



ARTICLE

Enacting efficient markets: The performativity of quantitative easing

Dylan Cassar 

University of Malta, Msida, Malta
Email: dylan.cassar@um.edu.mt

Abstract

Quantitative easing (QE) has been a favourite tool of central banks in their post-financial crisis monetary policy apparatus. Social science literature has interpreted QE as a shift away from performative governance characterising pre-crisis monetary policy. With reference to the Bank of England's experience, I offer a reinterpretation of QE as a performative intervention in the conditions of financial markets, as an attempt to alter the state of financial markets away from dysfunctionality and towards efficiency. I claim that, following the financial crisis, the model of complete and efficient markets – a mainstay in central banking prior to the crisis – was transformed from a real-world approximation to a 'performative object' to be achieved. In deploying the balance sheet, central banks attempt to performatively enact complete and efficient markets. The article rejects the claim of discontinuity between pre-crisis and post-crisis monetary policy, arguing that QE is a continuation of inflation targeting though with important innovations. While pre-crisis performativity relied on central bankers' communicative framing of market expectations, QE is performative via the ontological shaping of financial markets, driven by epistemic models. The article relies on a set of 51 interviews with central bankers and financial market participants and a corpus of documents.

Keywords: central banks; financial markets; performativity; monetary policy; quantitative easing

Introduction

In August 2020, at the Jackson Hole Economic Policy Symposium, the Governor of the Bank of England Andrew Bailey (2020) suggested that the effectiveness and transmission mechanisms of monetary policy – specifically quantitative easing (QE) – are highly contingent on the 'state of the world' (i.e., the conditions of financial markets) in which it functions. Two years earlier, Ben Broadbent, the Deputy Governor, had remarked that the states of the world can change and immediately proceeded by 'One might even say this is intrinsic to the effects involved' (Broadbent, 2018: 10). In other words, not only is QE contingent on the state of the world, but that indeed it can *alter* the states of the world in which it is deployed. These signify that the very intervention of central banks can shape the states of the world in which it is implemented. QE is, as such, performative.

The performative powers of central banking have been put under the scrutiny of sociology and political economy scholarship. At the heart of the performative powers of central banks stand expectations, a fundamental feature of modern capitalism (Beckert, 2016). Central banks are said to have developed a performative form of governance by

which they bring private (financial) market expectations under control (Braun, 2015; Wansleben, 2018). This form of performativity operates by virtue of central bank's intervention in the economy through utterances, such that the economy is *made* 'as a communicative field and as an empirical fact' (Holmes, 2013: 21). While this is tantamount to reducing central banks' performativity to a sovereign speaker, namely, the Governor, whose utterances bring about the effects of which they speak, more recent scholarship has delineated the broader apparatus through which these utterances find their performative force in framing expectations (Braun, 2015; Wansleben, 2018; Walter and Wansleben, 2020). Braun (2015) claims that beyond mere verbal utterances, central banks developed a communicative *apparatus* by means of which central banks frame the future, and financial market participants' expectations are performed. Private market expectations became 'a performative effect of the interaction between the central bank and [financial market actors]', such that by influencing them they influence short-term interest rates and a whole swathe of other related financial assets, all the way into the 'real economy' by shaping the macroeconomic conditions for household consumption and firm's investment decisions, in order to control inflation. This apparatus was erected on epistemic and material technologies and techniques through which the economy is modelled and rendered governable. What has come to be known as the *performative* governance of central banking in this literature, then, is often claimed to be able to explain the Great Moderation as a period in which central banks held sway over expectations and could thus hold control over inflation rates.

In turn, post-crisis central banking has been interpreted as heralding a schismatic break with the past. Braun (2018) has claimed that QE entails a return to a hydraulic form of state governance, defining the implementation of QE as the key moment in which central banking broke away from performative governance. Similar to post-war macroeconomic policy, which took the form of Keynesian mechanical lever-wielding on the macroeconomy (hence, hydraulic), QE-led central banking is seen to involve the leveraging of the balance sheet through asset purchases, which pushes upwards their prices (and yields downwards) mechanically. The central bank 'purchase[s] financial assets in order to stimulate the financial economy directly' as a hydraulic macroeconomic agency, though this time adapted to a financialised economy, and hence fundamentally different from the 1990s/2000s central bank governance that was 'communicative and performative' (Braun, 2018: 207). Hence, QE entails a shift away from performative governance and a return to hydraulic governance.

In this article, I argue that there is more continuity between pre-crisis and post-crisis than previously believed, with central bank interventions remaining performative, albeit through a different modality of performativity from pre-crisis monetary policy. I show how the claim to schismatic discontinuity between pre-crisis and post-crisis central banking hinges on a restrictive conception of performativity – one which is limited to communication-based expectational governance. This article proposes instead a broader view of performativity, one which speaks to the relationship between the epistemic and socio-technical models adopted by central banks and enrolled in the practices of governance, as well as the particular textures of financial markets (i.e., states of the world), which said models are intended to represent. In doing so, the article highlights the continuities and discontinuities between pre-crisis and post-crisis monetary policy. While this entails going back to the classic version of performativity laid out by Callon (2007) and MacKenzie (2006), it also means paying closer attention to the alignments (and misalignments) between central banks' operative governance and financial markets as proposed by Walter and Wansleben (2020).

Indeed, the pre-crisis mode of communicative governance was contingent on a process of alignment between the conditions of financial markets and the governance practices and the socio-technical tools enrolled within them. Walter and Wansleben's (2020)

influential work has laid out the ways in which central banks were involved in shaping the very textures of financial markets while building their edifice of governance in alignment with those conditions of financial markets. For instance, the Bank of England embarked on institution-building in the 1990s through the development of repo markets in gilts that allowed the enactment of deeper and liquid markets. But it also developed the socio-technical devices through which it could ‘see’ the market, such as the government bond yield curve as the representation of the gilt market (Cassar, 2022), and dynamic stochastic general equilibrium (DSGE) models through which the economy (and the impact of monetary policy) becomes calculable (Braun, 2014; Wansleben, 2023). These were enrolled in practices and routines of governance, such as inflation forecasting, and their communication to financial market participants via inflation reports and regular monetary policy conferences. Communication became a key practice of governance for central banks as they realised that the very *announcement* of future policy rates often had an effect on market prices as market participants would adjust their expectations, which come to be reflected in interest rates (Krippner, 2011).¹

Using the case of the Bank of England, I show how post-crisis central banking required central banks to perform alignment work along these lines, as the texture of financial markets detached from the underlying model operative in monetary policy. I therefore offer a reinterpretation of QE as a performative intervention in the very texture of financial markets as these come to be shaped into the image of the model. Walter and Wansleben (2020: 647) have offered a similar interpretation, in which QE is an attempt to ‘safeguard and restore the technical presuppositions of inflation targeting’. While they do not provide any further explication, their point rests on the argument that central banks are tied to market-based finance because it is through that infrastructure that they can govern. Hence QE would entail an attempt to support these infrastructures such that central banks can govern through them. Similarly, Wansleben (2023: 220) claims that central banks have sought to repair dysfunctional markets in various ways, while simultaneously attempting to provide monetary accommodation to a real economy stuck in a persistently low inflation.

I follow Wansleben’s (2023) and Walter and Wansleben’s (2020: 647) continuity argument by which unconventional policy ‘constitutes not so much a (temporary) return to an older form of ‘hydraulic’ economic policy (cf. Braun, 2018) or a direct attempt to subsidise finance (Jacobs and King, 2016). Rather, through QE, central bankers have sought to safeguard and restore the technical presuppositions of inflation targeting’. But while Wansleben’s (2023) work channels the recent focus of the social science literature on infrastructures (Braun, 2020; Coombs, forthcoming; 2022), it glosses over the place of economic theory and models in central banks’ edifice of monetary policy, as recently argued by Coombs (2024). Paying attention to economic theory and models is necessary to properly elucidate the performative dimensions in the relationship between programmes of governing and financial markets.

