Bank of England

Are the effects of quantitative easing and tightening state contingent?

Staff Working Paper No. 1,185

May 2026

Michael Ellington, Costas Milas and Ryland Thomas

Bank of England

Staff Working Paper No. 1,185

Are the effects of quantitative easing and tightening state contingent?

Michael Ellington,[1] Costas Milas[2] and Ryland Thomas[3]

Abstract

We provide evidence that quantitative easing (QE) and quantitative tightening (QT) policies are state contingent. Using 60 years of UK data on public-sector debt sales to the banking system we identify a novel bank funding shock that indicates the impact of unconventional monetary policies changes significantly over time, and that regimes are non-recurrent. Our approach also permits an appraisal of state contingency at different stages of transmission. Over successive QE rounds, we find the responsiveness of government bond yields to a given amount of QE falls. However, demand becomes more responsive to yield changes, while inflation exhibits more persistence and greater sensitivity to the output gap. In the presence of such state contingencies, our results suggest careful monitoring is needed when assessing the impact of QE policies.

Key words: Quantitative easing, quantitative tightening, unconventional monetary policy, bank funding, vector autoregression.

JEL classification: C11, C32, E52, E58.

---

(1) University of Liverpool Management School. Email: m.ellington@liverpool.ac.uk
(2) University of Liverpool Management School. Email: costas.milas@liverpool.ac.uk
(3) Bank of England. Email: ryland.thomas@bankofengland.co.uk

The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. We appreciate the valuable comments of Christiane Baumeister, Jonas Arias, Iryna Kaminska, Simon Lloyd, Chris Martin, David Ronicle, Martin Seneca and Tim Congdon. We are also grateful to seminar participants at the Bank of England, the University of Liverpool, and the Royal Economic Society 2025 Annual Conference at the University of Birmingham. We acknowledge funding from the University of Liverpool Impact Development Fund.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Email: enquiries@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

This paper presents new empirical evidence on the state-contingent impact of Quantitative Easing (QE) and Quantitative Tightening (QT) policies in the UK. We adopt a novel approach by including earlier historical debt management operations that have comparable balance sheet impacts to unconventional monetary policies (UMPs). We provide a thorough assessment of how each UMP and UMP–like episode over the past 60 years affects the UK economy using data on public sector debt sales to the banking system. To the best of our knowledge, we are the first to quantify the state contingencies of each QE episode in the UK along with an initial assessment on the implications of QT.

Our definition of state contingency within this study is agnostic. We aim to capture state contingencies using time-varying parameter structural vector autoregressive (VAR) models (Petrova, 2019). This is inherently different from the structural VAR literature identifying state contingencies in macroeconomic data using either monetary policy regimes (i.e. high/low interest rates as in Hayashi and Koeda (2019), among others), or episodes of high and low macroeconomic uncertainty (e.g. Alessandri and Mumtaz, 2019), or times of high/low financial stress (see e.g. Hubrich and Tetlow, 2015). These models typically impose an *a priori* a restriction that there are a small number of recurrent regimes. However, our approach offers substantial flexibility by allowing for the possibility of many different, non-recurring regimes, which in turn allows us to investigate whether the recurrent regime assumption is valid.

The paper makes three key findings. First, we show as a general feature, that the impact of UMP policies in the UK are state contingent, based on the past 60 years of data. However, our analysis indicates the drivers of such state contingencies are dependent on a mix of factors, reflecting the policy framework, the structure and state of financial markets and conditions in the real economy. We provide evidence that regimes are non-recurrent and do not appear to be related to a particular state variable. Second, we distinguish between the unanticipated and systematic effects of UMP policies on the real economy. Our results indicate the unanticipated component contributes relatively small amounts to the total impact, suggesting that QE was largely anticipated given the shocks hitting the economy. But the total effect of QE provides substantial support to output during each episode and increases inflation. Third, we disentangle the transmission mechanism at work during each UMP episode. Over successive QE and QT episodes since 2009, the impact on gilt yields for a given quantum of QE falls. However, such a fall is offset by increases in both the responsiveness of the output gap to gilt yield changes, and the responsiveness of inflation to the output gap.

The general consensus on QE policies is that they are effective tools for stimulating and stabilizing economies (for an in-depth review of the literature on UMPs, see e.g. Bhattarai and Neely, 2022). A number of studies confirm this for various economies across the globe and quantifies the effects of early large-scale asset purchases using small samples (see e.g. Kapetanios et al., 2012; Baumeister and Benati, 2013; Weale and Wieladek, 2016). Another subsequent branch of the literature on UMPs suggests that the effects and channels through which such policies operate are likely to be state contingent. However, the drivers of such state contingencies are uncertain and difficult to pin down. (see e.g. Joyce and Tong, 2012; Haldane et al., 2016; Du et al., 2024; Braun et al., 2025).

Existing work on the transmission mechanism of UMPs provides robust evidence of a portfolio rebalancing channel (investors rebalancing their portfolios into other assets following gilt purchases, see Joyce et al. (2012); Joyce and Tong (2012)), and a signaling channel (lower expectations of future short-term rates) both of which have an effect on gilt yields and other asset prices. QE can affect the economy more generally by reducing uncertainty (Kapetanios et al., 2012; Joyce et al., 2012). Studies suggest the strength of the portfolio rebalancing and signaling channels depend on a range of factors such as: the state of financial markets; the capital position of market makers and arbitrageurs in government bond markets (e.g. Vayanos and Vila, 2021)[1]; market participants' anticipation of the amounts announced (e.g. Bernanke, 2020; Braun et al., 2025)[2]; and the level of interest rates within the economy.[3]

[1]More specifically, if market stress is high and arbitrageurs capital constraints becoming binding, the effects on term-premia are likely to be larger.

[2]Note this affects the relationship between the announced amount of QE and financial market yields. The underlying impact on yields from the unanticipated component of amounts announced might be unchanged.

[3]Note also the aggregate demand effects of UMPs will largely depend on how gilt yields affect the cost of capital to firms and household wealth, which will depend on spillovers to other asset prices and yields such as corporate bond yields, equities, the exchange rate and house prices. These spillover effects are also likely to be a function of the state of financial markets and the perceived substitutability between gilts and these assets.

An issue that hinders analysis of state contingencies is the lack of historical precedents for such policies. To resolve this issue, we look at Bank of England debt management operations in periods prior to 2009. We focus on operations that had similar effects on financial balance sheets as QE and QT. These "underfunding" and "overfunding" operations were often undertaken for different policy purposes and in evolving financial market structures, which should shed light on the various types of state contingency identified in the literature. We use these operations to identify a novel "bank funding" shock consistent with the economic rationale that underpins the channels through which UMPs transmit to the real economy. Our bank funding shock captures the impact of QE/QT from 2009 onward, as well as the UMP-like debt management operations in earlier periods.

To identify the shock, we use 60 years of historical data on public sector debt sales to the banking system, expressed as a contribution to broad money growth, as a proxy variable that captures the financial balance sheet effects of UMPs. The benefit of this proxy is threefold. First, it is easy to construct for long time-periods; thereby allowing us to track state-contingent effects on financial markets as they evolve over time.[4] Second, our measure contains significant variation prior to 2009; an inherent shortcoming of alternative central bank balance sheet measures (see e.g. Weale and Wieladek, 2016). Third, it directly relates to the broad money supply impact of QE, which, as King (2024) and Congdon (2025) argue, is the key feature of QE/QT operations and an integral part of their transmission to the real economy. However, we view the impact via broad money as a representation of the portfolio rebalancing channel. This emphasises the imperfect substitutability between money and bonds on the asset side of non-bank investor balance sheets rather than then "local supply" or "duration" channels, which focus more on the composition of the assets purchased by the central bank (see Busetto et al. (2022) for a discussion).

Our paper relates primarily to the literature using VAR techniques to analyze the macroeconomic impact of UMPs. Kapetanios et al. (2012); Bridges and Thomas (2012); Haldane et al. (2016); Weale and Wieladek (2016) all use various structural identification schemes to show that QE policies stimulate the UK economy and produce positive inflationary impacts. Kaminska and Mumtaz (2022) study the various rounds of QE prior to

[4]The period since the 1960s have seen major changes in the micro structure of the UK gilt market such as the "Big Bang" reforms of the 1980s and the development of a gilt-repo market in the mid-1990s both of which increased the liquidity of the gilt market and its ability to absorb large shocks.

2020. Their findings suggest that the economic impact of QE policies varies in strength with the mix of different channels and motivates further investigation of state contingencies.[5] Baumeister and Benati (2013) consider pure spread shocks for the US and UK to analyze the first round of asset purchases and conclude that such shocks capture a range of influences. Ellington (2022) considers dynamics of UMPs for the US economy to the point where the Federal Reserve moves away from the effective lower bound.

In addition to shedding light on the state-contingent effects of UMPs, our results also disentangle the total and unanticipated impact of UMPs in a clean and consistent manner. This is because our models implicitly contain a time-varying UMP/QE "reaction function" to standard macroeconomic and financial (gilt) market shocks. Isolating the unanticipated component allows for a comparison with the results from event studies of different QE/QT episodes (see e.g. Busetto et al., 2022; Braun et al., 2025). It also allows us to gauge the broader effect of UMPs through their systematic response to shocks as well. We also diagnose state contingencies at different stages of the transmission mechanism using the reduced-form semi-elasticities between key variables from our structural VAR responses. We back out model-consistent portfolio rebalancing and signaling channels that highlight a weakening relationship between financial market yields and the quantities of asset purchases across QE episodes, which is largely attributable to a declining signaling channel. However, we find that aggregate demand in the later QE periods and the (most recent) QT period is more responsive to gilt yields. We also provide evidence that the inflation/output gap trade-off worsens over QE episodes. Our results suggest this reflects variability in underlying nominal and real rigidities as a result of global supply chain problems and other supply-side issues that emerged as the UK and other major economies came out of the COVID-19 pandemic. It also reflects increased inflation persistence.

The remainder of our paper proceeds as follows. Section 2 presents our proxy measure of public sector debt sales to the banking system and discusses the economic data we use. Section 3 outlines the methodology of our baseline analysis. Sections 4 and 5 report our results and robustness analysis, respectively. Section 6 summarises our estimated impacts across QE/QT episodes and Section 7 concludes.

[5] Specifically the signaling channel can have a persistent impact on interest rates at the short end of the yield curve and so affect the pricing of a broad range of loans and financial assets. In contrast, local supply and duration effects have their main impact at the longer end of the curve.

2 Tracking Quantitative Easing/Tightening through public sector debt sales to the banking system and state contingency

Public sector debt sales to the banking system are a feature of historical debt management operations in the UK. Until the mid-1980s, debt management was viewed by the Bank of England as part of its monetary policy toolkit alongside short-term interest rates. Throughout the 1960s and 1970s, there was frequent recourse to fund government deficits by sales of debt to the commercial banking system, especially as large deficits became difficult to finance from traditional investors in the non-bank financial sector (which we associate today with the "preferred habitat" investors in portfolio balance models such as Vayanos and Vila (2021)). This policy, known as "underfunding", implies very similar portfolio effects on non-bank financial balance sheets and the money supply to those of central bank purchases of government debt. Figures 1 and 2 show how the purchase (sale) of government bonds by commercial banks has an almost identical effect on non-bank investor balance sheets as a QE (QT) operation does (i.e. where a central bank buys (sells) such securities). In addition, because commercial bank reserves were unremunerated in this period, Bank of England purchases of government debt also typically led to increased debt holdings by the commercial banking system. To maintain control of short-term interest rates, the Bank would have to offset any purchases it made by sales of Treasury bills to the commercial banks. The Bank would in effect perform an "operation twist" and the commercial banks would be net purchasers of government debt, as shown in Figure 1c.

Subsequently, the early 1980s saw the reversal of such policies as the authorities switched to "overfunding" the deficit, by selling more government bonds to the non-bank private sector ("preferred habitat" investors) than necessary to fund the deficit. Ronicle (2026) provides an in-depth analysis of this policy. Figure 2 shows that this has a similar effect on portfolios to QT and directly reduces the growth of broad money, which at the time was the key intermediate target of monetary policy.

From Figures 1 and Figure 2, it is clear that the initial impact of underfunding and overfunding operations on non-bank investors' portfolios are the same as QE and QT respectively. That suggests they should work through similar transmission mechanisms. In the case of underfunding and QE operations, they should trigger the same portfolio balance effects as non-bank investors attempt to reinvest their deposits in other risky securities, pushing down yields and increasing asset prices in the process. The only difference is that commercial bank ends up holding gilts rather than interest-bearing reserves on its balance sheet (i.e. slightly different forms of interest-bearing public sector debt).[6]

Figure 1c, shows what happens when the Bank of England made purchases in periods with unremunerated reserves. As mentioned, the Bank would have to drain the reserves created by its purchases, typically with offsetting sales of Treasury bills to the commercial banks. The only difference in this case to 1a and 1b is that commercial banks would have Treasury bills to match the deposits created.

From Figure 2, we see the same analogy applies in reverse for gilt sales. The non-bank private sector holds fewer deposits and more gilts. In the case of an overfunding operation, the over-issuance of gilts initially leads to a rise in government deposits at the central bank at the expense of reserves. The complication is that the overfunding operations of the 1980s also occurred during a period when reserve balances were unremunerated, and banks held relatively small amounts. Thus, a significant overfunding operation pushes reserve balances close to zero, putting upward pressure on money market rates. To maintain control of short-term interest rates the BoE purchases short-term bills from banks to replenish their reserves; such operations are in red. Similar to QT, the commercial banks hold fewer interest-bearing liquid assets as a counterpart to lower deposit liabilities.[7] As the intention at the time was to help meet official monetary targets, overfunding may also have had "signaling" effects on inflation and interest rate expectations over and above the effects through portfolio balance channels.

The late 1980s to early 2000s saw further changes in debt management and the move to full funding of deficits over either the financial year or over the course of a cycle. So underfunding and overfunding episodes were rarer and there was a gradual separation

[6]This may imply some differences in the commercial bank response to QE if the perception of liquidity differences between the two operations affect bank lending. However, the UK evidence for QE suggests little impact on bank lending, see Busetto et al. (2022). We review the impact on bank lending in Section 5

[7]Initially, the bills the Bank bought were Treasury bills. When such bills matured the government redeemed them using its excess deposits at the central bank (this transaction is not shown in the diagram), so the Bank of England balance sheet would return to its initial state as in Figure 1b. However, the stock of Treasury bills held by the banks was limited and the Bank of England then turned towards purchasing commercial bills. So a side-effect of overfunding was that the BoE ended up with a "bill mountain" of short-term private sector paper which it had to roll over in the market to maintain banks' reserves.

