deglobalization, reshoring, onshoring, nearshoring, friend-shoring, localization, regionalization. Data are obtained from NL analytics and are based on the methodology described in Hassan et al (2019).

As global tensions continue unabated, the financial consequences of fragmentation for firms may intensify, amplifying macro-financial turbulence. This could manifest in crossborder effects, including capital shifts away from exposed firms, reduced asset valuations, and heightened market volatility. ■

Alessandro D'Orazio is a PhD candidate in Statistics at the University of Rome 'La Sapienza', Fabrizio Ferriani is Deputy Head of International Financial Markets and Commodity Division at the Bank of Italy, and Andrea Gazzani is a Senior Economist at the Bank of Italy

References

Aiyar, S and A Ilyina (2023), "Geo-economic fragmentation and the world economy", VoxEU.org, 27 March.

Aiyar, S, D Malacrino and AF Presbitero (2024), "Investing in friends: The role of geopolitical alignment in fdi flows", European Journal of Political Economy.

Albrizio, S, J Bluedorn, C Koch, A Pescatori and M Stuermer (2023), "Sectoral Shocks and the Role of Market Integration: The Case of Natural Gas", AEA Papers and Proceedings 113: 43-46.

Attinasi, MG, L Boeckelmann and B Meunier (2023), "The economic costs of supply chain decoupling", ECB Working Paper.

Borin, A, G Cariola, E Gentili, A Linarello, M Mancini, T Padellini, L Panon and E Sette (2024), "Inputs in geopolitical distress: A risk assessment based on micro data", VoxEU.org, 28 January.

Campos, RG, J Estefanía-Flores, D Furceri and J Timini (2023), "Geopolitical fragmentation and trade", VoxEU.org, 31 July.

D'Orazio, A, F Fabrizio and A Gazzani (2024), "Geoeconomic fragmentation and firms' financial performance", SSRN Working paper.

Emiliozzi, S, F Fabrizio and A Gazzani (2024), "The European Energy Crisis and the Consequences for the Global Natural Gas Market", VoxEU.org, 11 January.

Feng, C, L Han, S Vigne and Y Xu (2023), "Geopolitical risk and the dynamics of international capital flows", Journal of International Financial Markets, Institutions and Money 82.

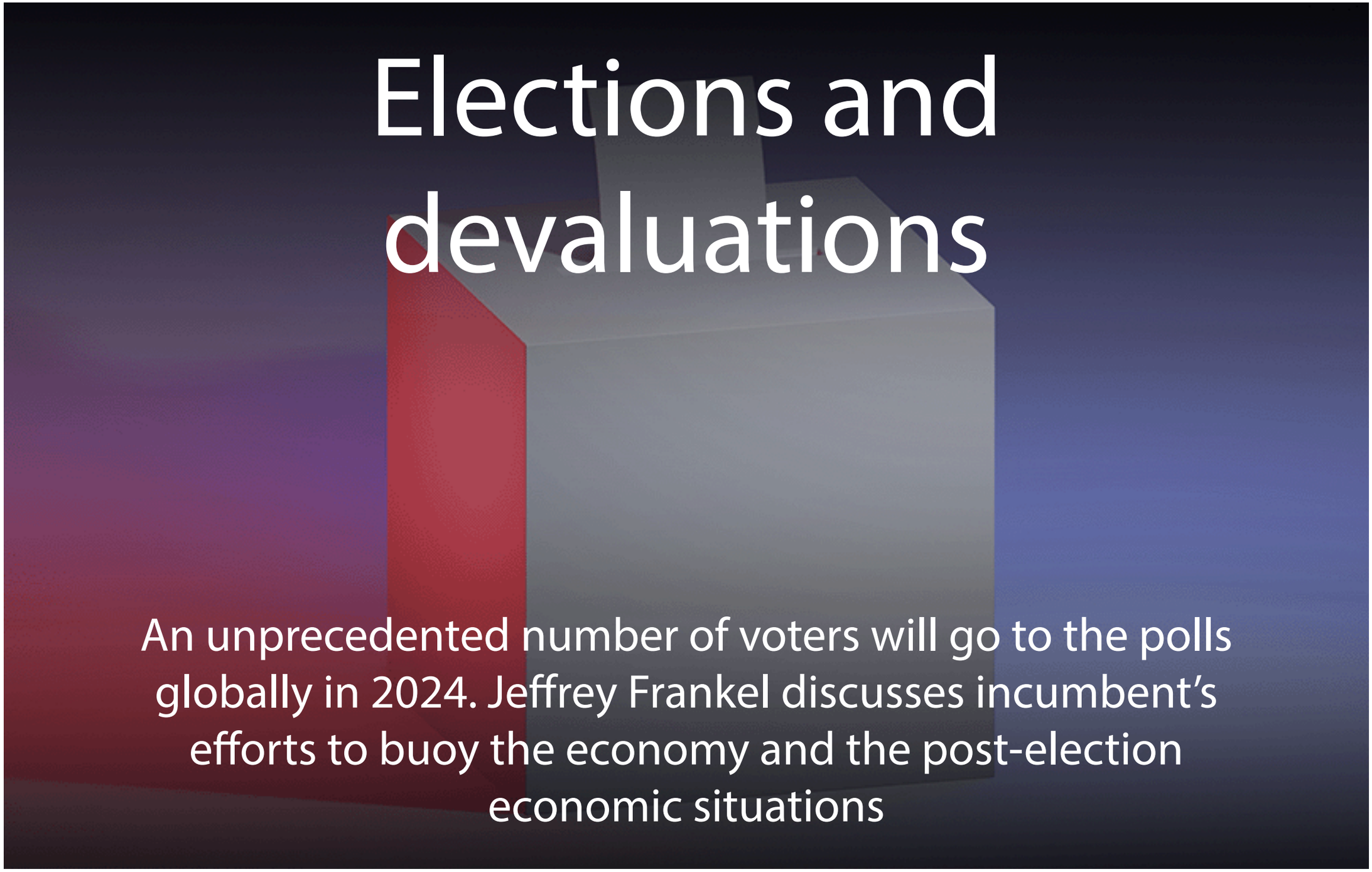
Hassan, TA, S Hollander, L Van Lent and A Tahoun (2019), "Firm-level political risk: Measurement and effects", The Quarterly Journal of Economics 134(4): 2135-2202.