Indeed, I claim that, as central banks sought to restore the edifice of inflation targeting, the contingent intervention of QE was driven by an *epistemic model* of financial markets, hence revealing its performative dimension. While QE was truly an attempt to stimulate the economy by lowering yields, it was also intended as a performative intervention in financial markets in order to shape them in the image of a particular model. Following the financial crisis, the model of complete and efficient markets – embedded deeply in the epistemic and socio-technical procedures of central banks – was transformed from a real-world approximation to a ‘performative object’ to be achieved. QE is therefore an attempt to alter the state of financial markets away from dysfunctionality (e.g. in crisis times) and towards efficiency – that is, to performatively enact complete and efficient markets.

From this viewpoint, post-crisis monetary policy marks less of a schismatic break than previously believed. However, the two periods also differ on two counts: in the *process* of

alignment between models and markets and in the *kind* of performativity operative. First, unlike the 1980s and 1990s, when the tools of governance (e.g. socio-technical devices) were built to *reflect* the texture of financial markets in the erecting of the inflation targeting regime, after the crisis it was the epistemic tools of governance themselves that were driving interventions into financial markets proper – that is, an inverse process where financial markets themselves come to reflect the model. In classic performative fashion, the model is not a camera, but an engine (MacKenzie, 2006) – one which served to restore the alignment between the model and market.

Secondly, the mode of performative governance following the crisis differed from pre-crisis performativity. The pre-crisis communicative form of governance relies on a network of actors acting as centres of calculation whose behaviour (e.g. expectation formation) is shaped by the central banker's utterances and forecasts within a shared epistemic world and relying on another set of actors – the arbitrageurs – whose conduct transmit the effects of monetary policy far and wide. However, the repeated breakdown of this shared world following the crisis impairs the functioning of monetary policy as the arbitrageur and the market-maker face liquidity constraints. The form the financial market takes – incompleteness and inefficiency – is no longer accurately represented by the inscriptions in the models. The performative power of a QE intervention lies in its world-making functions driven by the epistemic model – that is, an *ontological* shaping of the network. The central bank reorganises the network as it shifts (temporarily) to the centre of the network, absorbing risk and channelling liquidity and confidence to the rest of the network, modifying the network's behaviour in the process. The texture of financial markets is hence realigned with the models underpinning monetary policy, which are enacted performatively. I exemplify this with two instances when the Bank of England intervened in financial markets – the first being March 2020's 'dash for cash', and the second refers to developments following Liz Truss' mini-budget announcement – in which the Bank sought to restore market functioning and bring markets closer to efficiency.

As the financial crisis disrupted financial markets, the Bank was compelled to alter its model of markets. While post-crisis monetary policy could have implied the *permanent* abandonment of the model of efficient markets, central banks instead converted the pre-crisis model into a 'performative object' to be *achieved*, particularly in those moments when real-world conditions threatened the permanent failure of the model. In deploying the balance sheet as part of QE, the Bank attempts to performatively foster and enact complete and efficient markets. I show how central banks' modelling of the state of the world compels them to craft specific interventions by which they can alter those states of the world. In other words, the various theories and knowledge claims adopted by central banks, as embedded in epistemic devices, drive contingent performative interventions in their respective state of the world, and the very interventions are aimed towards enacting a particular state of the world. This process restores the pre-crisis modality of performativity, primarily expectation governance. Central banks act, therefore, as market sculptors in their own right, as they attempt to restore a particular state of the world (efficient markets) and bring into being the conditions under which communicative-based performative governance works best.

Data and methods

The article relies on a set of 51 interviews with central bank personnel and financial market participants primarily in the UK and Germany (see Table 1). Twenty-one of these interviews were carried out with central bankers, primarily Bank of England personnel, including former members of the Monetary Policy Committee and staff in the Markets and Monetary Policy divisions of the Bank of England. Thirty interviews with financial market

Table 1. Interviews broken down by professional and organisational category.

Current/former members of central bank decision-making bodies (all Bank of England)	11
Central bank staff economists (Bank of England and ECB)	10
Portfolio managers in asset management	9
Traders/market-makers in investment banks	8
Traders in hedge funds	8
'Quants' in investment banks	5

participants in the UK and Germany were also highly informative. Of note are a significant number of interviews with interviewees whose role at the time of the interview was that of a market participant but who also held a central banking experience and the inverse for other interviewees. Importantly, these interviews provided rich detail spanning the fluid 'boundaries' of central banks and financial markets (Coombs and Thiemann, 2022), with interviewees discussing in detail the oft-concealed relations between central banks and markets.

This article draws more directly on interviews with central bankers, primarily those on the Bank of England's Monetary Policy Committee and those central bankers in the engine rooms of central banks who design and run the models, especially DSGE models. In many cases, the latter are experts on DSGE modelling and have spent years working on, tweaking, and in some cases even revamping these models given that their experiences span the pre-crisis and post-crisis periods. Equally important are those interviews with traders in hedge funds and in investment banks who experienced the market disruptions of the financial crisis and the March 2020 debacle. The rest of the interviews informed the article indirectly.

The interviews were conducted between April 2019 and January 2021. Those interviews conducted between April 2019 and March 2020 were done in person, while those interviews conducted following the onset of the pandemic were carried out via audio/video call. All interviewees, irrespective of seniority and position, were guaranteed full anonymity. Interviews lasted between 28 and 180 minutes, with the average interview taking 69 minutes. During the interviews, I asked the interviewees for views about the relationship between central banks and financial markets and the role of the socio-technical infrastructures that constituted this relationship. In this light, I also asked how central bankers and market participants made sense of the relatively new policies known as unconventional monetary policies, including QE and forward guidance, among others, and how these shaped their practices.

I triangulate my interview data with a large set of documents from digital archives and the Bank of England website, ranging from 2009 to 2023, primarily involving Bank of England documents. These include Monetary Policy Committee meeting minutes, speeches given by MPC members, *Inflation Reports* and *Monetary Policy Reports*, press conference transcripts, public statements and letters, research papers, *Asset Purchase Facility Reports*, *Quarterly Bulletins*, oral evidence to Treasury Committee, and other relevant documents.

I use the case of the Bank of England because it was one of the first central banks in which the programme of expectation governance was erected alongside the development (and making) of deep and liquid financial markets during the 1980s and 1990s (Wansleben, 2018), and it was also one of the first central banks to turn to QE in response to the financial crisis of 2007/08, which led to internal conflicts on the form QE would take and to undertake significant relational and political work around it (Cassar, 2024). The process of scientisation of the Bank of England since the 1990s has also meant alignment work

between its socio-technical devices and financial market developments, making it an appropriate case study for an investigation of post-crisis monetary policy. Its generative mould, as a technocratic central bank implementing procedures and policies in line with other central banks, also allows a degree of generalisability to other central banks across the world, though contextual and institutional conditions will necessarily differ, and in spite of comparative assessments not being within the scope of this article.

The model of efficient markets in pre-crisis monetary policy

This section explores how conventional monetary policy – revolving around the performativity of expectations – was conditioned on a model of complete and efficient financial markets and on this model being an acceptable real-world approximation for the central banker. While a lot of work has been done on how central banks govern market expectations (Holmes, 2013; Braun, 2015; 2014; Wansleben, 2018), the role of the model of efficient markets as the backbone of this policy arrangement has not been given much attention. I show how, in treating their model as an acceptable approximation of reality, central banks came to rely on ‘the arbitrageur’ *from a distance*. That is to say, central bankers assigned much of the work in the transmission mechanism of monetary policy to the arbitrageur (i.e., fixed income relative value traders such as those in leveraged relative value hedge funds). This analysis serves as pre-ambles to understand precisely how post-crisis QE implementation implied a performative intervention in financial markets.

The arbitrageur is a direct descendent of modern financial theory developed in the second half of the 20th century, and the no-arbitrage ‘epistemic culture’ to which this gave rise in financial markets (MacKenzie, 2006; MacKenzie and Spears, 2014). As MacKenzie (2006) shows, these arbitrageurs enacted market efficiency as they exploited deviations from their theoretical models embodying efficient markets. In exploiting deviations, therefore, arbitrageurs pushed market prices towards the theoretical prices as exemplified by models. The socio-materiality of arbitrage, particularly of the kind this article deals with, has been the subject of study in the social studies of finance, in what has been called ‘a material sociology of arbitrage’ (Beunza and Stark, 2004; Beunza et al., 2006; Beunza, 2019). Because it is not within the scope of this article to take up a discussion on the murky definitions of what constitutes arbitrage (see Hardie, 2004, for such a discussion), for the purposes of this article, I will restrict the term to the way market participants and central bankers themselves refer to it: any form of market-neutral trading that may involve (non-directional) risk and that seeks to profit from price discrepancies between two (or more) securities that in principle should trade close to each other.