Figure 1. Quantitative easing and bank funding balance sheet mechanisms
Notes: The top panel of this figure, Figure 1a, shows a QE operation by the central bank where it purchases government bonds from non-bank investors. The bottom panel of this figure, Figure 1b, shows what happens when commercial banks do the same. Figure 1c shows what happens when the central bank makes purchases from non-bank investors and reserves are unremunerated forcing it to sell Treasury bills to the commercial banks.

of the UK's monetary policy and debt management objectives; we provide full details in Online Appendix A, Table A.1.

The evolution of debt management operations occurred alongside other structural changes in the policy paradigm, financial markets, and economic conditions, each of which are likely to have influenced the transmission of different funding policies to the broader economy. This motivates our agnostic prior on state contingency. We distinguish between several regimes spanning the evolution of debt management policies and UK financial markets.

The first episode covers 1964Q2–1971Q3, which we label "leaning against the wind". The

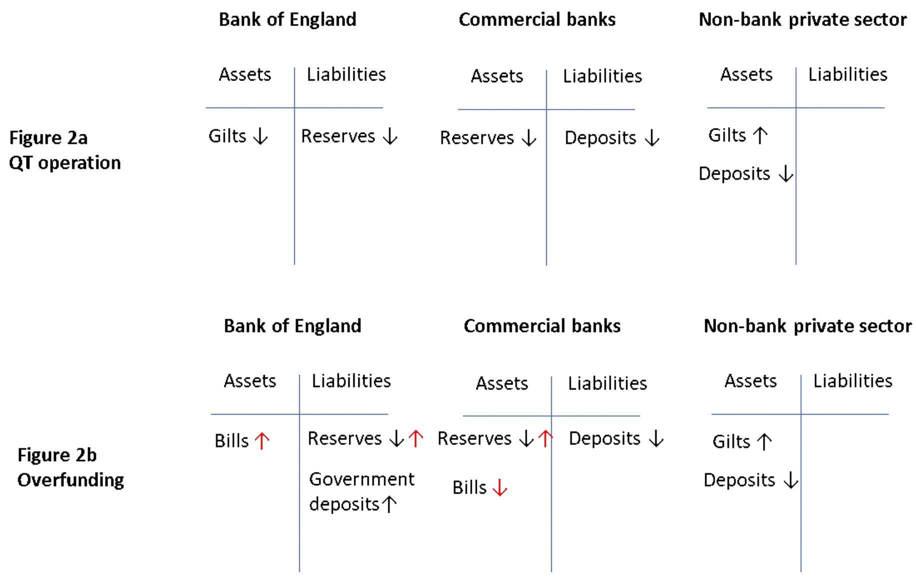

Figure 2. Quantitative tightening and bank funding balance sheet mechanisms
Notes: The top panel of this figure, Figure 2a, shows a QT operation by the central bank where it sells government bonds to non-bank investors. The bottom panel of this figure, Figure 2b, shows what happens when commercial banks do the same. The arrows in red relate to Bank actions to a significant overfunding operation. For example the overfunding operation of the 1980s saw no renumeration of reserve balances and banks held relatively small amounts. Therefore a significant overfunding operation pushes reserve balances close to zero, putting upward pressure on money market rates. Operations in red represent the Bank of England interventions to replenish reserves.

general BoE policy over this period was to smooth long-term yield fluctuations through purchase (sale) of government debt when gilt yields rise (fall).[8] Our next episode spans 1971Q3–1981Q4 that we label "Underfunding". Throughout this period, the BoE allowed for more market determination of yields. When the BoE deemed frictions impeded sufficient secondary market gilt sales to preferred habitat investors, the financing of government deficits was, by residual, from the banking system. As we show below, these were on a scale equivalent to post-2009 QE operations .[9]

The next two episodes span 1982Q1–1986Q4 and 1987Q1–1992Q4 and are labelled "Overfunding" and the "Big Bang", respectively. The former relates to the period when debt management operations were geared towards a direct reduction in money supply growth

[8]As in Figure 1c, the Bank would typically drain the reserves created by these operations by sales of bills and short-maturity debt to the banking system.

[9]During the 1960s and 1970s, smoothing yield fluctuations was undertaken to protect the undercapitalized stock market "jobbers" who acted as market makers and were largely small partnerships vulnerable to large movements in yields (see Allen (2019)). The subsequent market determination of yields which put pressure on the jobbers as market makers and made the gilt market more sensitive to economic and financial conditions. In practice, this period witnesses gilt sales to non-bank investors increasingly in opportunistic bursts.

to meet official monetary targets (this represents the only major period where funding policy has analogous impacts on financial balance sheets at QT). The latter relates to the period after major stock exchange reforms that increased the capitalization of market makers in the gilt market. Our final two episodes span 1993Q1–1997Q2 and 1997Q3–2008Q4. The former relates to the period that saw increased gilt market liquidity with the development of the gilt repo market in 1996. The latter relates to the period when the operational independence of monetary and debt management policies was established and respectively given to the Bank of England and the Debt Management Office to undertake. Table A.1 in Online Appendix A provides further details of our funding episodes, what makes them distinct from one another, and how they relate to financial balance sheet effects of QE/QT policies. In each regime, the purpose and role of debt management operations and their relationship with monetary policy differs and there are also structural differences in the depth and liquidity of financial markets.

Given the above, we use a measure of UK public sector debt sales to the banking system to track the impact of unconventional monetary policies over time. Our measure, psc_t, is the contribution of public sector borrowing from the banking system to the annual growth of the UK broad money supply, M4/M4ex. These data, part of the credit counterparts to broad money growth (see Bridges and Thomas (2012), are available from the Bank of England website as well as the Economic Statistics Centre of Excellence ESCoE historical data repository, and go back to 1963 at a quarterly frequency. The "public sector contribution" to broad money growth captures both over- and underfunding and reflects the holdings of government debt of both the central bank and commercial banks. Therefore, our measure encapsulates the operations we outline in Figures 1 and 2.[10]

Figure 3 plots psc_t against a backdrop of the historical funding regimes we outlined above, and the QE/QT operations from 2009. As we can see from the time profile of psc_t, the measure increases strongly during the various QE periods in 2009, 2012, 2016 and 2020. From a historical perspective, this measure also shows the positive and negative contributions to money growth as a result of the under- and overfunding episodes in the

[10] The link between funding from the banking system and broad money aggregates was key to the credit counterparts analysis of the money supply starting in the late 1960s as central banks began to pay more attention to the money supply. M4/M4ex is a break-adjusted series that splices conventional M4 with M4 excluding other financial corporations (OFCs); this is the Bank of England's preferred measure of UK broad money.

1960s–1970s and 1980s, respectively. The highest values coincide with the largest rounds of the Bank of England's asset purchase facility, namely, QE1, QE2/3, and QE5. But earlier operations in the 1960s and 1970s are on a similar scale. Then, towards the end of sample, psc_t becomes negative, in conjunction with the initial phase of QT.

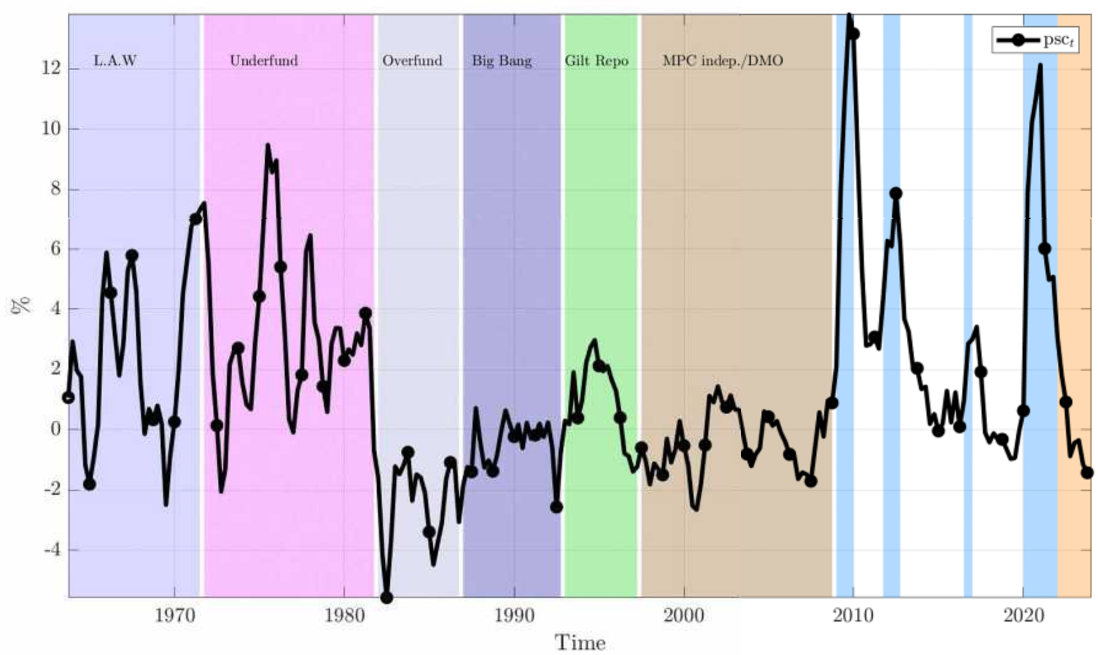

Figure 3. Public sector borrowing from the banking system as a proportion of UK broad money supply from 1963–2024

Notes: The figure plots public sector borrowing from the banking system as a proportion of UK broad money supply from 1963Q4–2024Q1, psc_t. Shaded bars show historical funding regimes where L.A.W is the "Lean Against the Wind" episode from 1963Q4–1971Q3. "Underfund", reflects the underfunding of deficits during the period of Competition and Credit control, the mid-1970s gilt strikes and the beginning of monetary targeting from 1971Q3–1981Q4. Overfund is the overfunding episode from 1982Q1–1986Q4, Big Bang reflects the introduction of the "Big Bang" reforms and the return to "full funding" over the period 1987Q1–1992Q4. Gilt-repo covers the more relaxed "full financing" rule and the development of the gilt repo market from 1993Q1–1997Q2. MPC ind is when the Bank of England obtained operational independence and the Debt Management Office takes over debt management from 1997Q3–2008Q4. Unconventional monetary policy episodes (i.e. QE/QT periods) are the bars from 2009 to 2024. QE1 is from 2009Q1–2010Q1; QE2/3 is from 2011Q4–2012Q4; QE4 is from 2016Q3–2017Q1; QE5 is from 2020Q1–2021Q4; QT1 is from 2022Q1–2024Q1.

Our measure captures both past under- and overfunding operations, and recent QE/QT operations in terms of the flow of purchases or sales scaled by their contribution to broad money growth. This is an appropriate quantum to compare the size of operations over time, given it is in line with a portfolio re-balancing approach where shares of portfolios relate to relative rates of return. When money is exchanged for gilts on the non-bank private sector's balance sheet, this implies rates of return will be a function of the percentage change in the money stock, as in Pill (2022). It is also consistent with a monetarist view of QE (QT), where the effects depend directly on the quantitative impact on the broad money supply, as in Bridges and Thomas (2012) and King (2024).

We note here that nominal GDP is often used as a scaling variable for QE in other studies (see e.g. Weale and Wieladek, 2016), and that: i) a measure of the size of government bond market could be another candidate for the denominator of psc_t from a portfolio balance perspective; and ii) we might focus only on central bank holdings of government debt going back in time. We formally address such issues by constructing alternative proxies of psc_t in Section 5. Notice that Figure A.2 in Appendix A shows that scaling bank purchase of gilts by the UK broad money supply is quantitatively similar to scaling by nominal GDP for the period since 2009 and hence our impulse response analysis is comparable to prior studies of recent QE episodes (e.g. Weale and Wieladek, 2016).

In addition to psc_t, we use UK macroeconomic data relevant for assessing the impact of QE (QT) over time, from 1963Q3–2024Q1. The end of the sample covers the passive phase of QT from 2022Q1 and the active sales phase beginning in 2022Q4 up to the first quarter of 2024. To proxy economic activity, we use the output gap data series provided by the Office for Budget Responsibility (OBR), y_t. The output gap is available from 1972 onwards. We extend the data back to 1963 by constructing an average output gap measure using: (i) the Hodrick and Prescott (1997) filter; (ii) the Christiano and Fitzgerald (2003) band pass filter; and (iii) a quadratic trend filter. We also obtain, from the Bank of England's Millennium of Macroeconomic Database, the annual rate of consumer price inflation, π_t; Bank Rate, r_t; and the 10-year gilt yield, lty_t. We plot the economic data in Online Appendix A in Figure A.3.

3 A flexible approach to modeling state contingencies

Our approach to tracking state contingency in the impact of UMPs allows for the possibility of non-recurring regimes whilst retaining the possibility of recurring regimes if the data dictates. We specify a time-varying coefficient VAR model (TVP-VAR) with p lags and N variables as

$$\mathbf{Y}_t = \beta_{0,t} + \sum_{p=1}^{k} \beta_{p,t}\mathbf{Y}_{t-p} + \mathbf{e}_t, \quad (1)$$

$$\mathbf{e}_t = \mathbf{\Omega}_t^{-\frac{1}{2}}\boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \backsim \mathrm{N}(0, \mathbf{I}_N) \quad (2)$$

where $\mathbf{Y}_t \equiv [\mathrm{r}_t, \mathrm{psc}_t, \mathrm{lty}_t, \pi_t, y_t,]'$. $\mathbf{Y}_{t-p}$ contains lags of $\mathbf{Y}_t$. We estimate the model based on Quasi-Bayesian Local Likelihood methods following Petrova (2019), and provide further estimation details and relative benefits of such an approach in Online Appendix B.

3.1 Structural Identification

We conduct structural inference using time-varying impulse response functions that at each point in time assume that parameters remain constant over the impulse horizon (60 quarters). This is because time variation in the autoregressive parameters and covariance matrices evolves in a non-parametric manner and hence possesses no laws of motion. We use the algorithm of Arias et al. (2018) to identify structural shocks using contemporaneous sign restrictions and zero restrictions.[11] Specifically, we identify: a monetary policy shock, $\mathbf{v}_t^{\mathbf{mp}}$; a bank funding shock, $\mathbf{v}_t^{\mathbf{q}}$; a supply shock, $\mathbf{v}_t^{\mathbf{s}}$; a (non-policy) demand shock, $\mathbf{v}_t^{\mathbf{d}}$; and a long-term yield shock $\mathbf{v}_t^{\mathbf{lty}}$. Table 1 reports the contemporaneous response of our variables to each identified shock.