IMF (2023a), "Fragmentation and Commodity Markets: Vulnerabilities and Risks", World Economic Outlook, October.

IMF (2023b), Geopolitics and Financial Fragmentation: Implications for Macro-Financial Stability, Global Financial Stability Report, April.

Authors' note: This column does not necessarily reflect the view of the Bank of Italy, the European Central Bank, or the European System of Central Banks. This article was originally published on VoxEU.org.

Elections and devaluations



An unprecedented number of voters will go to the polls globally in 2024. Jeffrey Frankel discusses incumbent's efforts to buoy the economy and the post-election economic situations

Lots of countries are voting, with 2024 an unprecedented year in terms of the number of people who will go to the polls. Recent elections in a number of emerging market and developing economies (EMDEs) have demonstrated anew the proposition that major currency devaluations are more likely to come immediately after an election, rather than before one. Indeed, Nigeria, Turkey, Argentina, Egypt, and Indonesia are five countries that have experienced post-election devaluations within the last year.

The election–devaluation cycle

Economists will recall a 50-year-old paper by Nobel Prize winning professor Bill Nordhaus as essentially initiating research on the political business cycle (PBC). The PBC refers to governments' general inclination towards fiscal and monetary expansion in the year leading up to an election, in hopes of the incumbent president, or at least the incumbent party, being re-elected.

The idea is that growth in output and employment will accelerate before the election, boosting the government's popularity, whereas the major costs in terms of debt troubles and inflation will come after the election.

But the seminal 1975 paper by Nordhaus also included the prediction of a foreign exchange cycle particularly relevant for EMDEs. That is the proposition that countries generally seek to prop up the value of their currencies before an election, spending down their foreign exchange reserves, if necessary, only to undergo a devaluation after the election.

Nordhaus wrote: *"It is predicted that the concern with loss of reserves and balance of payments deficits will be greater in the beginning of electoral regimes, and less toward the end.…The basic difficulty in making intertemporal choices in democratic systems is that the implicit weighting function on consumption has positive weight during the electoral period and zero (or small) weights in the future."*

The devaluation may be undertaken deliberately by an incoming government, choosing to get the unpleasant step – with its unpopular exacerbation of inflation – out of the way while it can still blame it on its predecessors. Or the devaluation may take the form of an overwhelming balance-of-payments crisis soon after the election.

Either way, a government has an incentive to hoard international reserves during the early part of its term in office, and to spend them more freely to defend the currency toward the end of its term.