So how do arbitrageurs enact complete and efficient markets, thus supporting the central bank’s governance of expectations? The answer lies in a sociomaterial device that is a key tool on the desk of the arbitrageur – the yield curve (Zaloom, 2009; Cassar, 2022), a tool around which *both* central banks and markets are cognitively and materially entangled. The yield curve is a technical illustration of yield levels and a representation of market pricing in bonds and their derivatives. The graph is a comparative representation of yields on similar-quality bonds with different maturities. As a representation of market pricing, most of the analysis and activity happening in the bond market revolves around it (Zaloom, 2009; Choudhry, 2019).

The yield curve is the arbitrageur’s world, existing as a ‘material system’ in the sense of mechanical engineering or physics – a mathematical space holding a set of relationships along the curve, that should, in theory, be mean-reverting (Interview 32). For the arbitrageur, what matters is not the future direction in which yields move (e.g. via a central bank announcement). Indeed, the latter would be the remit of *directional* traders who place trades on the future yield (or price) of a security, such as a bond, which action is entirely reliant on the formation of expectations which central banks seek to govern. (It is

to these directional traders, or central bank watchers, that the social studies of central banking refer to when they speak about market participants and their expectations.) On the directional trader's desk, the yield curve functions as an agglomeration of market expectations, along or against which a view (or a bet) is expressed. However, on the arbitrageur's desk, the yield curve is ontologically an entirely different object, one that represents statistical and mathematical relationships between yields. For the arbitrageur, then, what matters is that yields move in some predictable fashion *relative* to similar securities, hence *relative value* (Interview 25).

While there are multiple yield curves (swaps curve, money market curves, government bond market curves, etc.), it is often the bond market yield curve that concerns central bankers. But this restricted focus is only possible because of the existence of arbitrageurs who connect yield curves together (Woodford, 2003) via the analysis and trading-based exploitation of spreads (Beunza and Muniesa, 2005) across curves. Hence the transmission of monetary policy can move relatively seamlessly across financial markets. They do so by exploiting (and therefore bringing about) converging relationships, or associations (Beunza and Stark, 2004), between on-the-run and off-the-run bonds; basis trades between government bonds and futures; and spreads between government bonds and swap curves. The result of such trading is that different markets become connected, not unlike the socio-materially constituted 'world market' that Beunza et al. (2006) speak of. Central bank's action on short-term interest rates will, *in theory*, reverberate to other markets (see also Gabor, 2014) because of the work of the arbitrageur. Indeed, this is an aspect of the *perlocutionary* dimension of monetary policy which Walter (2024) emphasises and which has often been neglected in the literature.²

Central bankers could then adopt this conception of markets as an acceptable 'real-world approximation' underlying their policy arrangement. This model is inscribed in the material devices through which central banks make policy. One interviewee who has worked for years in modelling at the Bank claimed that it is precisely for this reason that central banks' key macroeconomic models of DSGE did not, at least not until after the crisis, incorporate the financial and banking systems within them (Interview 17). In this model world, financial markets are absent because they are treated as largely efficient and frictionless, and hence there is no need for them – in other words, the transmission of monetary policy would seem to happen efficiently and instantaneously, and hence one could abstract away financial markets in the model (Woodford, 2003; Dale, 2010; Braun, 2014; Bailey et al., 2020). DSGE models involve a range of models, and many participants cautioned against lumping all kinds of DSGE models into one basket. There are some, for instance, that incorporate different kinds of financial frictions in them. Others, like those dominant in central banks prior to the crisis, however, do not. These come in the form of an Arrow-Debreu style of models, in which markets clear instantly and in efficient ways. They include an intertemporal real interest rate channel that is powerful and which can be split into inflation and policy expectations (Interview 4; see also Braun, 2014).

The very fact that DSGE models did not incorporate frictions also suggests that central banks largely assumed away those frictions. In practice, central bankers are well-aware that markets are never fully complete and efficient and always entail some degree of frictions (Shleifer and Vishny, 1997; Nyborg and Östberg, 2014). The very existence of arbitrageurs would suggest that markets are not completely arbitrated, since arbitrageurs depend on the identifying of frictions in order to waive them away via trading. And the transmission mechanism of central bankers, therefore, is never instantaneous. Even so, central bankers often accept these marginal frictions if they are able to deem markets as efficient and complete *to an approximation*.

For central bankers, the degree to which markets approximate perfectly complete and efficient markets is largely down to their assessment of the leverage abilities of arbitrageurs. Because price deviations are typically small, the arbitrageur needs to deploy

several magnitudes of leverage to extract reasonable profit. Leverage is crucial for arbitrage trading; leverage-constrained market participants would not exploit the arbitrage opportunities because the profits generated from such trades are so low that it is only leverage, which makes these trades worth exploiting. In fact, central bankers tie the notion of complete and efficient markets directly to leveraged arbitrage. As one central banker told me:

It's the efficient market hypothesis. So one way of characterising that would be, financial market prices are such that there are no arbitrage opportunities left. But one of the ways that that would happen in reality would be there's no constraints for someone in financial markets taking an infinitely large or arbitrarily large position in order to take advantage of those arbitrage opportunities. In reality... even the largest financial institutions have got leverage constraints, [but] then it comes down to the debate about how big they can be. Some people will say, look, there are massive hedge funds out there. Sure, they can't take infinite positions, but they can take really, really big ones. And so it's only going to be a few basis points that anything can be away from where it should be. It's an approximation. (Interview 8)

Their assessment of leveraged arbitrage, and the model of complete and efficient markets, allows central bankers to restrict their focus to market expectations. The yield curve itself, taken as a representation of expectations, has been enrolled into the expectations management policy arrangement of pre-crisis central banking. In other words, the yield curve became the world of the Bank's monetary policy, supported by an interpretive space that redefined the yield curve in terms of expectations. Assessing that arbitrage relations are functioning allows central banks to engage in what Jessop (2010) calls complexity reduction, by carving and opening up an epistemic space around the yield curve of the government bond market and the expectations hypothesis. Rather than being overly concerned with multiple markets, with a multitude of expectations and how to tame them, and with how that is transmitted to the 'real economy', they could therefore assume away the transmission mechanism of monetary policy 'outside' of the government bond yield curve. Concealed beneath the Bank's expectations management policy arrangement, therefore, was an understanding that a muted process was labouring in the background, led by arbitrageurs whose trading made markets more complete and efficient, and inscribed in the very models that render the economy visible and the governance of expectations calculable.

Modelling incomplete and inefficient markets in post-crisis monetary policy

The central banks' policy arrangement of expectation governance met its limits at the height of the 2007/08 financial crisis. As central bankers pushed their policy rates close to zero, they hit the zero-lower bound. The model by which central bankers operated, as argued in the previous section, was a New Keynesian DSGE model which incorporated an idea of efficient and complete markets (by omission), and importantly, one in which only interest rates would have some bite on the real economy (Woodford, 2003). Given that the policy rate was at zero, central bankers needed to turn to other policy tools for monetary policy purposes.

This required central bankers to design a new policy toolkit, which itself depended on the model of financial markets underlying the transmission mechanism. And this is precisely where the question of the model of financial markets becomes important. David Miles, speaking in 2009, had argued that '[t]here are economic models which would imply that the impact [of QE] is small – though I believe they rely upon unrealistic assumptions, specifically about the efficiency and completeness of financial markets' (Miles, 2009: 13).

Indeed, at the height of the crisis, disruptions in financial markets were glaringly evident for most economists. One MPC member (Interview 12), referring to 2008, claimed that it was obvious that markets were not perfectly arbitrated and that there would hence be supply effects.