Our restrictions for $\mathbf{v}_t^{\mathbf{q}}$ are consistent with the channels through which unconventional monetary policies are thought to affect an economy (see e.g. Bridges and Thomas, 2012; Kapetanios et al., 2012; Weale and Wieladek, 2016). It is distinct from a conventional monetary policy shock through the zero impact restriction on the short-term interest rate similar to the pure spread shock in Baumeister and Benati (2013). The positive impact response to the output gap captures the expansionary effects UMPs have on the real economy, and the response of the long-term yield proxies the financial market impact of bank funding shocks, although we check the impact on other financial variables in Section 5. We also im-

[11] There is an ongoing debate around this identification procedure, with Baumeister and Hamilton (2015) critique the prior does not vanish asymptotically and others, including for example Arias et al. (2025), defending such an approach. They show, for a various VAR models identified using sign restrictions, that the point-wise posterior median and credible sets are qualitatively similar to those using our identification scheme. Technically taking these points from the posterior are just as likely as any point of the distribution. Given the findings in the former, we proceed in this case, noting to the reader one must interpret our results with this in mind.

pose a zero "nominal rigidity" restriction on inflation in our baseline specification, although we relax this in our robustness analysis.

Table 1. Contemporaneous Sign Restrictions
Notes: This table reports the contemporaneous response of: the Bank Rate, r_t; public sector debt sales to the banking sector as a percentage of the broad money stock, psc_t; the long-term (i.e., 10-year gilt) gilt yield, lty_t; inflation, π_t; and output gap, y_t to structural shocks. $\mathbf{v}_t^{\mathbf{mp}}$ is a monetary policy shock; $\mathbf{v}_t^{\mathbf{q}}$ is a bank funding shock; $\mathbf{v}_t^{\mathbf{s}}$ is a supply shock; $\mathbf{v}_t^{\mathbf{d}}$ is a demand non-policy shock; and $\mathbf{v}_t^{\mathbf{lty}}$ is a long-term yield shock. 0 denotes a zero response on impact, "x" denotes no restriction.

	$\mathbf{v}_t^{\mathbf{mp}}$	$\mathbf{v}_t^{\mathbf{q}}$	$\mathbf{v}_t^{\mathbf{s}}$	$\mathbf{v}_t^{\mathbf{d}}$	$\mathbf{v}_t^{\mathbf{lty}}$
r_t	$\geq$	0	x	$\geq$	0
psc_t	x	$\geq$	x	x	$\geq$
lty_t	x	$\leq$	x	$\geq$	$\geq$
π_t	$\leq$	0	$\leq$	$\geq$	$\leq$
y_t	$\leq$	$\geq$	$\geq$	$\geq$	$\leq$

Canova and Paustian (2011) argue the importance of identifying all theoretically plausible shocks in order to refine inference for the shock of interest. Therefore, the restrictions we impose to identify monetary policy, $\mathbf{v}_t^{\mathbf{mp}}$, supply $\mathbf{v}_t^{\mathbf{s}}$, and demand non-policy shocks, $\mathbf{v}_t^{\mathbf{d}}$, stem from a standard New Keynesian model and are popular within the literature (see e.g. Ellington, 2022, among others). It is important to note that in using the output gap, $\mathbf{v}_t^{\mathbf{s}}, \mathbf{v}_t^{\mathbf{d}}$ are more precisely interpretable as a "cost push" and "net demand" shock, respectively; we name these as supply and demand for simplicity.[12]

The long-term yield shock, $\mathbf{v}_t^{\mathbf{lty}}$, is identified as a fall in the demand for risky assets and a "dash for cash" by non-bank investors in the gilt market. This raises gilt yields and public sector debt sales to the banking system either because the debt manager shifts by default towards bank funding (given the difficulty of funding from traditional investors), or because the central bank responds with more QE to stabilize long-term yields.

Collectively, the restrictions above are sufficient to separate the shocks from one another. As an initial diagnostic, Figure C.1 in Online Appendix C plots the time-series of posterior draws for our bank funding shock, $\mathbf{v}_\mathbf{t}^{\mathbf{q}}$, alongside the QE shocks Braun et al. (2025) identify based on high frequency analysis of intra-day announcements. Direct comparisons here should be made with caution given our shocks are based on quarterly surprises. Overall,

[12]We also use annual real GDP growth in place of the output gap. Results and conclusions are consistent with those we present in the main text and are in Online Appendix D. The only notable difference relates to unrealistic counterfactual analyses as a result of outliers due to the COVID-19 pandemic.

our bank funding shocks display some similarity to those of Braun et al. (2025) during QE periods, albeit the correlation coefficients at different lag lengths are relatively low at 0.1-0.2 over the UMP period.[13] However, their shock seldom falls out of the range of posterior draws of our bank funding shocks. Also, the relative size of QE shocks is similar across episodes in both cases, suggesting that earlier rounds of QE were perhaps less of a surprise than previously thought, given the high degree of anticipation suggested by surveys in later rounds (see Bernanke (2020)).[14]

4 Results

4.1 Model evaluation

We first evaluate the fit of the model using the deviance information criterion (DIC) in Spiegelhalter et al. (2002). The DIC contains two terms, one evaluating fit and a penalty term for model complexity. The DIC is given by

$$\text{DIC} = \overline{D} + p_D \tag{3}$$

where $\overline{D} = -2\mathbf{E}[\ln L(\mathbf{\Lambda}_i)]$ evaluates model fit with $\mathbf{E}[\ln L(\mathbf{\Lambda}_i)]$ denoting the average log likelihood over posterior draws, and $p_D = \overline{D} + 2\ln L(\mathbf{E}[\mathbf{\Lambda}_i])$ is model complexity with $\ln L(\mathbf{E}[\mathbf{\Lambda}_i])$ denoting the log likelihood at the posterior mean of parameter draws. The data better supports models with a lower DICs.

Table 2 shows the expected value of the log likelihood over posterior draws, $\mathbf{E}[\ln L(\mathbf{\Lambda}_i)]$, the penalty term for model complexity, p_D, and the Bayesian Deviance Information Criterion, DIC (Spiegelhalter et al., 2002). The first column reports results for our baseline, that is, the TVP-VAR we outline in 3. Column two of Table 2 shows analogous results for a standard BVAR using a Minnesota inverse Wishart prior. The next three columns consider models with different sources of time variation in parameters, namely a BVAR with stochastic volatility (BVAR-SV), a time-varying parameter VAR with constant covariance

[13] Specifically, the full-sample correlation (i.e. from 1997Q2–2024Q1) between the QE shock averaged (summed) over months in the quarter in Braun et al. (2025), and the posterior mean of $\mathbf{v}_t^{\mathbf{q}}$ lagged by two periods is 0.23 (0.19), if we use two lags to account for the VAR dynamics. Taking the sample 2009Q1–2024Q1 the analogous correlation is 0.24 (0.20). Based on a contemporaneous comparison the correlation drops to 0.1

[14] We consider alternative identification schemes for the bank funding shock and discuss these in our robustness analysis.

matrix (TVC-VAR), and a TVP-VAR with stochastic volatility (TVP-VAR-SV). We estimate such models following Chan and Eisenstat (2018); in all cases we use the same lag length and number of simulations as in our baseline case. Since these models require MCMC methods, we estimate five chains to produce DIC estimates and report the mean over the chains, along with the standard deviations in square brackets.

Table 2. Model evaluation
Notes: This table reports the expected value of the log likelihood over posterior draws, $\mathbf{E}[\ln L(\boldsymbol{\Lambda}_i)]$, the penalty term for model complexity, p_D, and the Bayesian Deviance Information Criterion, DIC (Spiegelhalter et al., 2002) for: our baseline Petrova (2019) TVP-VAR in the first column; a traditional Bayesian VAR with Minnesota inverse Wishart prior, BVAR, in the second column; while columns three to five show results for a BVAR with stochastic volatility, BVAR-SV; a time-varying parameter VAR with a constant covariance matrix, TVC-VAR; and a time-varying parameter VAR with stochastic volatility. The three latter use the estimation approach in Chan and Eisenstat (2018) and consider 5 chains to generate DIC statistics. Therefore the values we report are averages over the chains, and values in square brackets are the standard deviations. Models in columns one and two do not require Markov Chain Monte Carlo Simulation.

	Petrova (2019)		Chan and Eisenstat (2018)		
	TVP-VAR	BVAR	BVAR-SV	TVC-VAR	TVP-VAR-SV
$\mathbf{E}[\ln L(\boldsymbol{\Lambda}_i)]$	-647	-2027	-1136 [1.79]	-1357 [0.86]	-1260 [0.26]
p_D	192	111	150 [3.56]	68 [0.41]	79 [1.15]
DIC	1486	4164	2422 [7.14]	2782 [1.83]	2599 [1.50]

It is clear our baseline TVP has the most support from the data with a DIC considerably lower than all competing models. The TVP-VAR has the highest penalty term of 192, yet the log likelihood averaged over posterior draws is substantially larger, at -647, than the alternative specifications we consider. The BVAR possesses the lowest support from the data with $\mathbf{E}[\ln L(\boldsymbol{\Lambda}_i)]$=-2027, p_D=111, and a DIC of 4164. Looking at the specifications with time variation the TVC-VAR has the next highest DIC (averaged over chains) of 2782, then comes the TVP-VAR-SV with a DIC of 2599, and the BVAR-SV ranks second with a DIC of 2422. The standard deviations over the chains are relatively tight, with the BVAR-SV displaying the largest standard deviation across all statistics. Overall, Table 2 leads us to conclude that the data best supports our baseline TVP-VAR.

4.2 Bank funding shocks across QE/QT periods and historical funding episodes

First, we consider the impact of bank funding shocks during the historical debt management operations in the UK occurring before QE/QT. Figure 4 shows the impulse response functions that we average over historical funding episodes. We normalize the impulse responses such that a bank funding shock causes psc_t to increase broad money by 1% on impact. The green lines and shaded areas represent the episodes shown prior to the forward slash on the y-axis label. The black lines and shaded areas represent episodes following the forward slash on the y-axis label. This shows clear differences in the impact of UMPs over various historical funding regimes. The sensitivity of economic variables changes as we move through time as policy and the structure of financial markets change. In particular, the output gap and inflation are more sensitive on average during the underfunding episodes in the 1960s and 1970s relative to the overfunding episode of the 1980s. We explore changes in the transmission mechanism in more detail in Section 4.3.

Next, we examine the evolution of the impact of unanticipated bank funding shocks across the 5 phases of QE and initial phase of QT. Figure 5 reports the posterior median and 68% posterior credible intervals for the impulse response functions with respect to bank funding shocks that we average over each QE or QT episode.[15] We opt for average responses because the main focus of the paper is on differences *across* QE/QT episodes rather than differences *within* a given QE (or QT) episode. Again, we normalize the impulse responses such that a bank funding shock causes psc_t to increase broad money by 1% on impact, which is approximately equivalent to 1% on nominal GDP in the period since 2009.

The long-term yield declines in response to a unit bank funding shock in all QE episodes by up to 20 basis points; the impact is larger and more persistent during the earlier QE1 and QE2/3 episodes, which corresponds to higher uncertainty in financial markets over the 2008/09 financial crisis. The Bank Rate response is flat and insignificant in the early QE episodes, but increases during the later UMP episodes; peaking at 15 basis points in QE5. This gives us some scope to distinguish between the time variation in responses that result from changes in the portfolio rebalancing and signaling channels of QE, an issue we consider in Section 4.3.

Over QE1 and QE2/3, there is a similar peak impact on the output gap of around 0.16 percentage points and the impact remains significantly different from zero for 20 quarters. These are well within the range of estimates for QE1 Busetto et al. (2022) reports for GDP.

[15] We compute simple impulse response functions because parameter evolution is nonparametric in our model.

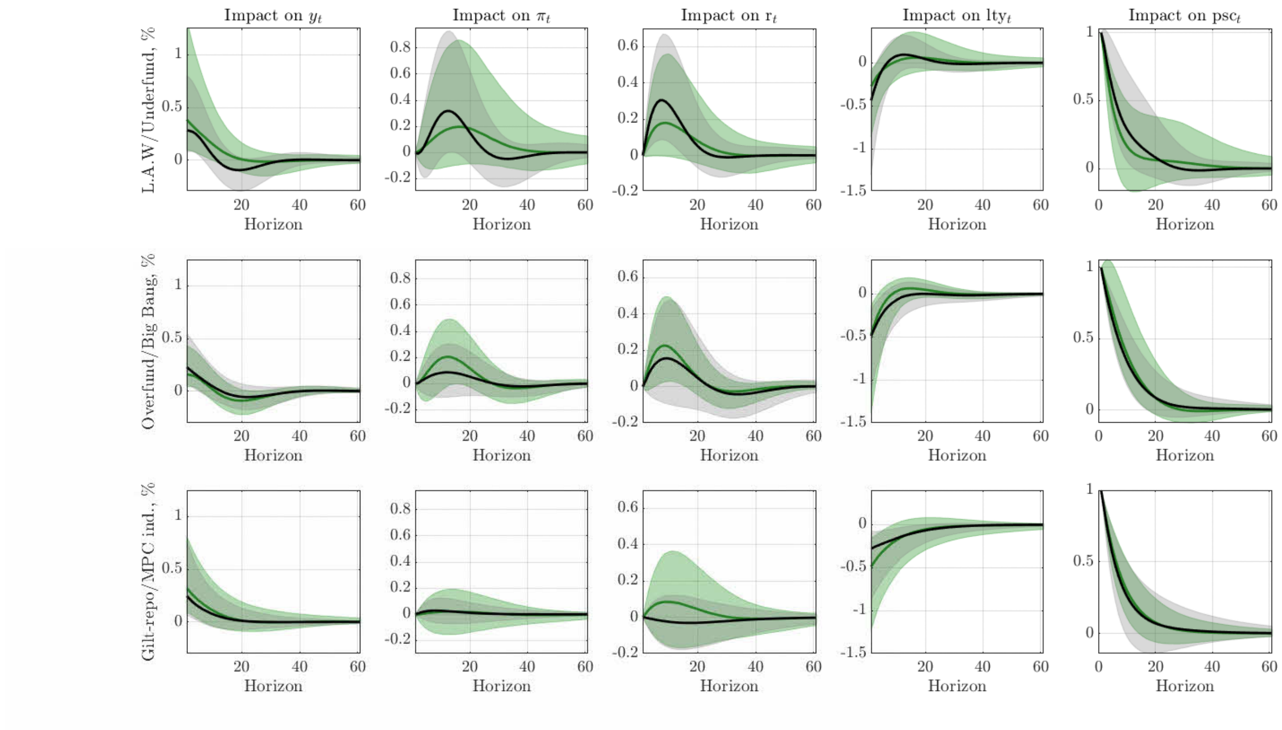

Figure 4. Impulse response functions to bank funding shocks: Averages over historical funding episodes

Notes: This figure plots the impulse response functions of all variables with respect to bank funding shocks averaged over historical funding episodes. The green lines and shaded areas represent the episodes prior to the forward slash on the y-axis label. The black lines and shaded areas represent episodes following the forward slash on the y-axis label. L.A.W is the "Lean Against the Wind" episode from 1963Q4–1971Q3. "Underfund", reflects the underfunding of deficits during the period of Competition and Credit control, the mid-1970s gilt strikes and the beginning of monetary targeting from 1971Q3–1981Q4. Overfund is the overfunding episode from 1982Q1–1986Q4, Big Bang reflects the introduction of the "Big Bang" reforms and the return to "full funding" over the period 1987Q1–1992Q4. Gilt-repo covers the more relaxed "full financing" rule and the development of the gilt repo market from 1993Q1–1997Q2. MPC ind is when the Bank of England obtained operational independence and the Debt Management Office takes over debt management from 1997Q3–2008Q4. y_t is the OBR's output gap; π_t is annual CPI inflation; r_t is the Bank Rate; lty_t is the 10-year gilt yield; and psc_t is public sector bank funding as a percent of M4. We normalize the distribution of the responses such that a one-standard deviation bank funding shock causes psc_t to rise by 1%.