Of course, the association between elections and the exchange rate is not inevitable. India is undergoing elections now and Mexico will in June. But neither seems especially in need of major currency adjustment

A political leader is almost twice as likely to lose office in the six months following a major devaluation as otherwise, especially among presidential democracies (Frankel 2005). Why are devaluations so unpopular that governments fear to undertake them before elections? In the traditional textbook model, a devaluation stimulates the economy by improving the trade balance. But devaluations are always inflationary in countries which import at least a portion of the basket of goods consumed.

Furthermore, devaluations in EMDEs often are contractionary for economic activity, particularly via the adverse balance sheet effects on those domestic borrowers who had incurred debts denominated in dollars.

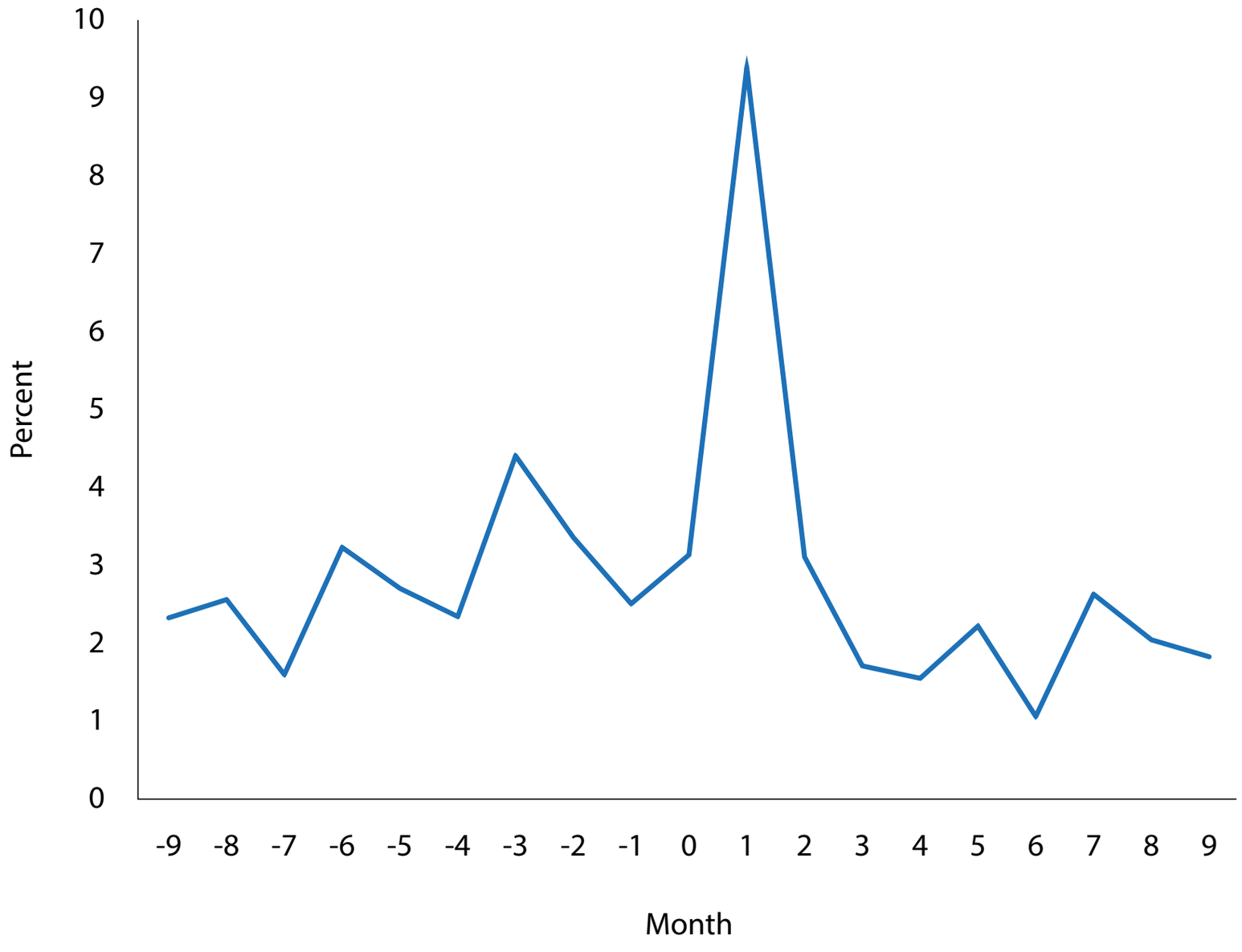
The theory of the political devaluation cycle was developed in a series of papers by Ernesto Stein and co-authors. One might think that voters would wise up to these cycles and vote against a leader who sneakily postponed a needed exchange rate adjustment. But given a lack of information about the true nature of the politicians, voters may in fact be acting rationally.