The model which came to the fore was one going all the way back to Culbertson (1957) and Modigliani and Sutch (1966) and recently reworked by Vayanos and Vila (2009), the latter of which became the key model for central bankers. A DSGE model specialist argued: ‘So then for us central bankers, this work by Vayanos and Vila is important because for the first time it established a link between quantities and supply effects’ (Interview 24). In this model, agents are heterogeneous, assets are *not* perfectly substitutable, and therefore in which there are *limits-to-arbitrage* because of preferred habitats or market segmentation. If market participants prefer to inhabit local parts of the curve, and arbitrageurs fail to undo those preferred habitat effects (e.g. because they have capital/liquidity/collateral constraints) (Tenreyro, 2023: 9), then there are limits-to-arbitrage because assets are no longer perfectly substitutable, and, for instance, rates at the shorter end bear little relation to rates at the longer end (Interview 17; Interview 24). If these preferred habitats are not undone (see Bernanke, 2020), QE can exploit local supply effects by influencing the risk component (risk premia) of the yield curve, rather than expectations.

And yet this was a world away from the model of efficient markets which central bankers had worked with previously. For decades Bank economists had slowly and incrementally built tools and devices that shaped the doing of monetary policy, such as DSGE models, which implied the existence of complete and efficient financial markets. Replacing the model of complete and efficient market was no straightforward mission, even *materially*. The model of efficient markets was so dominant that model specialists across central banks (and at the Bank of England) had never needed to formalise consensus around how to model financial frictions, market inefficiencies, and incompleteness. And the edifice of the post-crisis New Keynesian DSGE model known as COMPASS (which replaced BEQM in 2011) at the Bank of England was still built on an inscription of ‘perfectly functioning financial markets’ which retained an inability to allow for an impact of QE. One model specialist at the Bank argued:

So we made the decision not to have explicit financial market frictions in the model, which was quite . . . many people thought that was quite odd, given that we were bringing the model in, sort of after the financial crisis, but the position that we had at the time was that there wasn’t a single dominant model of financial frictions, or the banking sector. And so it wasn’t obvious . . . we could have included something if we wanted to, but it wasn’t obvious that including something was the right thing to do. (Interview 17)

The DSGE model itself, therefore, retained the assumption of efficient markets, hardcoded into the material model. And forecasters at the Bank of England resorted to introducing some of these frictions in their macroeconomic modelling practices:

[We would] mimic the effects of financial frictions, that occur in richer models through the shocks that we have in the COMPASS model. And so the idea was we sort of think through ‘Okay, what does the financial shock look like from the perspective of a model like COMPASS, and then you can mimic the effects by putting through usually a combination of different shocks in the in the COMPASS model, so we kind of made that decision for that reason. Namely, that that wasn’t a sort of off-the-shelf [. . .] for financial frictions, there wasn’t really a kind of dominant way of doing it within the DSGE approach. (Interview 17)

Table 2. The state contingency of monetary policy.

	Yield curve reading and target	Socio-technical model	Intervention	Transmission
Benign state	Expectation hypothesis	DSGE model, incorporating efficient markets by omission	Communication and adjustment of rates	Expectation formation (and less constrained arbitrage)
Less benign state	Expectation hypothesis + risk premia	DSGE model, satellite models, working within a view of segmented/preferred habitat model Constrained arbitrage	Communication and Quantitative easing	Investors' portfolio rebalancing + liquidity provisions

Source: Author's own.

Bank staff would superimpose QE effects *as if* they were expectational effects. They:

arrived at their judgements about the effects it would have, and then they just imposed that on the model. And then as time went on their judgements were refined by looking at the things like event studies, which would tell you about the short-end yields. Not much more. So then assumptions would be made about how permanent effects on yields could be, could be mimicked in a model, and then some assumptions were made about what effects that would have on the rest. (Interview 17)

Qualitative and non-model-based judgements about the impacts of QE were superimposed on the model via expectational effects, supported via the satellite suite of models around the DSGE (Burgess et al., 2013; Interview 17).

But while *materially* the device failed to directly include frictions, superimposing these frictions in the policy arrangements implied that the *cognitive* model of complete and efficient markets had been replaced by incomplete and frictional markets. Indeed, this is the conclusion reached by some scholars and observers of the central banking world (Borio, 2011; Gabor, 2014: 158). The question this raises is: have central banks permanently replaced the underlying model of the transmission mechanism from one of complete and efficient markets to one of frictional and imperfect markets?

Performing ‘states of the world’ via quantitative easing

The above empirical sections have elucidated how the state of the world in which central banks operate changed following the crisis and how the Bank scrambled to realign its operative procedures with the shifting state of the world (i.e., financial market conditions). This section will first lay out the state contingency of monetary policy from the perspective of the Bank of England, detailing how those very same interventions may shape the state of the world, thus giving rise to a *performative* element to monetary policy as world-making ontological shaping.

The notion that monetary policy interventions are state-contingent is reflected in the documents published by the Bank of England – see Miles (2009), Haldane et al. (2016), Broadbent (2018), Vlieghe (2018), Bailey et al. (2020), Vlieghe (2020), Ramsden (2021), and Tenreyro (2023). Interventions and their transmission channels mirror the state of financial markets (see Table 2 for a schematic illustration). The governance of expectations is most relevant in the benign state of the world when financial markets are more efficient and complete. In the less benign state of the world (when arbitrage relations break down),

QE takes stronger hold via risk premia as markets are less arbitrated (Buiter and Sibert, 2007; Mehrling, 2011; Bailey et al., 2020).

The Bank's underlying model of financial markets is therefore not static, but changes to reflect their assessment of the 'states of the world'. And this will, in turn, shape the contingent interventions. For instance, from this viewpoint, QE will be deployed more actively when financial markets are disrupted, but not when markets are seen to be well-arbitrated and hence approximating the efficiency and completeness of markets. This is not to say that the channels are mutually exclusive but that QE's effectiveness and channels rely much on the states of financial markets. Bailey et al. (2020: 18) argue:

Both the portfolio balance and the market liquidity channels of QE transmission may depend intimately on the state of financial markets and so vary with the degree of market dysfunction. For example, as markets become more dysfunctional, arbitrageurs become more constrained, strengthening the role of portfolio rebalancing effects. Similarly, impaired market functioning may give rise to an increased role for a liquidity channel of QE, if liquidity premia are larger and more sensitive to intervention.

Similarly, Vlieghe (2018: 17) argues: 'large quantity effects are mostly temporary, and larger when intermediaries are constrained, as in the financial crisis. But they have much smaller or no long-run effects once intermediaries are unconstrained'. And on the same lines, Tenreyro (2023: 14) claims that 'QE effects on yields depend on the risk capacity of arbitrageurs. If that risk capacity is state-contingent, QE effects will be state-contingent and largest when market liquidity dries up'. Indeed, it is perfectly reasonable to think that in a liquidity crisis, arbitrageurs become increasingly constrained by limited capital and liquidity, resulting in those preferred habitat/market segmentation effects to remain in place for longer.

Another central banker interviewee argued:

I think that [the effects] change over time. And I think if you if you speak to policymakers, they would probably all agree with that. But then, you know, how strong they are in each of those states of the world, they would probably have differences of view on that. So they would all agree, I think, that there's some form of... state contingency of the affairs. (Interview 20)

But while monetary policy is contingent on the state of the world, there is also a *performative* element underneath the process of monetary policy. To the extent that QE is state-contingent, the very act of purchasing assets in financial markets will have an effect on financial markets conditions, that is, the states of the world themselves. In line with a MacKenzie and Callon's conception of performativity, I mean an act or intervention (QE) that brings about the (state of the) world of which it speaks. The model of efficient markets that was assumed away in the benign state of the world (i.e., as one which explains the conditions under which central bankers govern) is transformed into a performative object to be achieved in the less benign state of the world.

While in the previous section we have seen how the Bank sought to replace its own model to realign it with the state of financial markets, the model of complete and efficient markets remained operative in the background – the model drives contingent interventions to shape financial markets in its image. The intervention of QE is hence more than mere monetary policy – it is an intervention in the very texture of financial markets that alters them and their conditions. And the intervention is intended to ontologically shape markets into the image of the model of complete and efficient markets. This is particularly visible in those particular instances when central bankers judge the model to have detached from reality and when there is a threat that the model remains

detached from reality in a state of permanent performative failure. Instances of severely impaired market functioning are one key example of this. The latter implies arbitrageurs who, despite having trades that seek to bring about some degree of efficiency and completeness in the market, are constrained from doing so. Even worse, they are pushed to wind up the trades they already have on, as assets/securities that normally trade close to each other break away from each other.

