

Balancing the productivity opportunities of financial technology and AI against the potential risks – speech by Randall S. Kroszner

Given at City Week 2024

21 May 2024

Speech

It's a pleasure to be at CityWeek 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.

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](#)

In this speech, 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 conducive 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

⁶ AWS Whitepaper, [Interpretability versus explainability](#)

⁷ See [Bank of England, The Bank of England’s supervision of financial market infrastructures Annual Report 2023](#)

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

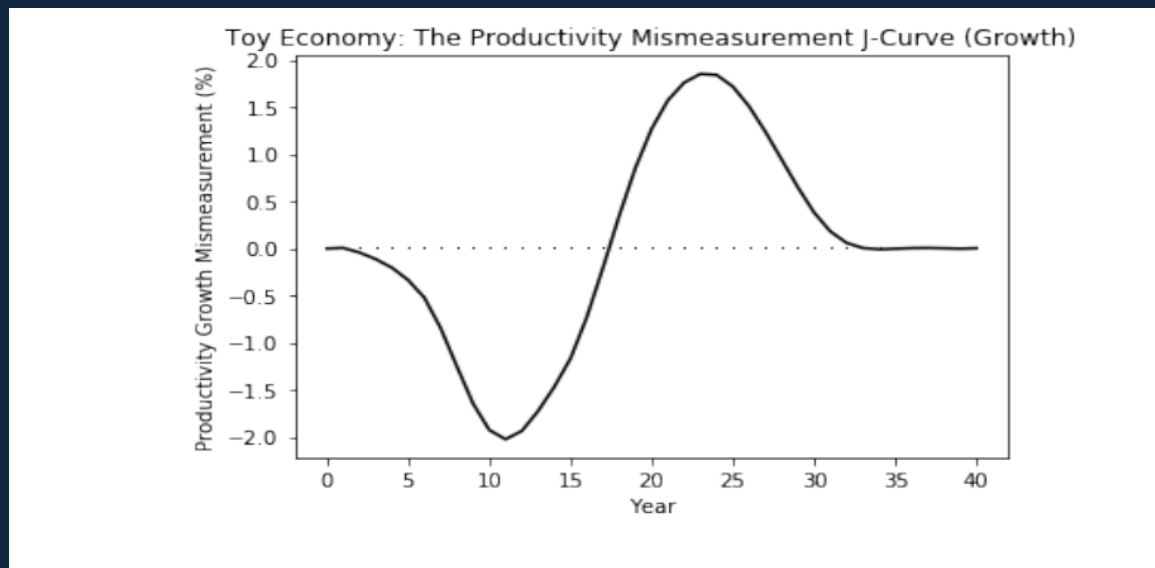
⁸ Briggs and Kodnani, (2023), [‘The Potentially Large Effects of Artificial Intelligence on Economic Growth’](#)

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’](#)

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](#)

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 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. My colleague Sasha Mills will be saying more about this later today.

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

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 U.S. 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’](#)

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, just a couple of weeks ago 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.

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.

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.

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'](#))

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 – and I look forward to that continuing in conferences such as this one.

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.