Figure 1, from Stein and Streb (2005) shows that devaluations are far more common in the immediate aftermath of changes in government. (The sample covers 118 episodes of changes, excluding coups, among 26 countries in Latin America and the Caribbean between 1960 and 1994)[1].

Some devaluations over the past year

Many EMDEs have been under balance-of-payments pressure during the last two years. One factor is that the US Federal Reserve raised interest rates sharply in 2022-23 and is now leaving them higher for longer than markets had been expecting. Consequently, international investors find US treasury bills more attractive than EMDE loans and securities.

Figure 1. Average devaluation pattern before and after elections



Source: Stein and Streb (2004).

A good example of the political devaluation cycle is Nigeria. Africa's most populous country held a contentious presidential election on 25 February 2023. The incumbent, who was term-limited, had long used foreign exchange intervention, capital controls, and multiple exchange rates to avoid devaluing the currency, the naira.

The new Nigerian president, Bola Tinabu, was inaugurated on 29 May 2023. Two weeks later, on 14 June, the government devalued the naira by 49% (from 465 naira/$, to 760 naira/$, computed logarithmically). It soon turned out that this was not enough to restore equilibrium in the balance of payments.

At the end of January 2024, the government abandoned its effort to prop up the official value of the naira, devaluing another 45% (from 900 naira/$ to 1,418 naira/$, logarithmically).

A second example is Turkey's election in May 2023. President Recep Tayyip Erdoğan had long pursued economic growth by obliging the central bank to keep interest rates low – a populist monetary policy that was widely ridiculed because of the president's insistence that it would reduce soaring inflation – while simultaneously intervening to support the value of the lira.

The government guaranteed Turkish bank deposits against depreciation, an expensive and unsustainable way to prolong the currency overvaluation. After the elections, the lira was immediately devalued, as the theory predicts. The currency continued to depreciate during the remainder of the year.

Next, on 19 November 2023, Argentina elected a surprise candidate as president, Javier Milei. Often described as a far-right libertarian, he comes from none of the established political parties. He campaigned on a platform of diminishing sharply the role of the government in the economy and abolishing the ability of the central bank to print money.

Milei was sworn in on December 10. Two days later, on 12 December he cut the official value of the peso by more than half (a 78% devaluation, computed logarithmically, from 367 pesos/$ to 800 pesos/$). At the same time, he took a chain saw to government spending such as energy subsidies rapidly achieved a budget surplus, and initiated sweeping reforms.

Argentine inflation remains very high, but the central bank stopped losing foreign exchange reserves after the devaluation, again as predicted by the theory.

A fourth example is Egypt, where President Abdel Fattah al-Sisi just started a third term, on 2 April 2024. The economy has been in crisis for some time. Nevertheless, the government had ensured its overwhelming re-election on 10-12 December 2023 by postponing unpleasant economic measures, not to mention by preventing serious opponents from running.

The widely expected devaluation of the Egyptian pound, came on 6 March 2024 depreciating 45% (from 31 Egyptian pounds/$ to 49 pounds/$, logarithmically). It was part of an enhanced-access IMF programme, which also included the usual unpopular monetary and fiscal discipline.

Finally, in Indonesia the widely liked but term-limited President Jokowi is soon to be succeeded by the Defence Minister Prabowo Subianto, who is less widely liked but was backed by the incumbent in the 14 February election. The rupiah has been depreciating ever since the 20 March announcement of the outcome of the contentious presidential vote. It fell almost to an all-time record low against the dollar on 16 April.

What next?

Of course, the association between elections and the exchange rate is not inevitable. India is undergoing elections now and Mexico will in June. But neither seems especially in need of major currency adjustment.

Venezuela is scheduled to hold a presidential election in July. As with some other countries, the election is expected to be a sham because no major opposition candidates are allowed to run. The economy is in a shambles due to long-time mismanagement featuring hyperinflation in the recent past and a chronically overvalued bolivar.

But the same government that essentially outlaws political opposition also essentially outlaws buying foreign exchange. So, equilibrium may not be restored to the foreign exchange market for some time.

To stave off devaluation, these countries do more than just spend their foreign exchange reserves. They often use capital controls or multiple exchange rates, as opposed to allowing free financial markets. That doesn't invalidate the phenomenon of post-election devaluations; it just works to insulate the governments a bit longer from the need to adjust to the reality of macroeconomic fundamentals.