The March 2020 ‘dash for cash’ is perhaps the most clear-cut example to illustrate the shift from one state of the world to the other, a shift brought about performatively by central banks. As the World Health Organisation declared Covid-19 a pandemic, markets responded quickly and intensely: a modern-day ‘bank run’ had set off. Global asset prices moved several basis points in the space of a few minutes and margin calls grew.³ The arbitrage relationships described earlier stood at the heart of the event. One of these in particular, known as the ‘cash-futures basis trade’, represented by far the most popular trade for relative value hedge funds in the Treasury market. This trade involves hedge funds taking large positions in government bonds funded on repo (and thus leveraged multiple times over) as the first leg of the trade, and shorting the corresponding futures contract as the second leg. In making the spread between the two, hedge funds normally enforce ‘the law of one price’ – that is to say, the exploiting of the opportunity moves the two rates (government bond and future rates) closer until there are no more opportunities left.

In the case of the March dash for cash, however, the same actors that are in other circumstances crucial in maintaining efficiency in financial markets, this time they turned into a source of systemic risk.⁴ Rising volatility as investors turned cash-hungry triggered margin calls as collateral prices fell, requiring them to either deposit additional funds or sell assets to maintain the required margin level. Non-banking institutions in particular, mostly leveraged hedge funds, saw the normally mean-reverting arb spreads moving against them, resulting in mark-to-market losses. For instance, the cash-futures basis that would, in efficient markets, trade at close to zero shot up to hundreds of basis points. This led to higher margin calls by dealers – that is, a demand by the broker for the trader to deposit additional securities as collateral. However, with not enough capital to meet those margins and to keep their positions on, traders were forced to unwind their positions by dealers. The result was further volatility and pressure on dealers’ balance sheets, leading to a lower intermediation capacity by dealers and hence a further contraction of liquidity. This was a looping effect, a ‘doom loop’ (Hauser, 2020), or margin spiral (Brunnermeier and Pederson, 2009) on those same spreads and the market-makers’ bid-ask spreads, in which a severe shortage of liquidity, price volatility, and rising margin calls fed on each other, thus threatening the very functioning of financial markets (Duffie, 2020; Schrimpf et al., 2020).

Severe market disruption, then, and perfect conditions for QE to make an impact. A decision to ‘go big and fast’ by major central banks, including the Bank of England, helped restore some stability to financial markets (FSB, 2020; Hauser, 2021; 2020). The central bank’s performative intervention worked insidiously through a (temporary) reformatting of the network that constituted the market (and the very liquidity that this made possible). The central bank enters the market, armed with a balance sheet, and makes bids for assets, mostly in the form of government bonds. Exchanging liquid cash for bonds, it first offers a reprieve for the conduits and intermediation between arbitrageurs. Dealers offload their inventories onto the central bank and hence also offload risk from their own balance sheet to the central bank’s, lowering risk premia. This frees up their capacity and appetite to intermediate between arbitrageurs. Lower margin calls remove some of the pressure off arbitrageurs to wind down their trades, thus allowing them to maintain them until they are successful.

The injection of liquidity was only possible by a *reconfiguration* of the network in which the central bank intervened, and the behaviour of the network was reformatted as market

participants regained confidence and risk appetite to make markets, at which point the central bank could slowly exit the market. An intervention such as QE targets confidence and risk, minimising those constraints on arbitrage, which is an attempt to bring about the performative object of complete and efficient markets.

From Callon's (2007) point of view, the formula/theory/model is not brought into being merely by the mere existence of the formula/theory/model itself. There are other forces and interests that cause the socio-technical network, or agencement, to be enacted. It is not enough, therefore, that both central bankers and market participants held a shared world of efficient markets in their activities. At certain points (such as the market impairment of March 2020), reality could stray away from the model, and it required efforts and an alignment of interests of the various actors to bring reality back to an approximation of efficient markets. Firstly, it required the effort of a central bank to exercise its balance sheet and to enter the network as it announced its intervention. And second, it required a condition in which it was in the interest of market participants – namely, arbitrageurs, dealers, and the central bank itself – to restore market efficiency. Indeed, the arbitrageurs and dealers willingly participated and cooperated with central banks in a reconstitution of the socio-technical agencement via which liquidity and confidence flowed once again. As risk abated, the yield curve could be once again interpreted primarily in terms of the expectation hypothesis (given the drop in risk premia), while the DSGE model became aligned more closely to the conditions of financial markets post-intervention.

QE, therefore, had a performative effect on the state of the world in which it was intervening. One member of the MPC explained that QE is primarily a monetary policy instrument, but given that the transmission mechanism relies on the state of the world, the way it operates has an element of world-making performativity to it, particularly in periods such as March 2020. One of my interviewees argued so:

It's a monetary policy instrument to hit the inflation target. So that's quite clear. I think it becomes more complicated because of this because the transmission mechanism [...] depends on what's going on in the economy and in markets at the time. And so, if you go back, if you go back to March [2020], you could reasonably argue that if the MPC hadn't done any QE at all, there would have been a very significant tightening in financial conditions in certain parts of the financial markets. [...] So it's monetary policy but the way that the actual transaction operates has a flavour of a cash flow alleviation type of instrument. And so there's a bit of ambiguity there about whether you should classify it as a market policy or as a monetary policy. (Interview 9)

The same process has been manifestly evident more recently in the Bank of England's conduct in 2022. As inflation in the UK skyrocketed, the MPC announced several rate hikes and later what it dubbed as 'quantitative tightening' (QT).⁵ The planned commencement of the programme for September 2022 came at an unfortunate moment in which the new British Chancellor, Kwasi Kwarteng, announced a 'mini-budget' involving planned freezes of household energy bills to be funded by the issuance of gilts and coupled with tax cuts for wealthier households. The announcement triggered a sell-off in the gilt market and a downward spiralling of the pound. The Bank, all set to start QT, reversed course on 28 September and announced a temporary programme of QE in an attempt to 'restore market functioning and reduce any risks from contagion to credit conditions for UK households and businesses' (BoE, 2022). In effect, the Bank was banking on the performative powers of QE to alter the state of the world in which it is effective.

But QE's performative powers are also evident when there is no severe shortage of liquidity in markets. For instance, the Bank provides empirical evidence that the portfolio

rebalancing channel was stronger in the first few years after QE was implemented (Joyce and Tong, 2012). As we have seen, prior to the crisis, central bankers were mostly concerned with governing expectations, and consequently, they interpreted interest rates as incorporating expectations about future rates. And while they were aware of the ‘risk premia’ that certain bonds entailed beyond the expectational component (especially at the longer end of the curve), they never turned them into an instrument of governance.⁶ Although not modelled in DSGE models, risk premia are primary targets of QE, which intervention is intended to minimise or eliminate risk premia (Lane, 2019). Central bankers seek to nudge market participants to rebalance their portfolios towards riskier assets, thus taking on more risk on their balance sheet. Rather than targeting expectations, the liquidity channel and portfolio rebalancing channel target risk.