The peak impact on annual CPI inflation is around half the size of the impact of GDP which is also consistent with the estimates of Joyce et al. (2011), Bridges and Thomas (2012) and Cloyne et al. (2015) .

During QE4 and QE5, the effect on the output gap is less persistent, but with the peak impact being slightly higher at around 0.18-0.20 percentage points, despite the lower and less persistent impact on gilt yields. Relative to earlier QE episodes, the peak response

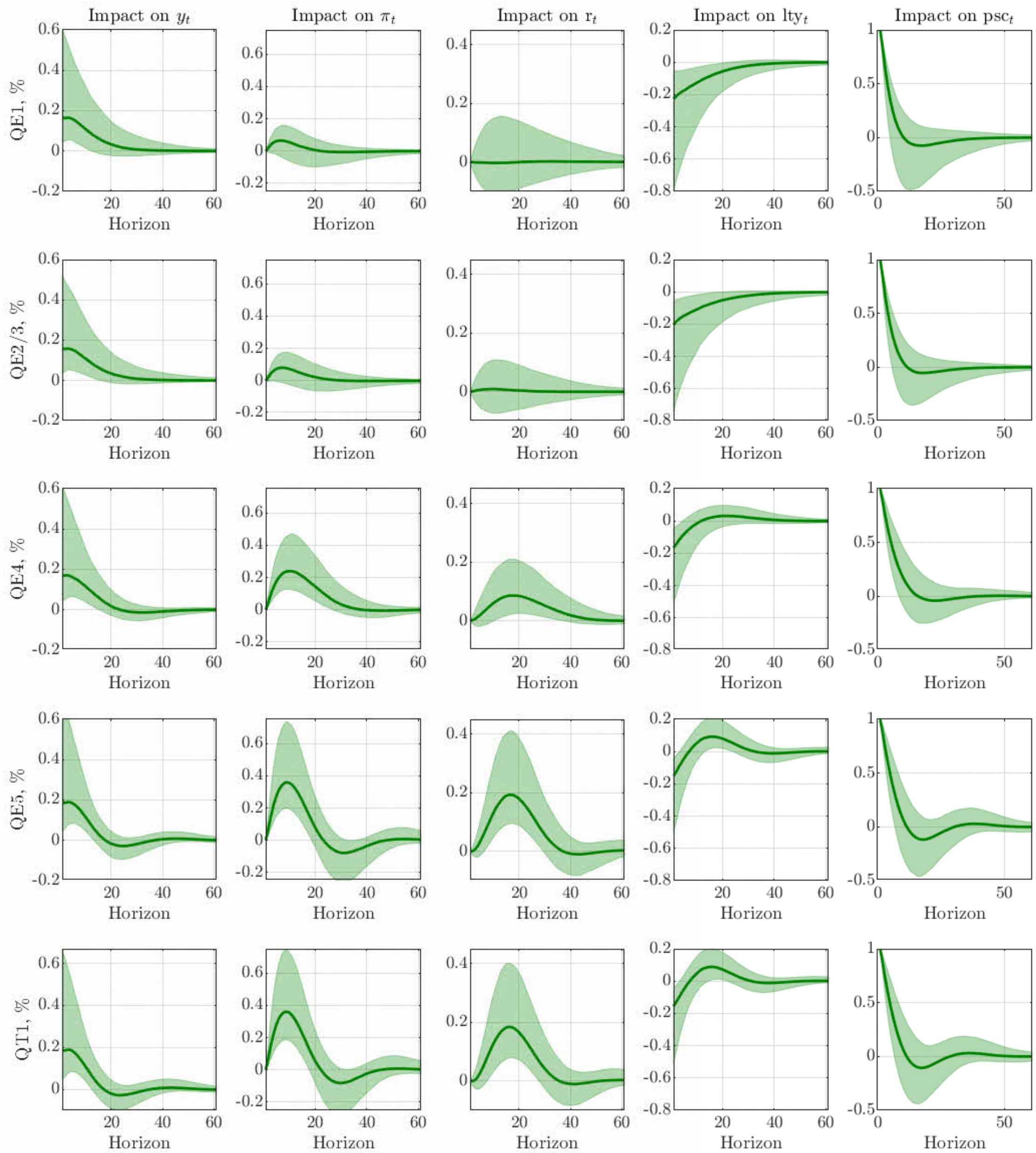

Figure 5. Impulse response functions to bank funding shocks: Averages over QE/QT episodes
Notes: This figure plots the impulse response functions of all variables with respect to bank funding shocks averaged over unconventional monetary policy episodes. QE1 is from 2009Q1–2010Q1; QE2/3 is from 2011Q4–2012Q4; QE4 is from 2016Q3–2017Q1; QE5 is from 2020Q1–2021Q4; QT1 is from 2022Q1–2024Q1. y_t is the OBR's output gap; π_t is annual CPI inflation; r_t is the Bank Rate; lty_t is the 10-year gilt yield; and psc_t is the public sector contribution to M4/M4ex growth. We normalize the distribution of the responses such that a one-standard deviation bank funding shock causes psc_t to rise by 1%.

of inflation is also stronger during QE4, QE5, and QT, at 0.35 percentage points. During QT, and for a shock of equivalent size, the output gap, inflation and Bank Rate rise in a similar manner to the QE5 episode; possibly due to the two periods' proximity to each other. The main takeaway point is there is considerable time variation in the magnitude and persistence of responses to different QE/QT episodes.

To investigate this formally, we consider the joint distribution of the accumulated responses of macroeconomic variables with respect to bank funding shocks averaged over the historical debt management operations and QE episodes. Let $\Theta_i = \left(\theta_i^{[1]}, \theta_i^{[2]}, \ldots, \theta_i^{[R]}\right)'$ and $\Theta_k = \left(\theta_k^{[1]}, \theta_k^{[2]}, \ldots, \theta_k^{[R]}\right)'$ be matrices containing the R posterior draws of the h horizon accumulated impulse response functions of macroeconomic variables in periods i and k respectively. Then, the Hellinger distance, $\mathcal{H}(\Theta_i, \Theta_k)$ and posterior probability differences, $\varpi_{i,k}$, are given by

$$\mathcal{H}(\Theta_i, \Theta_k) = \sqrt{1-\varphi} \tag{4}$$

$$\varpi_{i,k} = \frac{1}{R}\sum_{r=1}^{R} \mathbb{I}\left[\left(\theta_i^{[r]} - \theta_k^{[r]}\right)' \nu > 0\right] \tag{5}$$

where Equation (4) satisfies $0 \leq \mathcal{H}(\Theta_i, \Theta_k) \leq 1$ and $\mathcal{H}(\Theta_i, \Theta_k) = 0$ if and only if $\Theta_i = \Theta_k$ and φ is the Bhattacharyya coefficient.[16] Hence as $\mathcal{H}(\Theta_i, \Theta_k) \rightarrow 1$ the lower the overlap, with $\mathcal{H}(\Theta_i, \Theta_k) = 1$ implying the distributions are disjoint. In Equation (5) $\nu = \frac{\mu_i - \mu_k}{||\mu_i - \mu_k||}$, and $\mathbb{I}$ is an indicator function taking values of 1 if the condition is satisfied, zero otherwise. Such a metric provides a nonparametric projection of pairwise differences between posterior draws from accumulated impulse response functions during periods i and k relative to the line connecting their posterior means. Hence, it is a simple measure of directional posterior separation, thus $\varpi_{i,k} \rightarrow 1$ ($\varpi_{i,k} \rightarrow 0$) indicates strong separation (in the reverse direction).[17]

[16] In our case where the posterior distribution of impulse response functions follow a multivariate normal, the Bhattacharyya coefficient is analytical and given by $\varphi = \frac{|\Sigma_i|^{1/4}|\Sigma_k|^{1/4}}{\left|\frac{\Sigma_i+\Sigma_k}{2}\right|^{1/2}} \exp\left(-\frac{1}{8}(\mu_i-\mu_k)^\top \left(\frac{\Sigma_i+\Sigma_k}{2}\right)^{-1}(\mu_i-\mu_k)\right)$. Here. Σ_i, Σ_k denote the covariance matrices of Θ_i, Θ_k, and μ_i, μ_k are vectors of the row means of Θ_i, Θ_k.

[17] Available on request are statistics pertaining differences in impulse responses functions pertinent to each economic variable. Overall, the same conclusion holds true when looking at these metrics.

Table 3. Testing for state contingencies in bank funding shocks

Notes: This table reports Hellinger distances (in Panel A) and posterior probability distances (in Panel B) of the joint distribution of impulse response functions of macroeconomic variables with respect to bank funding shocks over different $\{i, k\}$ periods in the sample 1964Q2–2024Q1. Columns refer to period i and rows refer to period k, hence elements in provide results concerning the joint posterior distribution of period i versus period k accumulated impulse response functions. Let $\Theta_i = \left(\theta_i^{[1]}, \theta_i^{[2]}, \ldots, \theta_i^{[R]}\right)'$ and $\Theta_k = \left(\theta_k^{[1]}, \theta_k^{[2]}, \ldots, \theta_k^{[R]}\right)'$ be matrices containing the R posterior draws of the h = 61 horizon accumulated impulse response functions of macroeconomic variables in periods i and k respectively. Then, the Hellinger distance is given by $\mathcal{H}(\Theta_i, \Theta_k) = \sqrt{1-\varphi}$ and satisfies $0 \leq \mathcal{H}(\Theta_i, \Theta_k) \leq 1$ and $\mathcal{H}(\Theta_i, \Theta_k) = 0$ if and only if $\Theta_i = \Theta_k$; φ is the Bhattacharyya coefficient. The posterior probability difference is given by $\varpi_{i,k} = \frac{1}{R}\sum_{r=1}^{R} \mathbb{I}\left[\left(\theta_i^{[r]} - \theta_k^{[r]}\right)' \frac{\mu_i - \mu_k}{||\mu_i - \mu_k||} > 0\right]$, with $\mathbb{I}$ being an indicator function taking values of 1 if the condition is satisfied, zero otherwise. L.A.W is "Leaning Against the Wind" from 1963Q4–1971Q3. "Underfund", is the underfunding period from 1971Q3–1981Q4. Overfund is the overfunding episode from 1982Q1–1986Q4, Big Bang is the introduction of the "Big Bang" reforms from 1987Q1–1992Q4. Gilt Repo spans the development of the gilt repo market from 1993Q1–1997Q2. MPC Ind. begins when the Bank of England obtained operational independence and the Debt Management Office takes over debt management from 1997Q3–2008Q4. QE (QT) refer to quantitative easing (tightening) periods. QE1 is from 2009Q1–2010Q1; QE2/3 is from 2011Q4–2012Q4; QE4 is from 2016Q3–2017Q1; QE5 is from 2020Q1–2021Q4; QT1 is from 2022Q1–2024Q1. We highlight in bold font the posterior probability differences that equal 0.95 or higher.

A: Hellinger distances

	Underfund	Overfund	Big Bang	Gilt Repo	MPC Ind.	QE1	QE2/3	QE4	QE5	QT1
L.A.W	0.90	0.93	0.98	0.94	0.86	0.76	0.92	0.93	0.89	0.89
Underfund		0.94	0.99	0.97	0.95	0.94	0.99	0.98	0.96	0.97
Overfund			0.98	0.97	0.93	0.94	0.99	0.99	0.96	0.98
Big Bang				0.90	0.95	0.99	0.99	0.98	0.92	0.99
Gilt Repo					0.87	0.95	0.96	0.95	0.86	0.97
MPC Ind.						0.77	0.93	0.92	0.88	0.95
QE1							0.86	0.92	0.92	0.94
QE2/3								0.97	0.96	0.98
QE4									0.92	0.98
QE5										0.90

B: Posterior probability differences

	Underfund	Overfund	Big Bang	Gilt Repo	MPC Ind.	QE1	QE2/3	QE4	QE5	QT1
L.A.W	0.78	**0.96**	**0.96**	**0.96**	**0.97**	**0.97**	**0.97**	0.81	0.87	0.87
Underfund		0.85	0.91	0.94	**0.96**	**0.97**	**0.97**	0.79	0.91	0.92
Overfund			0.67	**0.95**	**0.98**	**0.97**	**0.98**	0.94	**0.98**	**0.97**
Big Bang				0.89	**0.95**	**0.94**	**0.95**	0.77	**0.97**	**0.96**
Gilt Repo					0.79	0.82	0.81	0.88	**0.98**	**0.96**
MPC Ind.						0.73	0.73	0.94	**0.97**	0.93
QE1							0.46	0.88	**0.96**	0.92
QE2/3								0.83	0.91	**0.97**
QE4									0.71	0.63
QE5										0.45

Table 3 reports the Hellinger distances in Panel A and the posterior probability differences in Panel B for the joint distribution of $h = 61$ quarter accumulated impulse response functions in periods i (columns) and k (rows) throughout our sample. We benchmark the n-th period against the remaining $n - 1$ periods we consider in Figure 4 and Figure 5. Comparing QE/QT periods with one another, Hellinger distances range from 0.85–0.99 thereby suggesting little overlap in the joint distribution of accumulated impulse response functions. Note also that the impact of bank funding shocks over QE/QT periods are distinct from historical funding episodes and policy regimes. The lowest Hellinger distances are 0.76 and 0.77. They stem from comparing QE1 with "leaning against the wind" and QE1 with the period 1997Q4–2008Q4 when the Bank obtained monetary policy independence (MPC Ind.), respectively; and still imply substantial differences.

Turning to Panel B, 56% of the 45 possible comparisons have a posterior probability difference greater than 0.95, with 73% exceeding 0.9. These statistics, albeit less so than the Hellinger distances, also provide strong evidence in favor of state contingencies over respective QE/QT periods. For instance, they suggest that the joint distribution of accumulated responses of our macroeconomic variables with respect to bank funding shocks in QE5 are larger than in QE1; the same holds true for QT1 versus QE4. Observe also that QE/QT periods, particularly QE1 and QE2/3 are distinct from funding regimes spanning the 1960s to early 1980s.

Taken together, Table 3 provides strong support in favour of state-contingent effects of QE and QT. Not only do we observe statistically credible differences over successive QE periods, but also substantial differences when comparing QE periods with the first phase of QT. In a historical context, we also observe clear state contingencies of recent UMPs with UMP–like episodes in the first part of our sample. We summarise the change in impulse responses over time in Figure C.2.