Unfortunately, many of these countries also fail to allow free and fair elections, which works to also insulate the government from the need to respond to the voters' verdict. ■

Jeffrey Frankel is an Economist and Professor at Harvard Kennedy School

Endnote

1. Including Frieden and Stein (2001) and Stein and Streb (1998, 2004, 2005). More recently, Quinn et al (2023) find that voters punish leaders who devalue, in particular, when the currency was already undervalued. Steinberg (2015) finds that they are more likely to welcome a weak currency in countries where the manufacturing sector is powerful.

References

Broz, JL, M Duru and J Frieden (2016), "Policy Responses to Balance-of-Payments Crises: The Role of Elections", Open Econ Review 27(2): 207–227.

Frankel, J (2005), "Contractionary Currency Crashes in Developing Countries", IMF Staff Papers 52(2).

Frieden, J and E Stein (2001), "The political economy of exchange rate policy in Latin America: an analytical overview", in J Frieden and E Stein (eds), The currency game: exchange rate politics in Latin America, pp. 1-20.

Nordhaus, W (1975), "The Political Business Cycle", Review of Economic Studies 42(2): 169–190.

Quinn, D, T Sattler and S Weymouth (2023), "Do Exchange Rates Influence Voting? Evidence from Elections and Survey Experiments in Democracies", International Organization 77(4): 789-823.

Stein, EH and JM Streb (1998), "Political Stabilization Cycles in High-Inflation Economies", Journal of Development Economics, pp. 159-180.

Stein, EH and JM Streb (2004), "Elections and the Timing of Devaluations", Journal of international Economics 63(1): 119-145.

Stein, EH, JM Streb and P Ghezzi (2005), "Real exchange rate cycles around elections", Economics & Politics 17(3): 297-330.

Steinberg, D (2015), Demanding Devaluation: Exchange Rate Politics in the Developing World, Cornell University Press.

Author's note: A shorter version of this column appeared on Project Syndicate. I thank Sohaib Nasim for research assistance. This article was originally published on VoxEU.org.

Changing central bank pressures and inflation

The factors which facilitated low average inflation for decades have started to reverse. Hassan Afrouzi, Marina Halac, Kenneth Rogoff and Pierre Yared argue that the growing tensions between central banks and politicians will have negative consequences for economic activity

The global spike in inflation after the pandemic has led to vigorous debate on how quickly central banks will bring it down to target. However, no one really seems to worry about a repeat in the foreseeable future; market expectations of long-term inflation have largely come into line with central bank targets across advanced economies.

And no wonder; Stantcheva (2024) finds that the US public deeply dislikes inflation even though most individuals have only a tenuous grasp of what economists mean by it. True, there has been some discussion of allowing inflation to settle at a higher rate of, say, 3%.

This discussion predates the pandemic and largely centres around the question of whether the monetary authorities will ultimately decide they require a higher inflation target to cushion against the zero lower bound on interest rates that constrained central banks during the 2010s and into the pandemic. The ability of central banks to make their targets credible, however, is generally taken for granted.

In Afrouzi *et al* (2024b), we take a different view, arguing that the political economy environment in which central banks operate has changed markedly after the pandemic. In particular, factors that for decades had made it easier to maintain low average inflation, including globalisation, demographics, and fiscal restraint, may have gone into reverse.

If so, this could reawaken a dormant inflationary bias that markets and researchers had considered eradicated by the advent of central bank independence and inflation targeting. Our view does not necessarily imply a world where inflation is always above target.

Rather, it would be one where it is more difficult for central banks to battle democratically elected politicians who realise that the public dislikes higher unemployment, lower social spending, and higher taxes as much as it dislikes inflation.

We believe that higher average inflation is not inevitable, and that if governments choose to strengthen the independence of central banks sufficiently, they will be able to resist the political pressures

Whereas most central banks have great operational independence over the short run – and public criticism of them can backfire as Demiralp (2024) shows – over the long run, politicians control appointments, budgets, and, in many cases, the central bank's mandate.

Thus, our paper challenges the 'end of history' narrative of modern central banking which views inflation as a largely solved technocratic problem.

We are not the first to make such a suggestion. A widely discussed book by Goodhart and Pradhan (2020) argued that persistent structural changes in the global economy will keep future global inflation higher on average than in the past, although they do not offer an analytical framework for explaining why central banks would not simply contract monetary policy as necessary to achieve their inflation targets.

Rogoff (2003) anticipates their analysis by arguing that central bankers had the wind at their backs during the period of globalisation and more conservative fiscal policies. However, the static Barro and Gordon style model he employed required making ad-hoc assumptions about why globalisation might lower political economy pressures, and indeed did not include any real explanation of why long-run inflation itself mattered for economic activity.