Both these channels may have a performative effect on the state of the world, as they diminish risk premia through a shift in the arrangements and behaviour within the market. This means that QE would have had a differential impact according to the changing state of the world. In fact, Busetto et al. (2022) claim ‘Therefore there were good reasons to believe that the impact of the March 2020 programme would be closer to the one from QE1 [in 2009], whereas the impact of subsequent programmes was likely to be lower, as market dysfunction and risk aversion had abated’.⁷

I have argued that the model of complete and efficient markets was temporarily abandoned as an epistemic tool on which monetary policy hinged. In other words, the pre-crisis model with which they operated no longer approximated the world that central bankers were observing. And yet, despite the temporary abandonment of the model in the sense just outlined, it was still labouring in the background of monetary policy. Not only did it remain the underlying model of the workhorse DSGE model, but it was turned into a performative outcome to be *achieved*. A clue can be found when central bankers spoke of ‘returning to normality’ or a variation thereof (Bailey et al., 2020; Lagarde, 2022; Schnabel, 2023). While this can be easily interpreted as a return to pre-crisis economic growth, stable inflation rates and the like, it also refers to the pre-crisis monetary policy of expectation management. As QE brings about the ‘performative outcome’ of efficient markets, it ends up enacting a world in which it itself is *less* effective. In other words, as it moulds financial markets into efficient ones, QE is seen to become less effective due to the neutrality result in efficient financial markets (Wallace, 1981).⁸ And this is perfectly in line with the Bank’s intention to return to the pre-crisis policy. This is how MPC members anticipated the return to a QE-less world, a process that the Bank of England started in February 2022. The wind down itself, it was argued, would not have significant effects because the state of the world under which it will be wound down would be different from that in which QE was first implemented. In this ‘state of the world’, balance sheets could be wound down in a context of efficient markets, as had been predicted by David Miles as early as 2014:

These ideas suggest that the impact on monetary conditions of gilt sales by the Bank of England could be very different from the effect of its purchases. This is because sales would be conducted in a very different environment. [The] unwinding of such asset purchases is likely to occur when financial markets are operating more normally. (Miles, 2014: 16)

By the time that central banks start the process of returning to the management of expectations, which was always an end goal for central banks (Bailey et al., 2020: 27–28), the state of the world would have altered in such a way as to approximate complete and efficient markets. Central bankers, then, have always aimed to – and worked towards a – return to a form of monetary policy that relies on the governing mechanism of expectations management in a world of efficient markets.

Conclusion

This article foregrounds the performativity of QE, by which the very intervention is an attempt to alter the state of the world in which it functions. All along this process, the model of complete and efficient markets was embedded as a performative object to be achieved, a *particular* state of the world that was to be enacted by the very intervention of QE. This account of central bank performativity goes one step further than the current literature. The latter has largely restricted the notion of performativity in central banking as a process of communication by which private expectations are influenced, thus bringing inflation into line. But I argue that this process is contingent on very specific conditions, namely, a state of the world that characterised the Great Moderation in which markets approximated completeness and efficiency.

This article makes the case for there being more continuity between pre-crisis and post-crisis central banking than previously understood. While post-crisis central banking has been characterised as a shift away from performative central banking and a return to hydraulic governance (Braun, 2018), I claim then that central bank governance remains performative – indeed, in many ways, it is *even more* so. Interpreting QE as hydraulic governance implies that central bank action mechanically influences yields, while leaving financial markets otherwise unchanged – in other words, that central banks have adjusted to the new state of the world. But this misses the other side of the coin, namely, that QE is also a way by which a *particular mould of finance is made (or restored)*. I claim that the performative intervention of QE is a form of world-making performativity that shapes financial markets, allowing for different transmission mechanisms to take hold.

This mode of performativity is hence detached from the kind of performativity characterising pre-crisis central banking, as documented in the social science literature. The latter relies on a restrictive conception of performativity in terms of a shared epistemic world between central banks and actors in financial markets who act as centres of calculation, with central bank communication framing private market expectations. However, this paper has shown that once this shared world collapses, the central bank intervenes to re-make the world, and interventions such as QE involve a world-making function through which the central bank shapes ontologically the network. To appreciate its performative dimension, we need to pay more attention to the economic-financial models and theories underlying central bankers' practices, in the sense of MacKenzie (2006), which so far have been neglected even in recent studies that offer a sensitivity to these performative processes (Coombs, 2024).

It is to the ways in which these models are enrolled in the relationship between the states of the world and the very interventions and programmes devised by central banks (Walter and Wansleben, 2020) that this article contributes. This article hence proposes a broader view of performativity defined by the co-constitutive relationship between the epistemic models of central banks and the specific conditions of financial markets (i.e., states of the world) represented by the epistemic models. This conception of performativity goes back to a Callonian and MacKenzie version of the idiom of performativity.

But here is one way in which post-crisis world-making interventions differed from the past: while in the 1980s and 1990s, models (and the practices in which they were enrolled) were constructed to formalise and operationalise the economy, financial markets, and the impact of monetary policy, in turn, post-crisis central banking implied an inverse relationship. Models that represent the economy and financial markets – primarily DSGE models in which efficient markets are hardcoded by omission, and the yield curve which relies on the expectation hypothesis – operated as engines that drove contingent interventions in financial markets. QE was a world-making performative intervention to shape markets in the image of the model, such that central banks could restore the

expectation-management programme of monetary policy that was successful prior to the crisis.

The vast amount of asset purchases carried out over the past decade was therefore intended not simply as a monetary policy tool in the strictest sense – that is, to bring back inflation to target – but also to restore market completeness and efficiency. But, at least according to one view (dominant at the Bank of England), the closer the approximation of the state of the world to the model of complete and efficient markets, the weaker the effects of QE would be. In adopting this conceptualisation of financial markets and the transmission mechanism, we can conclude that QE will be used especially in times of crisis and less so in ‘normal’ times when the model of complete and efficient markets is taken to be an acceptable real-world approximation, with the management of expectations via Bank Rate prevailing once again.

It is relatively uncontroversial to claim, for instance, that the very intervention of QE during March 2020 calmed markets and provided a degree of stability. It is also plausible to think that central banks are aiming to enact the model of complete and efficient markets. Indeed, there is also a political rationale for the return to pre-crisis monetary policy, following Krippner (2011; 2007). QE is visibly political, causing overt distributional outcomes. In returning to the pre-crisis state of the world, the central bank’s actions become less visible because it can let markets do its own work, and its actions become less politicised in terms of distributional outcomes (Dietsch et al., 2018). It is entirely within the interest of central banks, then, to restore the benign state of the world of efficient markets, and as long as they can leverage their balance sheet, they will seek to (re-)enact it performatively.

This transition to the benign state of the world – to ‘normal times’ – was underway by 2022 in most developed countries, as central banks were compelled to raise rates and put a stop to asset purchases in the context of rising inflation rates. Central banks also considered – and later followed through – shrinking their balance sheet. But QT presents a question of its own in light of the performativity claims I have made. Is QT merely the inverse of QE, and hence can we make the same argument of linear performative change from one state of the world to another, but in the opposite direction? In other words, does QT performatively enact *inefficient* and *unstable* markets?

As we have seen, the argument by the Bank is simply ‘no’ – in a state of the world of complete and efficient markets, the Wallace neutrality result of balance sheet models remains in place irrespective of the direction in which the balance sheet is being levered. As Huw Pill (2022: 13) has recently argued, the impact of QT is also state-contingent, and ‘[w]hen markets are functioning well (and substitutability between money and bonds is high), QT has a smaller overall effect on asset prices than when markets are functioning poorly. This is in line with the strong impact of QE at times of market stress in 2009 and 2020 relative to its impact in other more benign market environments’. In other words, QT will have no effect on the states of the world because assets are substitutable and will leave no mark on financial markets themselves. This view is, however, countered by other experts. Daniela Gabor (2022), for instance, has argued that QT may be equally performative, in that it may hold some effects on the state of the world via collateral management. QT, which involves the offloading of sovereign bonds from the central bank’s balance sheet, could impact government bonds’ valuation. Because these serve as collateral in repo markets, QT could disrupt repo markets and in consequence lead to market liquidity spirals. But given that QT is still in its infancy, and little empirical literature has scrutinised its impact, any performative claim or lack thereof on QT remains necessarily conjectural. Indeed, the jury is still out.

Acknowledgements. The research work for this publication was partially funded by the ENDEAVOUR Scholarships Scheme and by the School of Social and Political Science at the University of Edinburgh. I would like

to express my sincere gratitude to Nathan Coombs for carefully reading an earlier version of the paper and providing insightful feedback. I am also thankful to the participants at the SASE conference 2024 in Limerick, particularly those in 'The Politics of Central Banking' panel, and the participants at the Finance and Society conference 2024 in Sheffield, for their thoughtful discussion. I extend my appreciation to the editors of *Finance and Society* for their support, as well as to the three reviewers for their meticulous reviews and constructive feedback.