Overall, this section leads us to conclude that state contingencies are present across episodes of UMP and historical regimes, and hence, justifies the use of TVP-VAR models since by construction they allow for many regimes (Granger, 2008). This result also conforms with the work of, for example, Pettenuzzo and Timmermann (2017) who argue against the assumption of recurring regimes; we discuss results of models assuming recurring regimes in Section 5. The next step is to try and understand why and at what stage of transmission there is time variation in the magnitude and persistence of responses to different QE/QT episodes.

4.3 Transmission mechanism dynamics of unconventional monetary policies

To investigate the potential drivers of state contingency further, we break down the transmission mechanism of UMPs into different stages and calculate the time variation in some of the key semi-elasticities or slopes that theory suggests drive the response to bank funding shocks. This follows Kanngiesser and Willems (2024) who apply this method to the analysis of forecast errors.

At the first stage of the transmission mechanism, we look at the responsiveness of gilt yields to the quantity of asset purchases or sales. We distinguish between portfolio rebalancing and signaling channels by deriving a model-consistent measure of the term premium on government bonds as $\mathbf{tp_t} = \text{lty}_t - \mathbf{E}_t\left(\text{r}_{t:t+40}\right)$, where $\mathbf{E}_t\left(\text{r}_{t:t+40}\right) = \frac{1}{40}\sum_{i=1}^{40}\mathbf{irf}_{\text{r}_i}^{\mathbf{v^q}}$, where $\mathbf{irf}_{\mathbf{n_i}}^{\mathbf{x}}$ denotes the impulse response function of variable n at the ith horizon with respect to shock x. The first term is the current 10-year long-term yield rate, and the second term is the time t expectation of the average Bank rate over the next 40 quarters with respect to a bank funding shock. The sensitivity of the term-premium is the first-year average of the response of $\mathbf{tp_t}$, $\frac{1}{4}\sum_{i=1}^{4}\mathbf{tp_i}$. We interpret this as the strength of asset substitutability (i.e. portfolio rebalancing). We summarize the signaling channel by residual $\mathbf{E}_t\left(\text{r}_{t:t+40}\right)$ where we take the first-year average of $\frac{1}{4}\sum_{i=1}^{4}\mathbf{E}_t\left(\text{r}_{t:t+40}\right)$.

We also compute the slopes of the IS and Phillips curves (PC) implied by the impulse response functions to bank funding shocks. The former is the response of the output gap relative to the response of long-term yields, and the latter is the response of inflation relative to the response of the output gap at a suitable horizon. Such semi-elasticities allow us to examine at what stage of transmission the state contingency of unconventional monetary policies occurs. Specifically, we compute the slope of the IS curve and two different approaches for the slope of the Phillips curve as:

$$\Delta(\text{IS})_t = \frac{\max\left(\mathbf{irf}_t^{\mathbf{v^q},y_{0:60}}\right)}{\mathbf{irf}_t^{\mathbf{v^q},\mathbf{lty_0}}} \quad (6)$$

$$\Delta(\text{PC})_t = \frac{\max\left(\mathbf{irf}_t^{\mathbf{v^q},\pi_{0:60}}\right)}{\max\left(\mathbf{irf}_t^{\mathbf{v^q},y_{0:60}}\right)} \quad (7)$$

where $\Delta(\text{IS})_t$ takes the time t peak response of the output gap over the impulse horizon, $\max\left(\mathbf{irf}_t^{\mathbf{v^q},y_{0:60}}\right)$ relative to the time t impact response of long-term yields, $\mathbf{irf}_t^{\mathbf{v^q},\mathbf{lty_0}}$, with respect to bank funding shocks. $\Delta(\text{PC})_t$ takes the time t peak response of inflation, $\max\left(\mathbf{irf}_t^{\mathbf{v^q},\pi_{0:60}}\right)$, relative to the time t peak response of the output gap, $\max\left(\mathbf{irf}_t^{\mathbf{v^q},y_{0:60}}\right)$.[18]

We plot our breakdown of the transmission mechanism dynamics in Figure 6. First, we consider the sensitivity of gilt yields to bank funding shocks broken down into its term-premia and expected future rate components (i.e. the top left and right panels of Figure 6). We observe an increasing response (in absolute value) of the term-premium from 1964–1995. This likely reflects difficulties the BoE had in selling gilts during the volatile movements in financial markets during the 1970s. There is little initial impact from the Big Bang reforms in 1986 and the increases in sensitivity in the early 1990s might be linked to a large fiscal deficit creating uncertainty in the early 1990s. As fiscal policy tightens and the development of gilt repo market improves liquidity conditions, we observe a flattening trend in the 2000s.

The flattening of the term premium response continues over the QE episodes. The decline in the response of gilt yields we see in Figures 4–5 therefore appears predominantly due to changes in the responsiveness of expectations of future short-rates. This implies the signaling channel is stronger during QE1 than later episodes. During QE1, there is little response of short-term rates to the positive inflationary impact of a bank funding shock, whereas in QE2/3 and QE4 the response of the short-rate increases. This implies that in

[18] From a monetarist perspective the slope of the money demand function can be considered analogous to the response of **tp**. Additionally, we consider another variant of the slope of the Phillips curve as the 8-quarter ahead response of inflation relative to the 4-quarter ahead response of the output gap, $\Delta(\text{PC})_{1,t} = \frac{\mathbf{irf}_t^{\mathbf{v^q},\pi_8}}{\mathbf{irf}_t^{\mathbf{v^q},y_4}}$ where $\Delta(\text{PC})_{1,t}$ takes the time t two-year response of inflation, $\mathbf{irf}_t^{\mathbf{v^q},\pi_8}$ relative to the corresponding one-year response of the output gap, $\mathbf{irf}_t^{\mathbf{v^q},y_4}$ These results are in Online Appendix C.4.

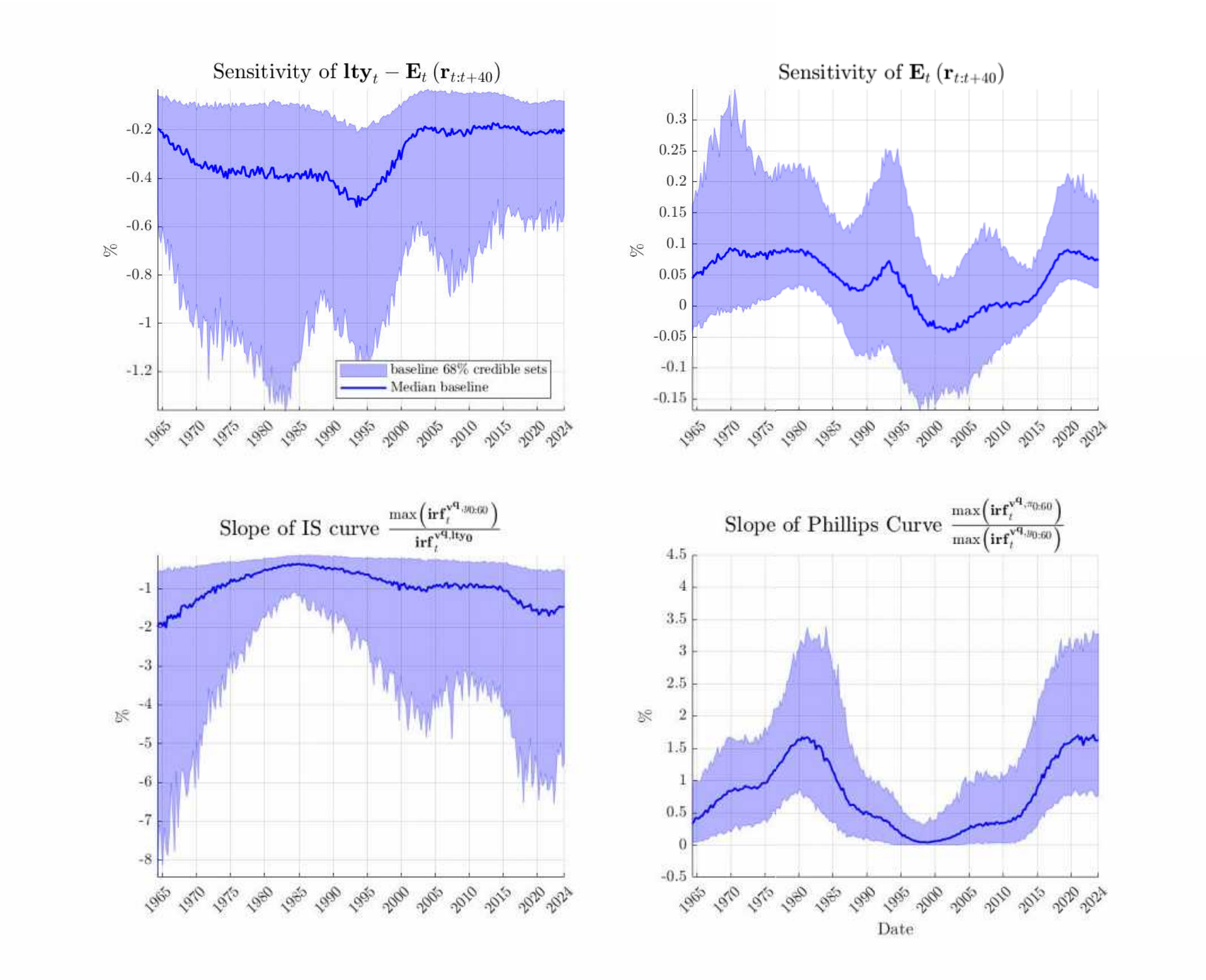

Figure 6. Transmission Mechanism of bank funding (v^q) shocks

Notes: The top left panel of this figure shows the sensitivity of the term premium which we proxy as the 1-year average of the current 10-year yield rate, lty_t, minus the time t expectation of the average level of Bank rate over the next 40 quarters. We back out the expectation from the impulse response functions of r_t with respect to a bank funding shock. This represents the asset substitutability channel throughout QE (QT) and other unconventional monetary policy episodes. The top right panel shows the 1-year average of the time t expectation of the average level of Bank rate over the next 40 quarters, which represents the signaling channel throughout QE (QT) and other unconventional monetary policy episodes. The bottom right panel shows the slope of the IS curve which represents the sensitivity of demand to yields. We proxy this as the ratio between time t peak horizon response of the impulse response function of the output gap to bank funding shocks, divided by the impact response of lty_t to the bank funding shock. The bottom right panel shows the slope of the Phillips curve which we proxy as the ratio of the peak horizon impulse response of inflation at time t with respect to bank funding shocks and the corresponding peak horizon response at time t of the output gap.

QE1 rates are expected to be kept "lower for longer" relative to the average policy reaction across all episodes. As the sensitivity of the term premium is flat across QE (QT) periods, the implication is that the impact of portfolio rebalancing was similar over each episode, but the signaling channel becomes progressively weaker. [19].

Turning to the slope of the IS curve (bottom panel of Figure 6), demand is more responsive to gilt yields during the underfunding episodes of the 1960s and 1970s. These are of similar magnitude to those during the latter QE episodes and subsequent QT policies. Overall, the sensitivity of demand is greater during later QE episodes relative to those in 2009–2013. This implies households and firms react more strongly to the higher wealth and lower cost of capital arising from lower long-term yields. Given the high levels of uncertainty in the financial crisis, this might indicate increases in confidence and reductions in economic uncertainty are a source of state contingency at this stage of the transmission mechanism. Alternatively, it might imply other asset prices became more responsive to falls in gilt yields or that transmission through non-bank financial intermediaries (NBFIs), which typically operate at longer maturities, may have been more important in later episodes (Bank of England (2021)); we examine this further in Section 5.

Finally, we consider estimates of the (reduced-form) slope of the Phillips curve in the bottom right panel of Figure 6. The Phillips curve slope is relatively flat from the mid-1990s to late 2010s with a relationship of just under 0.5. The slope increases to around 1.5 towards the end of our sample in 2024; similar to the late 1970s and early 1980s. The state contingency apparent in the slope of the Phillips curve might suggest there were changes in underlying real and nominal rigidities as a result of supply side conditions following the pandemic in the latter years of our sample. This is supported in Online Appendix C.4 where we show the increase in the slope of the Phillips curve in response to monetary policy shocks over this period. We also show in our robustness analysis that neither the exchange rate or inflation expectations appear to be driving the change in the reduced-form slope of the Phillips curve. However, the quantitative change in the slope does appear sensitive to using a measure of core inflation in the VAR, as we show later. On

[19]This is consistent with MPC guidance for QT that the stock of assets would be unwound gradually and predictably and that Bank Rate was the marginal instrument of monetary policy. Any financial market impact of QT would therefore be expected to involve an offsetting conventional policy response (see Bank of England (2024))

the whole, these findings resonate well with Costain et al. (2022), and Inoue et al. (2024). These papers respectively find a flattening in the UK and US Phillips curves since 1980 and an increase in the slope of the UK and US Phillips curve since 2020.

4.4 Quantitative Easing/Tightening counterfactuals

Here, we present counterfactuals to examine the state-contingent impact of the various rounds of QE and QT on the UK economy. As discussed in Bernanke (2020) and Busetto et al. (2022), it is important to distinguish between the unanticipated impact of QE and its total impact which incorporates the systematic or expected response of QE already built into expectations. For example, many event study analyses of QE are only able to pick up the unanticipated effect of QE in the short windows following QE announcements. If QE is already priced into financial market prices or only operates with a lag after policy announcements or auctions, the total impact of QT on gilt rates could be underestimated by these methods.

Our approach to QE/QT counterfactuals first involves assessing the conditional path of each economic variable, holding the amount of QE/QT in successive rounds constant. From this we can calculate the total implied response of each variable to the amount of QE/QT in each episode. We then decompose those impacts into the anticipated or "systematic policy" component of QE/QT and the unanticipated or "policy shock" component. The unanticipated component is the marginal impact of the actual bank funding shocks we observe in a given QE (QT) period (and assumed zero thereafter). The anticipated or systematic component of QE/QT is then the difference between the total and unanticipated path response of each variable.