Our 2024 paper, which offers a much-simplified version of Afrouzi *et al* (2023), aims to resolve these issues, and importantly offers a fully dynamic analysis. We reach several novel conclusions relative to both existing political economy models of monetary policy and the extremely large modern literature on New Keynesian models.

First, central banks' temptation to inflate comes from their concern for raising output and employment in the short run given monopoly distortions in the economy; this does not have to be an expressed concern of central bankers but rather can be transmitted through political pressures that interact with economic pressures.

Second, inflation itself is problematic because, in a dynamic New Keynesian model, higher inflation creates greater price disparities across firms who, unlike in the classical models of Milton Friedman and others, do not all coordinate on resetting prices and wages at the same time. These price disparities lead to misallocation and inefficiency in the economy.

From a technical perspective, our analysis is quite novel in that we solve our model analytically without having to linearise around zero inflation as is the ubiquitous practice in the academic literature. We show that the framework can be visualised as a simple diagram with long-run aggregate demand and supply curves, with an intersection point corresponding to steady-state inflation and output, as displayed in Figure 1.

The framework captures that although money is neutral in the New Keynesian model (a one-time monetary shock has no real effects), it is not 'super-neutral': the steady-state level of inflation affects the steady-state level of real output. The cost of higher inflation via price dispersion may be relatively modest if inflation only rises from 2% to 3% in the baseline model, though Afrouzi *et al* (2024a) provide an analysis that suggests this cost is non-trivial in a model with production networks.

Moreover, if one considers inflationary pressures as emerging via occasional (if rare) bursts of much higher inflation, then the average costs of the ensuing distortions are much greater.

Another interesting result we can demonstrate in the non-linear model is inflation overshooting, as depicted in Figure 1. When contractionary supply shifts lead to a higher average steady-state level of inflation, the short- and medium-term rates of inflation can be expected to substantially overshoot the long-term rate.

This means that even if median inflation remains at 2%, concern over such bursts can imply higher long-term interest rates (including say for mortgages and car loans) even in normal times.

Figure 1. Illustration of supply shock in long-run aggregate supply/long-run aggregate demand framework