Notes

1. Central banks had previously relied on open-market operations (OMOs) as the central instrument of monetary policy implementation, which involves the purchasing and selling of short-term securities to influence available reserves, influencing the cost of borrowing for banks in the interbank market. In doing so, a central bank could put pressure on short-term interest rates according to the level of the supply of and demand for reserves, supporting the provision of new lending in the economy. They later shifted to announcing the short-term rate. This is not to say that open-market operations became redundant. Indeed, central banks often do adjust and tweak reserves to support their 'open-mouth operations' (see, e.g. Hanes [2019] for the US case).
2. As Walter (2024) argues, the literature stops short of fleshing out in full the ways in which performativity is 'accomplished'. Even in the discussion of pre-crisis performative governance, the literature does not explain how the framing of expectations is successful, such that market participants transmit the policy to its intended outcomes – namely, firms' investment and households' consumption behaviour. Taking an Austinian (1962) conception of performativity, Walter (2024) claims that the illocutionary (i.e., the repeated speech acts and practices that set up the frame, as, for example, by the central bank) has taken precedence over the perlocutionary (namely, the conditions under which those utterances become successful, as, for example, the audiences' adoption of the frame the central bank is offering). The literature fails, therefore, to explain how that perlocutionary force is attained by way of framing. In other words, the focus is on monetary policy implementation, but not transmission. This article is also intended to rectify, in part, this omission in the literature by emphasising the transmission of monetary policy and the conditions in which this transmission is made successful (or not).
3. A margin refers to the difference between the value of the securities used as collateral and the loan amount provided by the broker against that collateral.
4. Among many lessons to be learned, Andrew Hauser (2020: 15) – Executive of the Markets division at the Bank of England – singled out the central role played by arbitrageurs and the breakdown in arbitrage, referring to 'the central role played by highly-leveraged but thinly capitalised non-banks in arbitraging between key financial markets' as the source of market disruption.
5. This is an inverse process to QE. In QT, the central bank *sells*, rather than buys, assets, in effect shrinking its balance sheet.
6. Risk premia are typically represented as deviations from the expectations hypothesis.
7. There exists a risk that significant purchases of assets induce a scarcity effect in specific markets, in which case financial markets become not more but less efficient. Central banks are aware of this risk and implement safeguards in both design and implementation of the programme of asset purchases (BIS, 2019). These measures are briefly outlined in Hauser (2021), and as yet, there has not been any indication that adverse effects on market functioning were manifested in the UK.
8. It should be noted, however, that such a view is contested even within the central banking community. For instance, Bernanke (2020) claims that this is not the case and that QE remains effective across various states of the world.

References

- Austin, J.L. (1962) *How to Do Things with Words*. Oxford: Clarendon Press.
- Bailey, A. (2020) The central bank balance sheet as a policy tool: Past, present and future. Speech given at the Jackson Hole Economic Policy Symposium, Federal Reserve Bank of Kansas City, 27–28 August.
- Bailey, A., Bridges, J., Harrison, R., Jones, J., and Mankodi, A. (2020) The central bank balance sheet as a policy tool: Past, present and future. Bank of England Staff Working Paper No. 899.
- Beckert, J. (2016) *Imagined Futures: Fictional Expectations and Capitalist Dynamics*. Cambridge, MA: Harvard University Press.
- Bernanke, B. (2020) The new tools of monetary policy. *American Economic Review*, 110(4): 943–83.
- Beunza, D. (2019) *Taking the Floor: Models, Morals, and Management in a Wall Street Trading Room*. Princeton, NJ: Princeton University Press.

- Beunza, D., Hardie, I., and MacKenzie, D. (2006) A price is a social thing: towards a material sociology of arbitrage. *Organization Studies*, 27(5): 721–45.
- Beunza, D., and Muniesa, F. (2005) Listening to the spread plot. In: Latour, B., and Weibel, P. (eds.) *Making Things Public: Atmospheres of Democracy*. Cambridge, MA: MIT Press, 628–33.
- Beunza, D. and Stark, D. (2004) Tools of the trade: The socio-technology of arbitrage in a Wall Street trading room. *Industrial and Corporate Change*, 13(2): 369–400.
- BoE (2022) Bank of England announces gilt market operation. *Bank of England News Release*, 28 September. <https://www.bankofengland.co.uk/news/2022/september/bank-of-england-announces-gilt-market-operation>. Accessed 18 December 2024.
- Borio, C. (2011) Central banking post-crisis: What compass for uncharted waters? BIS Working Paper No 353.
- Braun, B. (2014) Why models matter: The making and unmaking of governability in macroeconomic discourse. *Journal of Critical Globalisation Studies*, 1(7): 48–79.
- Braun, B. (2015) Governing the future: The European Central Bank's expectation management during the Great Moderation. *Economy and Society*, 44(3): 367–91.
- Braun, B. (2018) Central bank planning: Unconventional monetary policy and the price of bending the yield curve. In: Beckert, J. and Bronk, R. (eds.) *Uncertain Futures: Imaginaries, Narratives, and Calculation in the Economy*. Oxford: Oxford University Press, 194–218.
- Braun, B. (2020) Central banking and the infrastructural power of finance: The case of ECB support for repo and securitization markets. *Socio-Economic Review*, 18(2): 395–418.
- Broadbent, B. (2018) The history and future of QE. *Bank of England*. Speech given to the Society of Professional Economists, London. <https://www.bankofengland.co.uk/speech/2018/ben-broadbent-society-of-professional-economists>. Accessed 18 December 2018.
- Brunnermeier, M.K. and Pederson, L.H. (2009) Market liquidity and funding liquidity. *The Review of Financial Studies*, 22(6): 2201–38.
- Buiter, W.H. and Sibert, A. (2007) The central bank as the market maker of last resort: From lender of last resort to market maker of last resort. *VoxEU*, 13 August. <https://cepr.org/voxeu/columns/central-bank-market-maker-last-resort-lender-last-resort-market-maker-last-resort>. Accessed 18 December 2024.
- Burgess, S., Fernandez-Corugedo, E., Groth, C., Harrison, R., Monti, F., Theodoridis, K., and Waldron, M. (2013) The Bank of England's forecasting platform: COMPASS, MAPS, EASE and the suite of models. Bank of England Working Paper No. 471.
- Busetto, F., Chavaz, M., Froemel, M., Joyce, M., Kaminska, I., and Worlidge, J. (2022) QE at the Bank of England: A perspective on its functioning and effectiveness. *Bank of England Quarterly Bulletin*, 2022: Q1.
- Callon, M. (2007) What does it mean to say that economics is performative? In: MacKenzie, D., Muniesa, F., and Siu, L. (eds.) *Do Economists Make Markets? On the Performativity of Economics*. Princeton, NJ: Princeton University Press, 311–57.
- Cassar, D. (2022) *Inscribing Markets, Shaping Policy: A Sociological Investigation into the Yield Curve*. Unpublished dissertation, University of Edinburgh.
- Cassar, D. (2024) Economics as intervention: Expert struggles over quantitative easing at the Bank of England. *Socio-Economic Review*, 22(3): 1225–54.
- Choudhry, M. (2019) *Analysing and Interpreting the Yield Curve*. Chichester: John Wiley & Sons.
- Coombs, N. (Forthcoming) Infrastructural power in financial governance: Its meaning, applications, and varieties. In: Brandl, B., Campbell, M., and Westermeier, C. (eds.) *Cambridge Global Companion to Financial Infrastructures*. Cambridge: Cambridge University Press, in press.
- Coombs, N. (2022) Narrating imagined crises: How central bank storytelling exerts infrastructural power. *Economy and Society*, 51(4): 679–702.
- Coombs, N. (2024) Central bank power without central bank autonomy? *Finance and Society*, 10(1): 65–68.
- Coombs, N. and Thiemann, M. (2022) Recentering central banks: Theorising state-economy boundaries as central bank effects. *Economy and Society*, 51(4): 535–58.
- Culbertson, J. (1957) The term structure of interest rates. *Quarterly Journal of Economics*, 71(4): 485–517.
- Dale, S. (2010) QE one year on. Speech given at the CIMF and MMF Conference, Cambridge. <https://www.bankofengland.co.uk/-/media/boe/files/speech/2010/qe-one-year-on-speech-by-spencer-dale.pdf>. Accessed 18 December 2024.
- Dietsch, P., Claveau, F., and Fontan, C. (2018) *Do Central Banks Serve the People?* Cambridge: Polity.
- Duffie, D. (2020) Still the World's Safe Haven? Redesigning the US Treasury Market After the COVID-19 Crisis. Hutchins Center Working Paper 62. Hutchins Center on Fiscal and Monetary Policy.
- FSB (2020) Holistic review of the March market turmoil. Financial Stability Board Report to the G20. <https://www.fsb.org/2020/11/holistic-review-of-the-march-market-turmoil/>. Accessed 18 December 2024.
- Gabor, D. (2014) The ECB and the political economy of collateral. In: Goodhart, C.A.E., Gabor, D., Vestergaard, J., and Erturk, I. (eds.) *Central Banking at a Crossroads: Europe and Beyond*. London: Anthem Press, 157–76.