First, we evaluate the total impact of QE (QT) through isolating the set of bank funding shocks that deliver the total amount of QE/QT in a given period. We construct the flow of QE (QT) for each period which delivers the amount of QE's (QT's) contribution to the broad money stock, call this $\widehat{QF}$[20]. This is a vector of length $H = 61$, which corresponds to the impulse response horizon. This vector contains non-zero elements equal to the flow of QE (QT) in a given episode, and zeroes otherwise. We then define a $H \times H$ matrix Q

[20]Because $\widehat{QF}$ is a contribution to four-quarter broad money growth, there is an overlap in the effects of QE5 and QT in the 2022Q1 to 2022Q3 period. To minimise the effects of this overlap, we introduce QT in 2022Q4 when active sales start, noting that the four-quarter growth rate in this quarter will still incorporate the passive redemption of gilts in 2022Q1.

that contains inversions of the time-series median impulse response functions of psc_t with respect to a bank funding shock over the QE (QT) episode each row of the matrix Q. Then we find the set of shocks, $\epsilon_t^{\mathbf{v^q},*}$, that minimizes the following objective function:

$$\epsilon_t^{\mathbf{v^q},*} = \min_x |Qx - \widehat{QF}_h|^2, \quad \text{where}$$

$$Q = \begin{bmatrix} \text{irf}_{1,1}^{\text{psc}_t} & 0 & 0 & 0 & \cdots & 0 \\ \text{irf}_{2,1}^{\text{psc}_t} & \text{irf}_{1,1}^{\text{psc}_t} & 0 & 0 & \cdots & 0 \\ \text{irf}_{3,1}^{\text{psc}_t} & \text{irf}_{2,1}^{\text{psc}_t} & \text{irf}_{1,1}^{\text{psc}_t} & 0 & \cdots & 0 \\ \text{irf}_{4,1}^{\text{psc}_t} & \text{irf}_{3,1}^{\text{psc}_t} & \text{irf}_{2,1}^{\text{psc}_t} & \text{irf}_{1,1}^{\text{psc}_t} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ \text{irf}_{H,1}^{\text{psc}_t} & \text{irf}_{H-1,1}^{\text{psc}_t} & \text{irf}_{H-2,1}^{\text{psc}_t} & \text{irf}_{H-3,1}^{\text{psc}_t} & \cdots & \text{irf}_{1,1}^{\text{psc}_t} \end{bmatrix}$$

The minimization problem delivers the counterfactual set of bank funding shocks that match the flow of QE (QT) for each episode, such that $\epsilon_t^{\mathbf{v^q},*}$ is a $H \times 1$ vector. We assume the set of shocks represent "modest interventions" in each episode as in Leeper and Zha (2003). This is a strong assumption as it assumes that if QE had not responded at all to economic shocks, all the other VAR parameters would be unchanged. But it does provide a useful benchmark counterfactual to evaluate how big the total effects of QE/QT might be relative to the unanticipated component. Given the set of shocks, we then invert the impulse response functions of all other variables and perform an element-wise multiplication by the cumulative sum of $\epsilon_t^{\mathbf{v^q},*}$ over the impulse horizon. This delivers an estimate of how much each unconventional monetary policy episode contributes to the output gap, inflation and long-term yields based on the amount of QE (QT) in the period.

Next, to compute the marginal impact of unanticipated QE/QT, we take estimates of the actual bank funding shocks we observe in a given QE (QT) period and hold them to zero thereafter. We then invert the median impulse response functions over the given QE (QT) episode, and perform an element-wise multiplication by the cumulative sum of such shocks. This counterfactual delivers the marginal impact of unanticipated QE (QT) on each economic variable. Such impacts are a closer comparison to the effects that event studies of QE operations capture (see e.g. Busetto et al., 2022; Braun et al., 2025), which form the bulk of the range of estimates in the literature.

We plot the posterior median response of the output gap, inflation and long-term yields from our counterfactuals in Figure 7. The top row shows the total impact of QE (QT) (i.e. derived from our first counterfactual that assumes QE (QT) did not respond to any of the economic shocks in the model). The second row shows the unanticipated proportion of QE (QT) by episode; i.e. the amount of QE/QT attributable to the $\mathbf{v_t^q}$ shock alone. In Online Appendix C.3 we report corresponding plots with 68% posterior credible intervals in Figures C.6 and C.7. All counterfactual responses are credibly different from zero.

First, we make a comparison of the relative size of the unanticipated and systematic effects of QE/QT in each episode. This provides an estimate of how much larger the systematic effect might be relative to the unanticipated effects, captured by event studies and other methods that focus on QE surprises[21]. For QE1, the proportion of the unanticipated component attributable to the total response of each variable at respective peak is around 10%. We observe similar proportions attributable to the unanticipated component of QE2/3. In QE4, the proportion attributable to unanticipated components hover around 40–47%. However, the overall impact of QE4 is substantially smaller relative to other episodes given the low amount of gilts purchased in the scheme. For the QE5 and QT1 counterfactuals, the proportion of the output gap, inflation and gilt yield responses attributable to the unanticipated component are also higher than QE1 and QE2/3. In QE5, the unanticipated component accounts for 15%, 18%, and 15% of the total response of the output gap, inflation and yields, respectively. The corresponding percentages attributable to the total output gap, inflation, and gilt yield response during QT are 11%, 27%, and 20%, respectively.

Overall, this indicates the unanticipated component contributes relatively small amounts to total movements in our baseline results. Our estimates suggest one could have been able to anticipate the lion's share of QE/QT policy from the underlying shocks that hit the economy. This is also shown in the historical decomposition of psc_t in Figure C.5. If this is also reflective of financial market participants' expectations of the QE/QT response, as for ex-

[21] Note event study impacts can be scaled up to give a total impact if they can be combined with estimates of the surprise component of the announced quantities, for example, from surveys of market participants made immediately prior to the announcement.

ample Braun et al. (2025) suggest, it implies a good understanding of the reaction function of the MPC over the 2009–2024 period.[22]

Next we examine the total peak responses across episodes and compare them to other estimates in the literature. We focus on the median effects here but Figures C.6–C.7 indicate the significant degree of uncertainty around these central case estimates. During QE1 (i.e. 2009Q1–2010Q1), the total median impact of £200bn of gilt purchases (13 per cent of initial M4ex) increases the output gap and inflation by 2.8 percentage points and 0.9 percentage points, respectively; these are within the range of impacts in Joyce et al. (2011). However, QE1 compresses gilt yields by around 200 basis points on average over the period. This is much larger than the 100 basis points that, e.g. Joyce et al. (2011) and Busetto et al. (2022) report for the UK based on event studies. However, as discussed above, these studies are more comparable with the impact of the unanticipated component from our counterfactual analysis. In this case, the gilt yield impact of around 20-35 basis points for a £20bn surprise in QE is smaller than the 100 basis points impact from a £140bn QE surprise in Busetto et al. (2022).

The impact of QE2/3, taking place during 2011Q4–2012Q4 has a smaller impact on the output gap and yields than QE1, commensurate with a slightly lower amount of purchases (£175bn, or 11 per cent of M4ex), but with an identical peak impact on inflation, relative to QE1. The unanticipated impact shows a temporary spike in yields of 40bps but is relatively small over the period as a whole and similar to QE1 although shorter lived.

During QE4 the total response of both the output gap and inflation peak at around 0.6 percentage points from £60bn of gilt purchases (3 per cent of M4ex). The steepening of the Phillips curve slope to around 1 contributes to the relatively high inflation impact

[22]This is evident in Section C.2 of the Online Appendix. Here we report historical decompositions of all economic variables with a decomposition of the deterministic component into "equilibrium" quantities (i.e. natural rate of interest, trend inflation etc.) following Petrova (2019). We can see that the decomposition of psc_t itself assigns a small proportion of variation to bank funding shocks relative to the other structural shocks we identify. This shows the importance of the systematic element of funding operations in the historical episodes prior to 2009 and the QE/QT reaction function thereafter. We can also see that supply side conditions (and demand) are increasing in importance for the output gap and inflation across QE episodes, especially during 2022 and 2023. In general, the historical decompositions show a large proportion of variation attributable to equilibrium movements in inflation and real interest rates relative to structural shocks looking at the whole sample. Note also Bank funding shock contributions are different to the unanticipated QE/QT counterfactuals presented above, given they include the lagged effects of commercial bank purchases/sales of gilts prior to each QE episode.

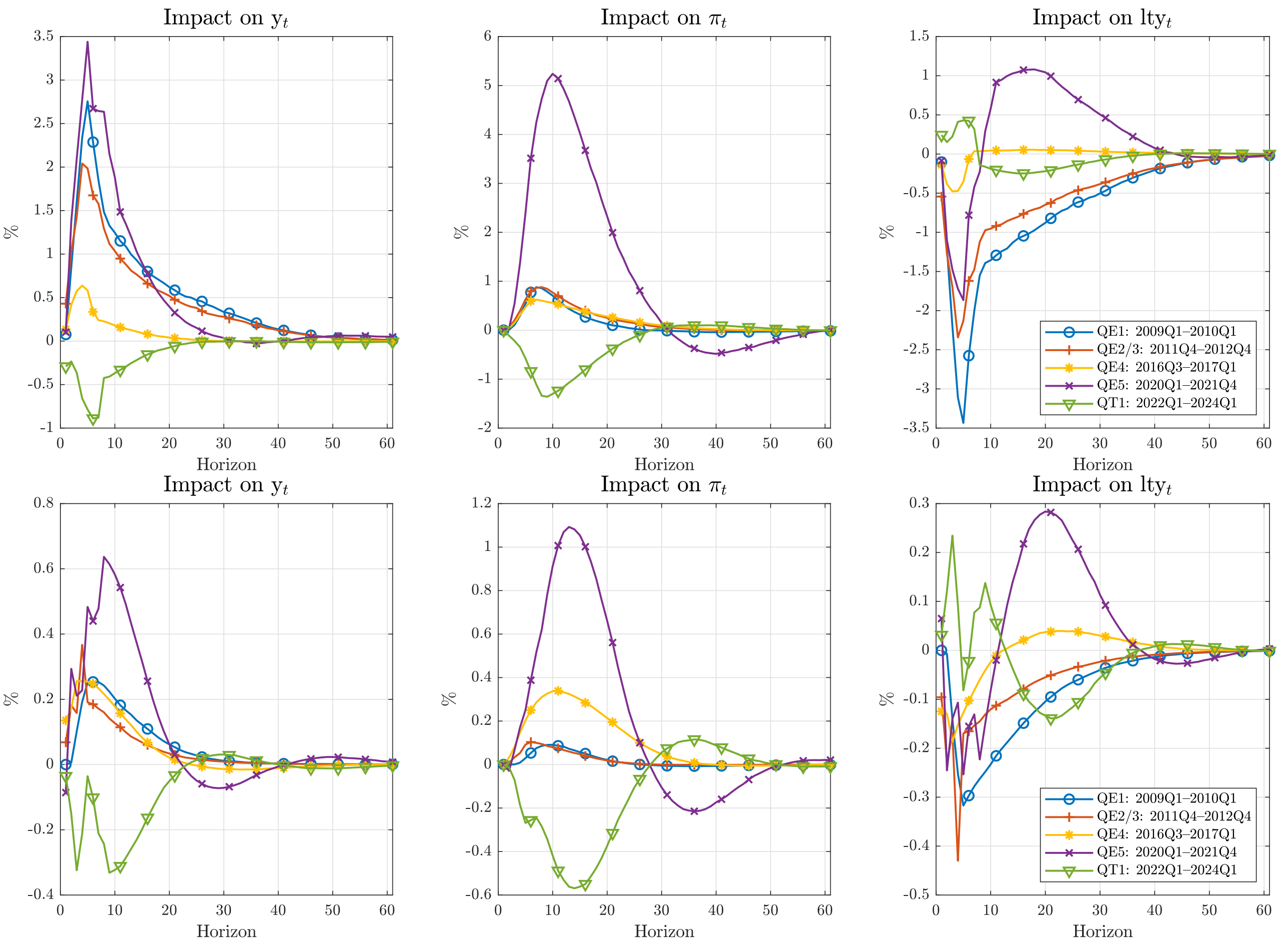

Figure 7. Counterfactual experiments: Quantitative Easing/Tightening episodes

Notes: This figure plots the posterior median counterfactual paths of the output gap, y_t; CPI inflation, π_t; and long-term yields, lty_t across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. The counterfactual in the first row shows the total impact that each round of QE (QT) as discussed in the text. The counterfactual in the second row shows us the unanticipated proportion of QE (QT).

compared to QE1. Yields are reduced by 48 basis points at their peak of which around 20bps is unanticipated.

QE5, from 2020Q1–2021Q4, which was undertaken in response to the pandemic, generates peak counterfactual effects on the output gap and inflation of 3.4 percentage points and 5.2 percentage points, respectively, as a result of £440bn of gilt purchases (18 per cent of M4ex); while compressing yields by an average of 125 basis points. Our estimates suggest most of that was anticipated given the size of the shocks during the pandemic. The large inflation impact reflects the further steepening of the Phillips curve over this period. However, this steepening occurred alongside large increases in international goods prices coming out of the pandemic reflecting bottlenecks in global supply chains. This was exacerbated by the Ukraine invasion in 2022Q1 that caused food and energy prices to increase. These factors are not explicitly captured in our VAR but other studies have emphasised their importance in this period (see Bernanke and Blanchard (2025), Haskel et al. (2025) and Bank of England (2026)). We test the robustness of our results to factoring in the impact of international commodity prices in Section 5. Further out, the yield impact from QE5 is estimated to overshoot given that, unlike QE1, Bank Rate is expected to respond to inflation and the output gap as they recover, reflecting the absence of a signaling effect in this episode.

QT started in 2022Q1, when the Bank stopped reinvesting maturing assets in the Asset Purchase Facility and active sales of gilts commenced in 2022Q4. This reduced the stock of assets by just under £150bn (5 per cent of M4ex) by the end of our sample in 2024Q1. The impact on yields is smaller compared to earlier estimates, because Bank Rate is expected to respond reflecting the lower signaling channel we report in Figure 6. The unanticipated impact suggests a 10-20 basis points rise in yields in line with the Bank of England's initial event study analysis (see Bank of England (2024)) but lower than the cumulated effect of 70bps from QT announcements found by Du et al. (2024).[23] This has a small impact on the

[23] Du et al. (2024) warn that their estimate includes announcement effects which may be insignificant. More to the point, our work is not directly comparable to that of Du et al. (2024). This is because we rely on a long history of quarterly data and use counterfactuals to assess the impact of QT on yields within a model which is flexible enough to capture time-variation and the endogenous effects of Bank Rate, output gap, inflation and the long-term yield to QE/QT shocks. On the other hand, Du et al. (2024) rely on daily data over the 2021-2023 period and a panel data framework to assess the impact of QT within a two-day window.

output gap of between 0.1-0.3 percentage points and a similar amount on inflation. The broader impact, including estimated anticipated effects, suggests a peak effect on yields of around 40 basis points that would reduce the output gap in the median case by almost 0.9 percentage points and inflation by 1.4 percentage points, given the steep estimated slope of the Phillips curve in 2023. However, one important caveat here is that commercial bank purchases of gilts partly offset the sales by the Bank of England over this period, unlike earlier episodes (see Bank of England (2025) and Figure A.1). That led to a lower reduction in broad money than if UK preferred habitat investors had been the sole buyers of gilts during QT. To the extent this was an endogenous response by commercial banks to QT (e.g. to offset the decline in liquidity from fewer commercial bank reserves), represents an additional source of stage contingency. There were also some offsets in the opposite direction during QE2/3 and QE5. This is an issue we address in the next Section.[24]

5 Robustness checks and further analysis of the drivers of state contingency

In this section we summarise the robustness of our findings to different data choices and modelling methods. The underlying details are provided in an Online Appendix. We also exploit the flexibility of our non-parametric TVP-VAR approach to examine expanded versions of the VAR that include extra variables that might provide information on the drivers of state contingency.