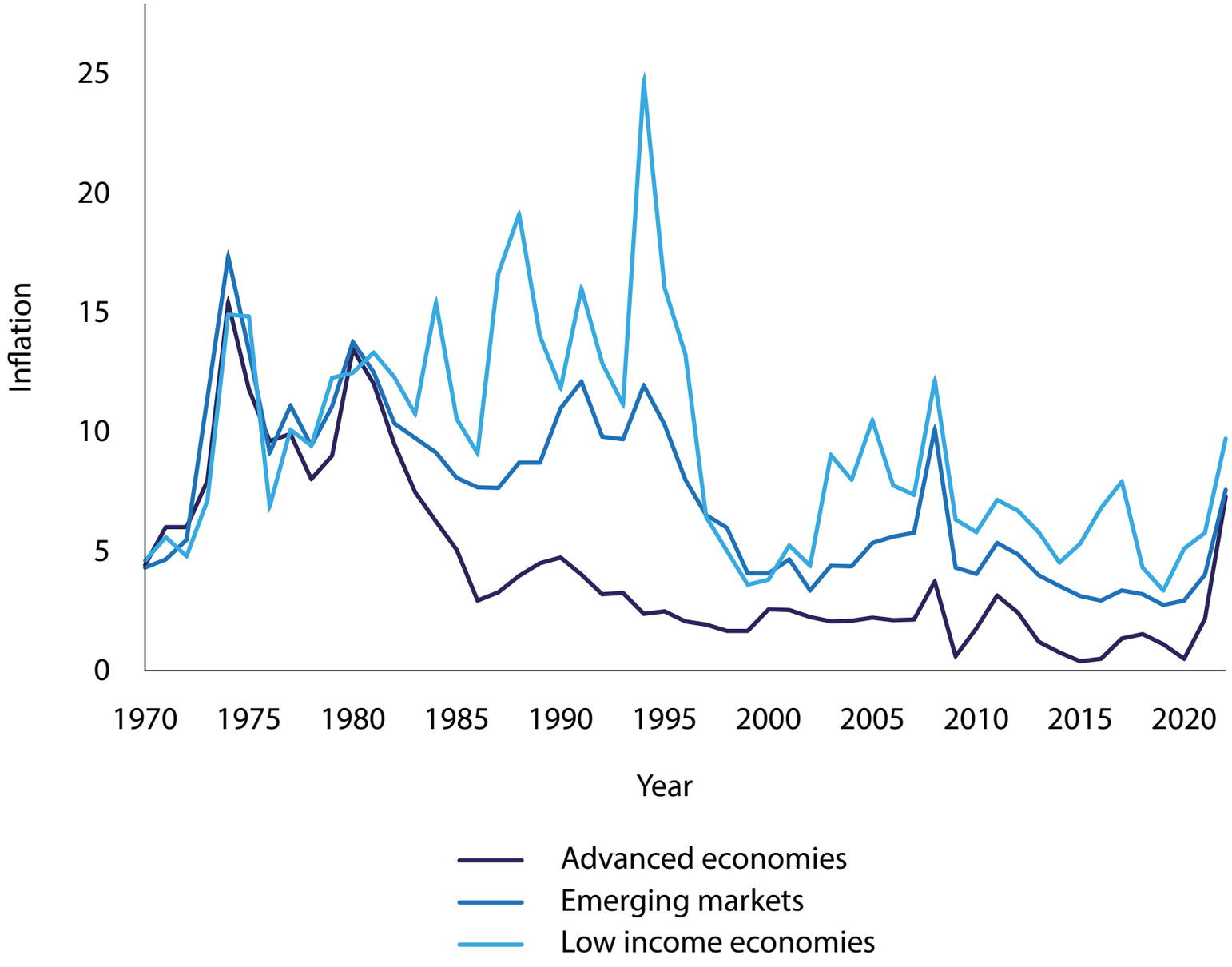
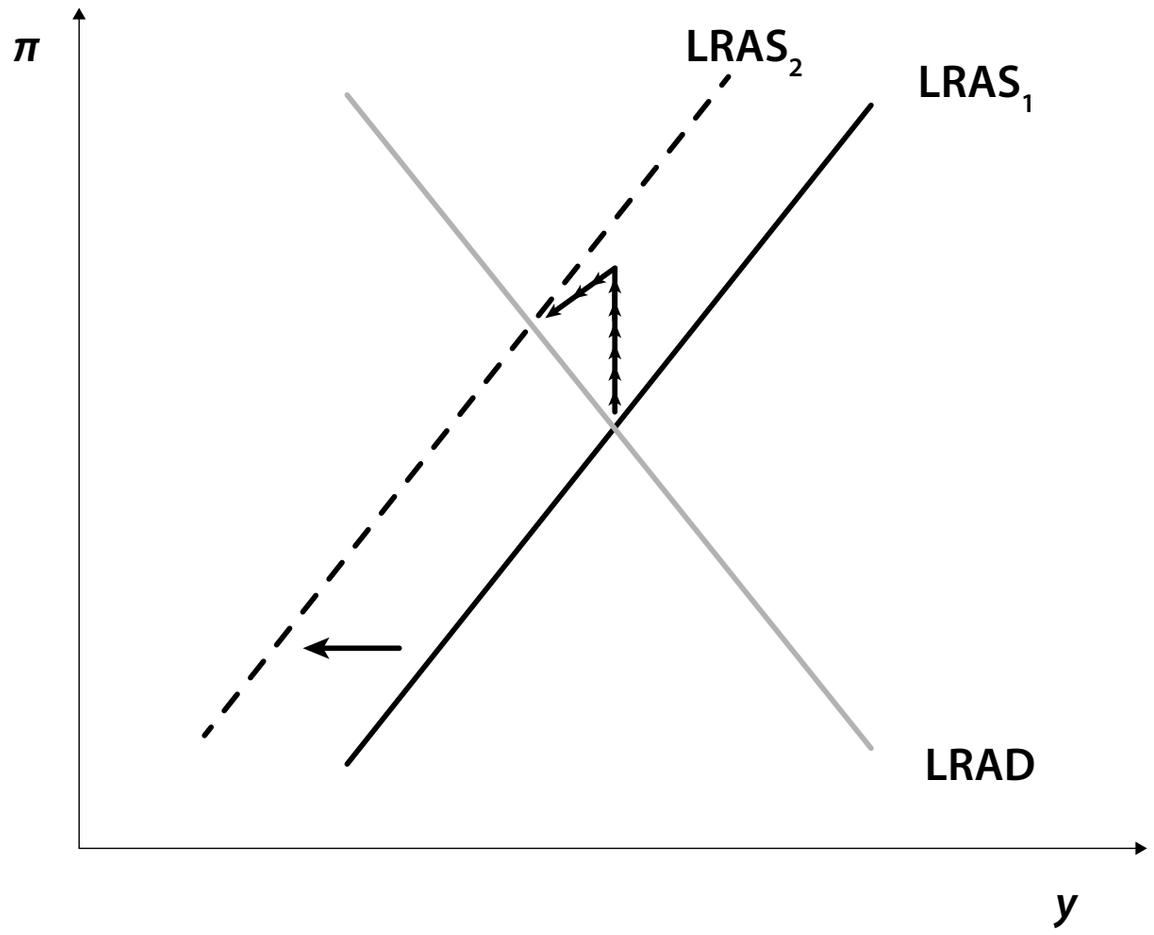