- Gabor, D. (2022) Zugzwang central banking (ECB edition). *Financial Times*, 8 September.
- Haldane, A., Roberts-Sklar, M., Wieladek, T., and Young, C. (2016) QE: The story so far. Bank of England Staff Working Paper No. 624.
- Hanes, C. (2019) Explaining the appearance of open-mouth operations in the 1990s U.S. *The North American Journal of Economics and Finance*, 48(C): 682–701.
- Hardie, I. (2004) 'The sociology of arbitrage': A comment on MacKenzie. *Economy and Society*, 33(2): 239–54.
- Hauser, A. (2020) Seven moments in spring: Covid-19, financial markets and the Bank of England's balance sheet operations. Speech given at Bloomberg, London. <https://www.bankofengland.co.uk/-/media/boe/files/speech/2020/seven-moments-in-spring-covid-19-speech-by-andrew-hauser.pdf>.
- Hauser, A. (2021) From lender of last resort to market maker of last resort via the dash for cash: Why central banks need new tools for dealing with market dysfunction. Speech given at Reuters, London. <https://www.bankofengland.co.uk/-/media/boe/files/speech/2021/january/why-central-banks-need-new-tools-for-dealing-with-market-dysfunction-speech-by-andrew-hauser.pdf>.
- Holmes, R.D. (2013) *Economy of Words: Communicative Imperatives in Central Banks*. Chicago, IL: University of Chicago Press.
- Jacobs, L.R. and King, D.S. (2016) *Fed Power: How Finance Wins*. Oxford: Oxford University Press.
- Jessop, B. (2010) Cultural political economy and critical policy studies. *Critical Policy Studies*, 3(3–4): 336–356.
- Joyce, M.A.S. and Tong, M. (2012) QE and the gilt market: A disaggregated analysis. *The Economic Journal (London)*, 122(564): F348–84.
- Krippner, G.R. (2007) The making of US monetary policy: Central bank transparency and the neoliberal dilemma. *Theory and Society*, 36(6): 477–513.
- Krippner, G.R. (2011) *Capitalizing on Crisis: The Political Origins of the Rise of Finance*. Cambridge, MA: Harvard University Press.
- Lagarde, C. (2022) Monetary policy normalisation in the euro area. *The ECB Blog*, 23 May. <https://www.ecb.europa.eu/press/blog/date/2022/html/ecb.blog220523~1f44a9e916.en.html>. Accessed 18 December 2024.
- Lane, P. (2019) The yield curve and monetary policy. Public lecture given at the Centre for Finance and the Department of Economics, University College London. <https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp191125~b0ecc8e6f0.en.html>. Accessed 18 December 2024.
- MacKenzie, D. (2006) *An Engine, Not a Camera: How Financial Models Shape Markets*. Cambridge, MA: MIT Press.
- MacKenzie, D. and Spears, T. (2014) 'The formula that killed Wall Street': The Gaussian copula and modelling practices in investment banking. *Social Studies of Science*, 44(3): 393–417.
- Mehrling, P. (2011) *The New Lombard Street: How the Fed Became the Dealer of Last Resort*. Princeton, NJ: Princeton University Press.
- Miles, D. (2009) Money, banks and quantitative easing. Speech given at the 14th Northern Ireland Economic Conference. <https://www.bankofengland.co.uk/news/2009/september/money-banks-and-quantitative-easing-speech-by-david-miles>. Accessed 18 December 2024.
- Miles, D. (2014) The transition to a new normal for monetary policy. Speech given at the Mile End Group, Queen Mary University of London. <https://www.bankofengland.co.uk/-/media/boe/files/speech/2014/the-transition-to-a-new-normal-for-monetary-policy.pdf>. Accessed 18 December 2024.
- Modigliani, F. and Sutch, R. (1966) Innovations in interest rate policy. *The American Economic Review*, 56(1/2): 178–97.
- Nyborg, K.G. and Östberg, P. (2014) Money and liquidity in financial markets. *Journal of Financial Economics*, 112(1): 30–52.
- Pill, H. (2022) What did the monetarists ever do for us? Speech given at the Walter Eucken Institut/Stifung Geld und Währung Conference. <https://www.bankofengland.co.uk/speech/2022/june/huw-pill-speech-at-the-walter-eucken-institute>. Accessed 18 December 2024.
- Ramsden, D. (2021) QE as an economic policy tool: What does it do and how should we use it? Lecture given at the Bank of England. <https://www.bankofengland.co.uk/speech/2021/february/dave-ramsdend-peter-sinclair-town-hall-lecture-series>. Accessed 18 December 2024.
- Schnabel, I. (2023) *Back to Normal? Balance Sheet Size and Interest Rate Control*. European Central Bank. https://www.ecb.europa.eu/press/key/date/2023/html/ecb.sp230327_1~fe4adb3e9b.en.html. Accessed 18 December 2024.
- Schrimpf, A., Shin, H.S., and Sushko, V. (2020) Leverage and margin spirals in fixed income markets during the Covid-19 crisis. BIS Bulletin 2. Bank for International Settlements.
- Shleifer, A. and Vishny, R.W. (1997) The limits of arbitrage. *The Journal of Finance*, 52(1): 35–55.
- Tenreiro, S. (2023) Quantitative easing and quantitative tightening. Speech given at the SES Annual Conference, Glasgow. <https://www.bankofengland.co.uk/speech/2023/april/quantitative-easing-quantitative-tightening-speech-silvana-tenreiro>. Accessed 18 December 2024.

- Vayanos, D. and Vila, J.-L. (2009) A preferred-habitat model of the term structure of interest rates. NBER Working Paper 15487. https://www.nber.org/system/files/working_papers/w15487/w15487.pdf. Accessed 18 December 2024.
- Vlieghe, G. (2018) The yield curve and QE. Speech given at Imperial College Business School, London. <https://www.bankofengland.co.uk/speech/2018/gertjan-vlieghe-imperial-college-business-school-london>. Accessed 18 December 2024.
- Vlieghe, G. (2020) Monetary policy and the Bank of England's balance sheet. Online webinar. <https://www.bankofengland.co.uk/speech/2020/gertjan-vlieghe-speech-monetary-policy-and-the-boes-balance-sheet>. Accessed 18 December 2024.
- Wallace, N. (1981) A Modigliani-Miller theorem for open-market operations. *The American Economic Review*, 71(3): 267–74.
- Walter, T. (2024) One future to bind them all? Modern central banking and the limits of performative governability. *British Journal of Politics and International Relations*, 26(3): 718–41.
- Walter, T. and Wansleben, L. (2020) How central bankers learned to love financialization: The Fed, the Bank, and the enlisting of unfettered markets in the conduct of monetary policy. *Socio-Economic Review*, 18(3): 625–53.
- Wansleben, L. (2018) How expectations became governable: Institutional change and the performative power of central banks. *Theory and Society*, 47(6): 773–803.
- Wansleben, L. (2023) *The Rise of Central Banks: State Power in Financial Capitalism*. Cambridge, MA: Harvard University Press.
- Woodford, M. (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.
- Zaloom, C. (2009) How to read the future: The yield curve, affect, and financial prediction. *Public Culture*, 21(2): 245–68.