5.1 Alternative data and identification choices

We first consider two alternative proxies to track unconventional monetary policy through the lens of public sector borrowing from banks. Our first alternative considers using only Bank of England government debt purchases in the numerator in our initial measure. The second alternative considers debt sales to the banking system as a proportion of the par value of gilts. This replaces the denominator in our initial measure. The data we use to construct such alternatives are from the Bank of England's Millennium of Macroeconomic Database.

Figure 8 plots these variables. In general, the time profiles are similar to our baseline

[24] Figure 9 shows, for the relevant QE and QT periods, using an alternative measure of psc_t that should in principle incorporate the offset, delivers a slightly smaller peak total impact on gilt yields.

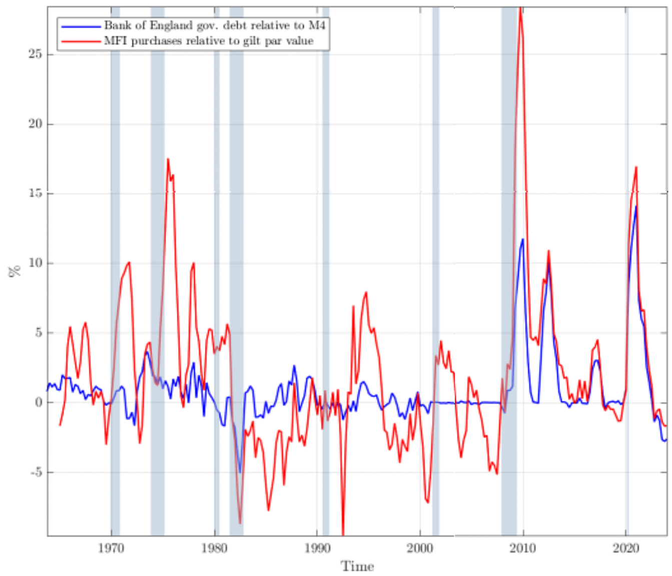

Figure 8. Alternative proxies to track unconventional monetary policy 1963Q4–2024Q1
Notes: This figure plots alternative funding measures to capture unconventional monetary policy. The blue line reports the annual amount of government debt purchased by the Bank of England only as a proportion of the broad money supply (i.e. M4). The red line plots the annual amount of total monetary financial institutions (MFI) purchases relative to the (total) par value of gilts. Grey bars indicate recessions.

measure of public sector bank funding. As we note in Figure 1c, the Bank of England typically had to offset its purchases of longer term securities by drains of reserves using sales of Treasury bills and other short-term debt in the period when reserves were unremunerated. So it's net purchases pre-2009 are on a much smaller scale than the QE operations which followed. However, this measure allows us to take into the account the potential response of commercial bank purchases of debt to QE/QT from 2009 discussed in the previous Section. The sample correlation between our baseline measure and Bank of England government debt purchases as a proportion of the broad money supply is 0.73. The sample correlation between our baseline measure and total monetary financial institutions' purchases as a proportion of the par value of gilts is far closer at 0.97.

We estimate the model in Equation 1 replacing psc_t sequentially with the above proxies in the exact manner as we outline in Section 3, as well as using the same identification scheme in Table 1. In Figure 9 we report the posterior median counterfactual paths of the output gap, y_t; CPI inflation, π_t; and long-term yields, lty_t across the QE and QT episodes in our sample. The counterfactual in the first row shows the total counterfactual impact implied by each round of QE (QT) using just Bank of England government debt purchases relative to M4 to estimate the model parameters. The counterfactual in the second row shows us corresponding results that uses total monetary financial institutions' purchases as a proportion of the par value of gilts outstanding. The counterfactual tells us the model-implied impact of each QE/QT episode and is directly comparable to the first row in Figure 7. Overall, it is clear that our alternative measures of bank funding imply little difference to

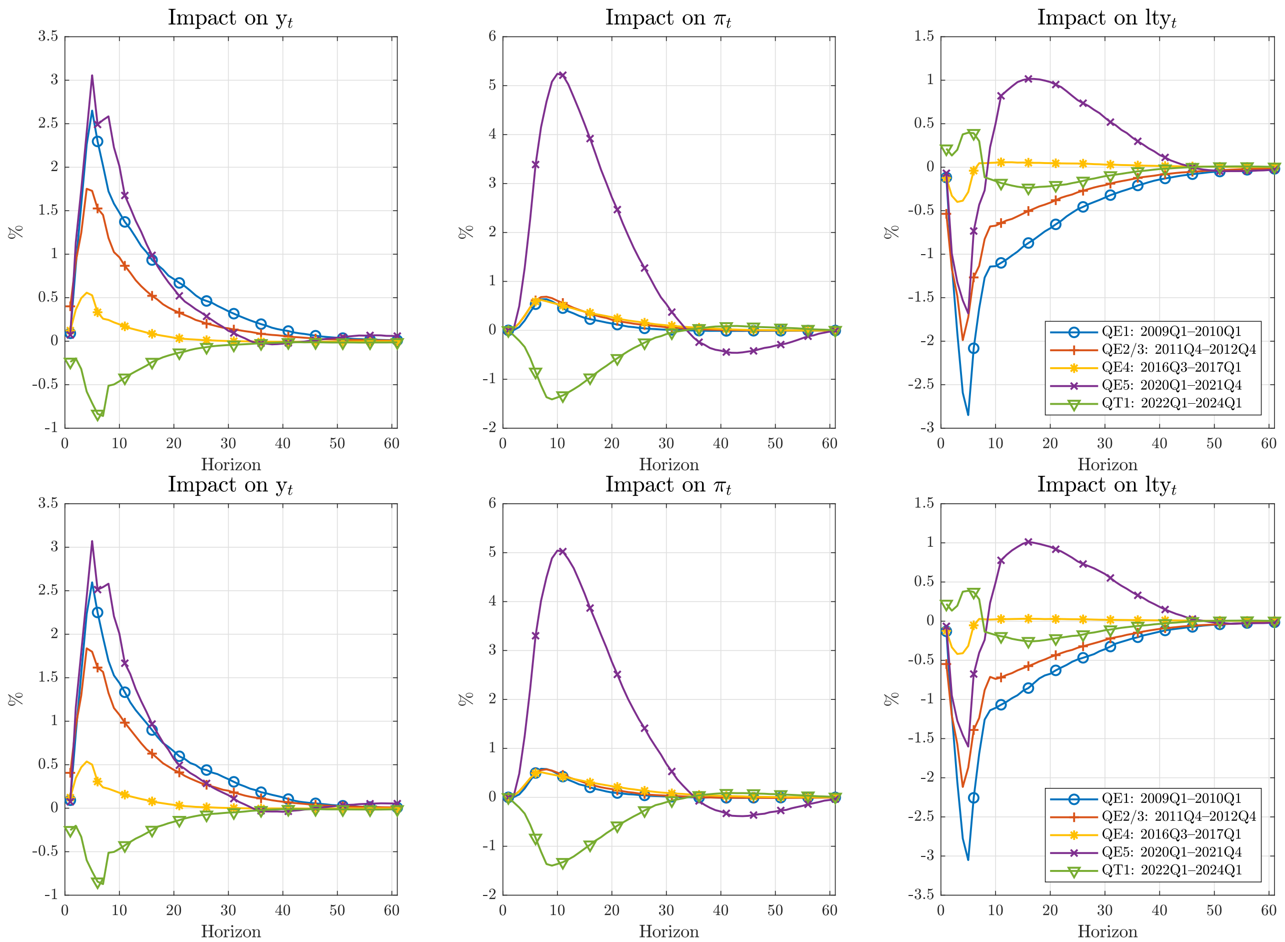

Figure 9. Counterfactual experiments with alternative bank funding proxies: Quantitative Easing/Tightening episodes
Notes: This figure plots the posterior median counterfactual paths of the output gap, y_t; CPI inflation, π_t; and long-term yields, lty_t across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. The counterfactual in the first row shows the impact that each round of QE (QT) has in isolation of all other shocks using the Bank of England government debt purchases relative to M4 as a proxy for public sector bank funding (we plot this measure in Figure 8). The counterfactual in the second row shows us corresponding results that uses total monetary financial institutions' purchases as a proportion of the par value of gilts (we plot this measure in Figure 8). The counterfactual tells us the model-implied impact of each QE/QT episode and is directly comparable to the first row in Figure 7.

the counterfactual paths over the QE (QT) episodes. The impact under our first alternative measure of psc_t delivers a slightly weaker effect on yields across most of the episodes, which may reflect the endogenous response of commercial bank purchases we discuss in the previous Section. In general, the posterior response of all variables exhibit marginal differences to those we report in Figure 7. In Online Appendix C.4, Figure C.10 shows the transmission mechanism plots analogous to those in Figure 6. Posterior median estimates are similar to those we report from our baseline specification. This leads us to conclude that our main results are robust when considering alternative measures to track asset the impact of unconventional monetary policies.

Next, we use annual real GDP growth as our proxy for economic activity rather than the output gap. As discussed earlier, this changes the interpretation of our demand and cost-push/supply shocks. However, impulse response functions for our bank funding shock yield the same conclusions to those we present in Section 4.

Next, we consider allowing for an explicit role for international commodity prices in the VAR. Arguably, the time-varying parameter estimates in our VAR might capture this through the responses to the cost-push shock, $\mathbf{v}_t^{\mathbf{s}}$, over time. However, the large role international commodity prices are argued to have played in particular periods, such as the oil price shocks of the 1970s, in the emergence from the Covid pandemic in 2021 and the Ukraine war after 2022, might act to distort our estimates of the slope of the Phillips curve unless they are explicitly taken account of. To assess this we replace total CPI inflation in the VAR with a measure of core CPI inflation that excludes the direct contributions from food, tobacco, alcohol and energy. The data for core inflation is taken from Bordo et al. (2025).

Figure 10 shows that the impact of QE and QT on core inflation is smaller in each episode than the impact on total CPI inflation in our baseline case. But our conclusion that the slope of the Phillips curve increases across episodes remains intact. Although, the total impact of QE5 on core inflation in QE5 is under 3 percentage points rather than 5 percentage points, it still suggests the slope of the Phillips curve rises to around 1 rather than 1.5 estimated in the baseline case. The estimates on other variables are broadly similar to the baseline case. These results suggest disentangling the effects of international prices from those of domestic supply constraints and price-setting behaviour in the period

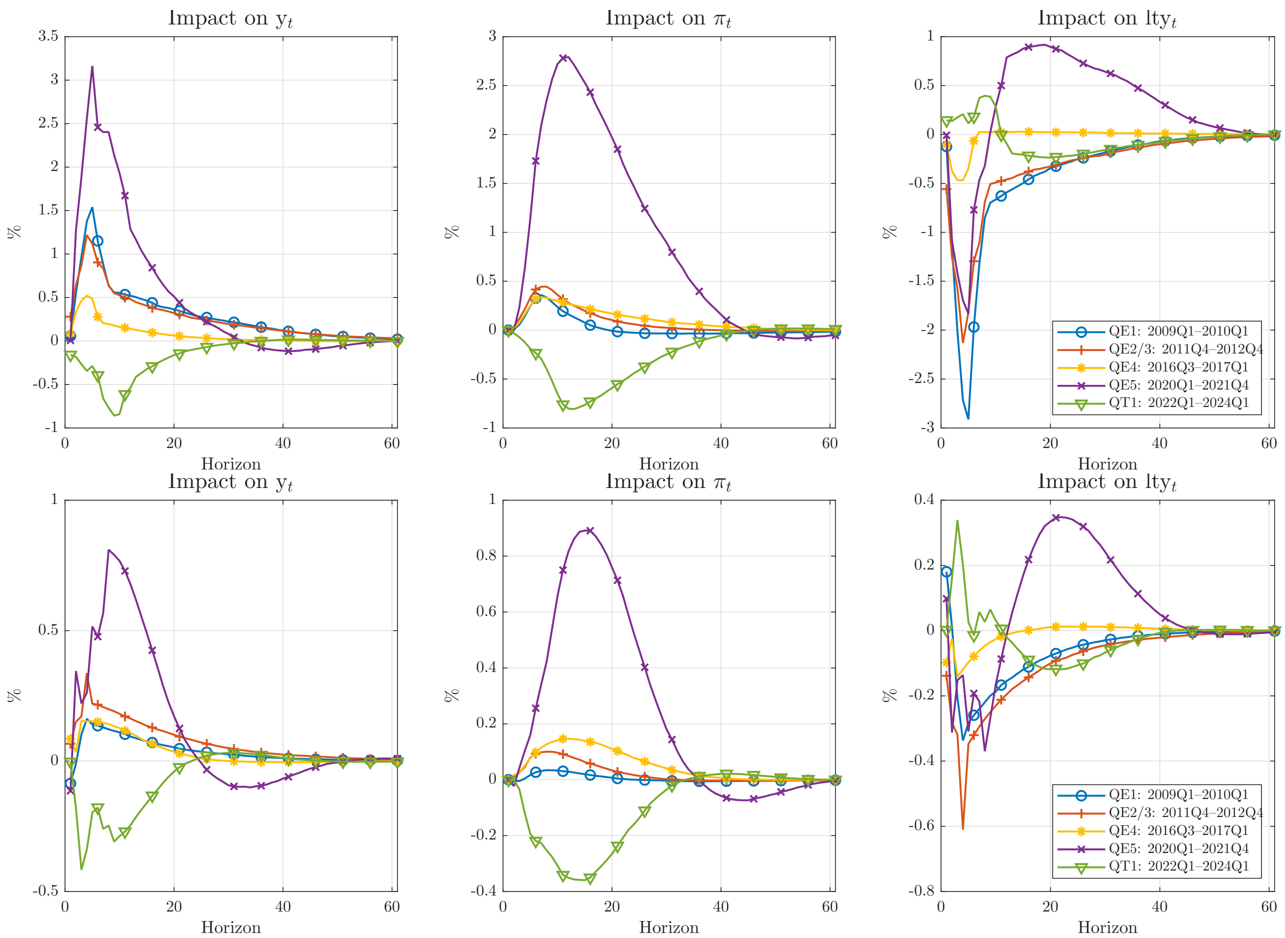

Figure 10. Counterfactual experiments using core inflation: Quantitative Easing/Tightening episodes

Notes: This figure plots the posterior median counterfactual paths of the output gap, y_t; core CPI inflation excluding food and energy, π_t; and long-term yields, lty_t across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. The counterfactual in the first row shows the total impact of each round of QE (QT) as discussed in the text. The counterfactual in the second row shows us the unanticipated proportion of QE (QT).

coming out of the pandemic is difficult (see Bank of England (2026) for a discussion) and sensitive to the choice of inflation rate used. But overall, our results suggests domestic supply conditions and price-setting behaviour do appear to play a significant role.