Figure 2. Inflation across the world



Source: Afrouzi et al (2024b) based on data from Ha et al (2021).

We recognise that our theoretical model implies predictions that are very different from the consensus discussed at the outset of this piece. Although we cannot formally test the model, our paper does revisit the disinflationary period of the 1980s through 2010s that occurred throughout the world, albeit with different timing across countries and country groups; see Figure 2.

In our paper, we describe how changes in globalisation, market liberalisation, fiscal policy, and unionisation helped support disinflation over much of this period. Many of these factors—which worked in tandem with the expansion of central bank independence over this period—may be reversing today.

We also note that the era of a zero lower bound on interest rates may have masked inflationary tendencies for an extended period, precisely because central banks lacked an effective instrument for inflating.

Now, with some arguing that a considerable portion of the recent rise in long-term real interest rates is a return to trend (Rogoff *et al* 2022, 2024) and with the experience of the pandemic that chastised bond markets, the zero lower bound may be a less frequent crutch for anti-inflation credibility in the future.

We believe that higher average inflation is not inevitable, and that if governments choose to strengthen the independence of central banks sufficiently, they will be able to resist the political pressures.

Despite the considerable progress of the past couple decades, as Romelli (2024) documents, further such progress appears unlikely in most countries given rising populism, not to mention enormous budgetary pressures from rising defence needs, the green transition, and servicing today's very high debt levels.

Thus, even if central banks do eventually get inflation back to target this round, the inexorable rise in political economy pressures implies that occasional bouts of inflation are likely to be a bigger risk than bouts of deflation in the coming decade, with overshooting and volatility becoming the new normal. If not contained, then, as Braggion *et al* (2023) show, the effects may be felt for generations. ■

Hassan Afrouzi is Assistant Professor of Economics at Columbia University, Marina Halac is the Stanley B Resor Professor of Economics at Yale University, Kenneth Rogoff is the Thomas D Cabot Professor of Public Policy and Professor of Economics at Harvard University, and Pierre Yared is a Professor at Columbia University

References

Afrouzi, H, M Halac, KS Rogoff and P Yared (2023), "Monetary Policy without Commitment", NBER Working Paper 31207.

Afrouzi, H, S Bhattarai and E Wu (2024a), "The Welfare Cost of Inflation in Production Networks", Working Paper.

Afrouzi, H, M Halac, KS Rogoff and P Yared (2024b), "Changing Central Bank Pressures and Inflation", NBER Working Paper 32308.

Braggion, F, F von Meyerinck, N Schaub and M Weber (2023), "The long-term effects of inflation on inflation expectations", VoxEU.org, 13 September.

Demiralp, S (2024), "Silence is Golden: How Public Criticism of Central Banks can Backfire for Leaders", VoxEU.org, 18 April.

Goodhart, C and M Pradhan (2020), The Great Demographic Reversal: Aging Societies, Waning Inequality, and an Inflation Revival, London: Palgrave-MacMillan.

Ha, Jongrim, M Ayhan Kose and F Ohnsorge (2021), "One-Stop Source: A Global Database of Inflation", World Bank Policy Research Working Paper 9737.

Rogoff, K (2003), "Globalization and Global Disinflation", Federal Reserve Bank of Kansas City Economic Review 88, no. Q IV: 45-78.

Rogoff, K, B Rossi and P Schmelzing (2022), "Was the post-global financial crisis collapse in real interest rates secular?", VoxEU.org, 24 October.

Rogoff, K, B Rossi and P Schmelzing (2024), "Long-Run Trends in Long-Maturity Real Rates 1311-2022", American Economic Review, forthcoming.

Romelli, D (2024), "Recent Trends in Central Bank Independence", VoxEU.org, 26 February.

Stantcheva, S (2024), "Why do We Dislike Inflation?", Working Paper, Harvard University.

This article was originally published on VoxEU.org.