Finally, we consider two alternative identification schemes to map to the structural model in Online Appendix F. In particular, we alter the timing assumptions through which QE affects inflation and activity. These results deliver qualitatively similar conclusions and provide evidence that our findings on the state-contingent impact of QE/QT policies is not fundamentally driven by the identification scheme.

5.2 Alternative models of state contingency

In this section we consider alternative VAR models that in principle could capture state contingency. A full set of results for the alternative models is in Online Appendix G and we only provide a summary here.

In Section 4 we show that our model has better fit that a standard linear Bayesian VAR with a Minnesota prior or a TVP-VAR with stochastic volatility of Primiceri (2005). The linear BVAR can loosely be thought of as averaging out results from time-varying TVP VARs. We observe little in terms of significance from the bank funding shock from this model which provides support for the state-contingent nature of how such shocks, and unconventional monetary policy, affect the real economy. The Primiceri (2005) TVP-VARs deliver qualitatively similar conclusions to our baseline analysis. The only notable differences are that the QE/QT counterfactuals are slightly larger relative to those from our baseline results, yet credible sets overlap.

We also consider whether a sequence of "repeated regimes" might be a better characterisation of the state contingency we find. For this we apply a threshold VAR (TVAR) model as in Alessandri and Mumtaz (2017). We consider each of our 5 baseline variables as candidates that could govern regime changes. We also consider a battery of exogenous variables to govern regime changes including: Economic policy uncertainty (Baker et al., 2016); the volatility of UK stock returns; the budget deficit (surplus); and financial stress. We find no evidence in favor of asymmetries in terms of size or sign of shocks in a given regime, and marginal differences in posterior median estimates across regimes. This leads us to conclude that TVAR specifications, where regime changes occur due to an (endogenous) estimate of a given variable, are not flexible (or capable) enough to capture the state contingencies during QE/QT that time-varying parameter models show. This is consistent with our main results that highlight the need to study the state contingencies of QE/QT policies using models with non-recurring regimes (see e.g. Pettenuzzo and Timmermann, 2017).

5.3 Additional diagnostic variables of state contingency

Here we consider the robustness of our main results by sequentially adding an extra variable to our original dataset. The additional variables we consider are candidates that might shed light on the drivers of state contingency we identified earlier: i) 1-year ahead inflation expectations; ii) annual UK effective exchange rate changes; iii) annual stock market returns; iv) annual house price changes; v) investment-grade corporate bond spreads and vi) bank and building society lending based on the M4Lex measure, such that $z_t = \{\mathbf{E}_t(\pi_{t+4}),\ xr_t,\ s_t,\ h_t,\ cs_t,\ M4Lx_t\}$; these additional variables are obtained from Bordo et al. (2025) and the Bank of England's Millennium of Macroeconomic Database.[25] To identify the sixth structural shock, $\mathbf{v^z}_t$, we impose a zero impact response of economic variables and a positive impact response of z_t. We impose no restrictions on z_t with respect our original five structural shocks. We summarize our identification for 6-variable models in Table 4.

Figure 11 reports counterfactual experiments using posterior median estimates that are directly comparable to the first row in Figure 7. Each row corresponds to $z_t = \{\mathbf{E}_t(\pi_{t+4}),\ xr_t,\ s_t,\ h_t,\ cs_t,\ M4Lx_t\}$. Along with the counterfactual paths of the output gap, inflation, and long-term yields, we also plot the counterfactual response of z_t. Overall, the responses of y_t, π_t, lty_t are almost indistinguishable from our main results. That means the addition of each individual variable and associated shock have little impact on the core responses.

We first look at the response of inflation expectations and the exchange rate across episodes, which might help explain the rise in the slope of the Phillips curve implied by our QE responses. We can see that inflation expectations initially fall rather than rise in

[25]Measures of money in UK macroeconomic models are considered in, among others, Florackis et al. (2014). We also consider additional variables including economic policy uncertainty (Baker et al., 2016), which Ellington et al. (2023) show relates to UK inflation, interest rates and output.

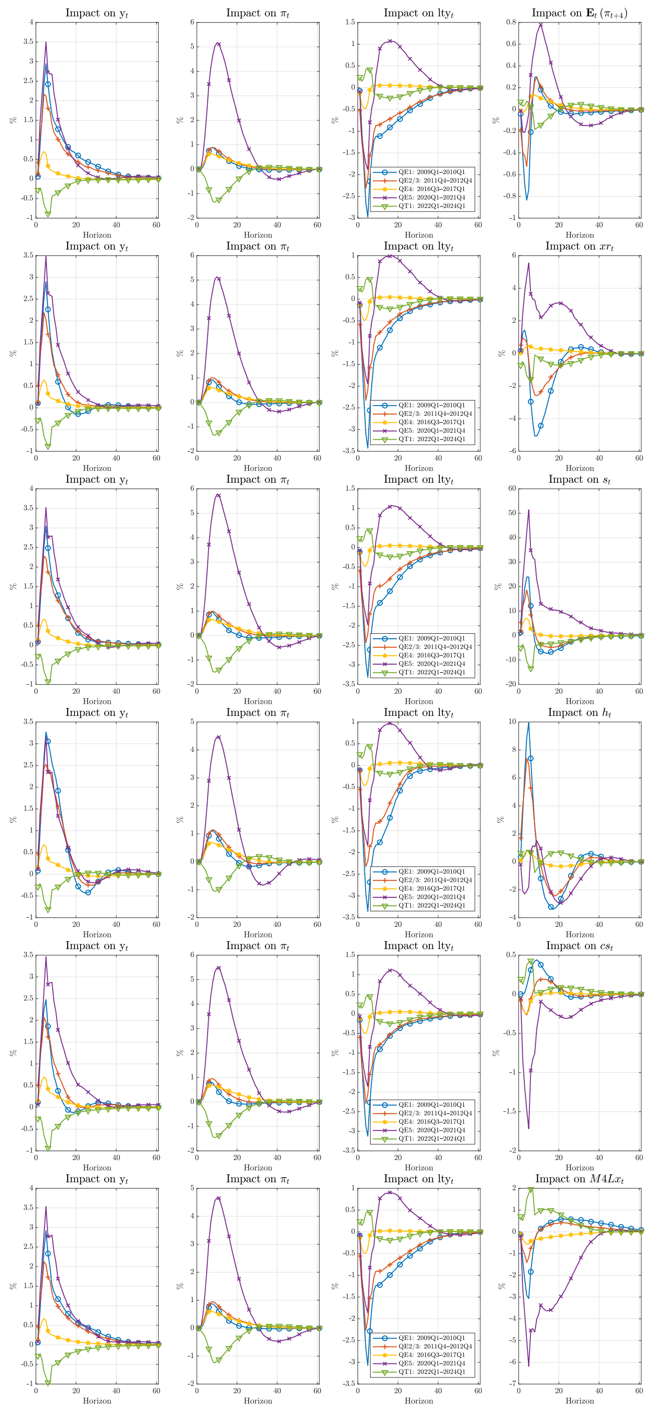

Figure 11. Counterfactuals with additional variables: Quantitative Easing/Tightening episodes
Notes: This figure plots the posterior median counterfactual paths of the output gap, y_t; CPI inflation, π_t; long-term yields, lty_t; and $z_t = \{\mathbf{E}_t(\pi_{t+4}), xr_t, s_t, h_t, cs_t, M4Lx_t\}$ across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. Rows 1–6 one add: one-year ahead inflation expectations to our baseline information set; UK effective exchange rate growth; annual stock returns; annual house price growth; corporate bond spreads; and bank & building society lending. They show the impact that each round of QE (QT) has in isolation of all other shocks through the lens of bank funding shocks. The counterfactual tells us the model-implied impact of each QE/QT episode and is directly comparable to the first row in Figure 7.

Table 4. Contemporaneous Sign Restrictions

Notes: This table reports the contemporaneous response of: the Bank Rate, r_t; public sector contributions to the banking sector as a proportion of broad money stock, psc_t; the long-term (i.e., 10-year gilt) gilt yield, lty_t; inflation, π_t; and output gap, y_t to structural shocks. $\mathbf{v}_t^{\mathbf{mp}}$ is a monetary policy shock; $\mathbf{v}_t^{\mathbf{q}}$ is a bank funding shock; $\mathbf{v}_t^{\mathbf{s}}$ is a supply shock; $\mathbf{v}_t^{\mathbf{d}}$ is a demand non-policy shock; $\mathbf{v}_t^{\mathbf{lty}}$ is a long-term yield shock; and $\mathbf{v}_t^{\mathbf{z}}$ is the sixth shock we identify . 0 denotes a zero response on impact, "x" denotes no restriction.

	$\mathbf{v}_t^{\mathbf{mp}}$	$\mathbf{v}_t^{\mathbf{q}}$	$\mathbf{v}_t^{\mathbf{s}}$	$\mathbf{v}_t^{\mathbf{d}}$	$\mathbf{v}_t^{\mathbf{lty}}$	$\mathbf{v}_t^{\mathbf{z}}$
r_t	$\geq$	0	x	$\geq$	0	0
psc_t	x	$\geq$	x	x	$\geq$	0
lty_t	x	$\leq$	x	$\geq$	$\geq$	0
π_t	$\leq$	0	$\leq$	$\geq$	$\leq$	0
y_t	$\leq$	$\geq$	$\geq$	$\geq$	$\leq$	0
z_t	x	x	x	x	x	$\geq$

response to QE episodes, perhaps because the use of QE signals a worsening in the state of the economy. But they increase further out and are most responsive during QE1, QE2/3 and QE5. The proportional response of inflation expectations relative to the response of actual inflation is larger in QE1 and QE2/3, with the increase being between 50-65% of the response of actual inflation; the proportional response in QE4, QE5 and QT1 is around 20%. The exchange rate depreciates in all episodes apart from QE5, with the largest falls in QE1 and QE2/3. That would be in line with portfolio re-balancing towards foreign currency assets. Neither result, however, suggests these variables are responsible for the increase in the reduced-form slope of the Phillips curve as we show from the corresponding slope estimates in Figure C.9 in Online Appendix C.4. These values are quantitatively similar to our main results. The appreciation of the exchange rate in QE5 might reflect the reduced signaling channel in QE5 and increased response of expected future Bank Rate. It might also reflect the effect of QE on confidence or foreign exchange rate risk premia.

Although we have used gilt yields to identify the impact of QE on asset prices, equity and house prices are likely to be more important in the transmission to demand as argued by Congdon (2025). Over QE episodes, equity and house prices generally increase in line with the portfolio balancing and signaling channels (Bridges and Thomas, 2012). The magnitudes in QE1 are similar to estimates in Joyce et al. (2011). Coupled with the appreciation of sterling in QE5, the relatively large increase in stock prices may suggest QE had a large effect on confidence during QE5 relative to others. This might explain part of the increase in responsiveness of demand to gilt yields we show in Figure 6, despite the decline in competitiveness the exchange rate suggests.

The response of corporate bond spreads is also different across episodes. Spreads widen slightly in QE1. However, corporate bond *yields* would have still fallen substantially given the response of the 10-year gilt rate; which is line with the UK event studies. In QE5, however, bond spreads decline significantly, suggesting that QE in this episode may have been working through lowering uncertainty and corporate credit risk. Coupled with the strong equity price response, it suggests that increased transmission though capital markets and NBFIs may partly explain the increased responsiveness of demand to gilt yields.

Looking finally at the responses of bank lending, we can see that QE shocks typically led to a decrease rather an increase in lending. Similarly the initial stages of QT provided a small boost to lending. This is in line with the idea that QE worked mainly through capital markets and causes some substitution away from bank credit given the fall in yields. The strong fall in lending growth in QE5, is also consistent with the substantial decline in corporate bond spreads noted earlier, which would have induced a higher degree of substitution away from bank debt compared to earlier episodes. The effects of QE on bank lending are ambiguous as discussed in Bridges and Thomas (2012) and Cloyne et al. (2015). The results here though are in line with the empirical evidence of Fatouh et al. (2025) based on an analasis of bond-level and corporate-level data over the QE episodes. We stress here that to arrive at any firm conclusion on such channels requires the proper identification of credit supply or banking sector shocks in the VAR. For instance, the capital market substitution channel is the result of corporate portfolio rebalancing away from existing debt. However, banks may respond to QE or QT by changing their supply of new loans, (see for example Kumhof and Salgado-Moreno (2024)) which may not be captured properly by our identification scheme.

We conclude here that our main results are robust to accounting for additional variables within the information set. They also point to some channels that might account for the increased response of demand, especially the role of capital market channels in QE5 However, a complete analysis requires full structural identification of these shocks and the imposition of more appropriate restrictions.

6 Summary of impacts across QE and QT episodes

To summarize the estimated impact of UMP policies from our analysis, Figure 12 provides an overview of our counterfactual QE and QT experiments over the period since 2009 for the output gap, inflation and the long-term gilt yield. We report the decomposition of the total QE impact into the unanticipated and anticipated/systematic components, the actual values of the data, and the corresponding counterfactual paths in the absence of QE/QT policies. For inflation, we show the contribution from the core inflation impact estimated in Section 5, together with the contributions from food, energy prices and other shocks to provide a comparison with other studies. Several main conclusions emerge regarding the quantitative impact of QE and QT, respectively.

First, QE policies provide substantial support to activity and avoided opening up large negative output gaps across each episode. The real economy outcomes would have been significantly worse without them.

Second, QE policies on average helped the BoE meet the inflation target of 2%, especially during QE5, when our results imply QE policies avoid deflation in late 2020 and early 2021. However, as we move through 2021, the deterioration in supply-side conditions, implied by the increases in our estimates of the slope of the Phillips curve, suggest that QE contributed around 2pps to the above-target inflation rate of 5% at the end of that year, with the remaining 3pps accounted for by food, energy and other shocks. Had the flexibility of the supply side of the economy been similar to that estimated for earlier QE episodes, inflation would have been closer to target. As noted earlier, the exact impact is quite uncertain and other variants that do not use core inflation in the VAR suggest the effect via domestic pricing factors could be larger, see Figure E.4, although these variants are less directly comparable with other studies that take explicit account of the role of international commodity prices. However, our results indicate that rising inflation following the Ukraine invasion is not directly attributable to QE5, with its contribution remaining flat throughout 2022. There would have been a substantial overshoot of inflation in 2022 and 2023 even if no QE had been undertaken during the pandemic, driven by the direct effects of food prices, energy prices and other fundamental shocks to the economy [26]. QT

[26] Note in the decomposition in Figure 12, the contribution of other shocks includes the indirect effects of food and energy prices working through the domestic supply chain on to prices in the core inflation basket.

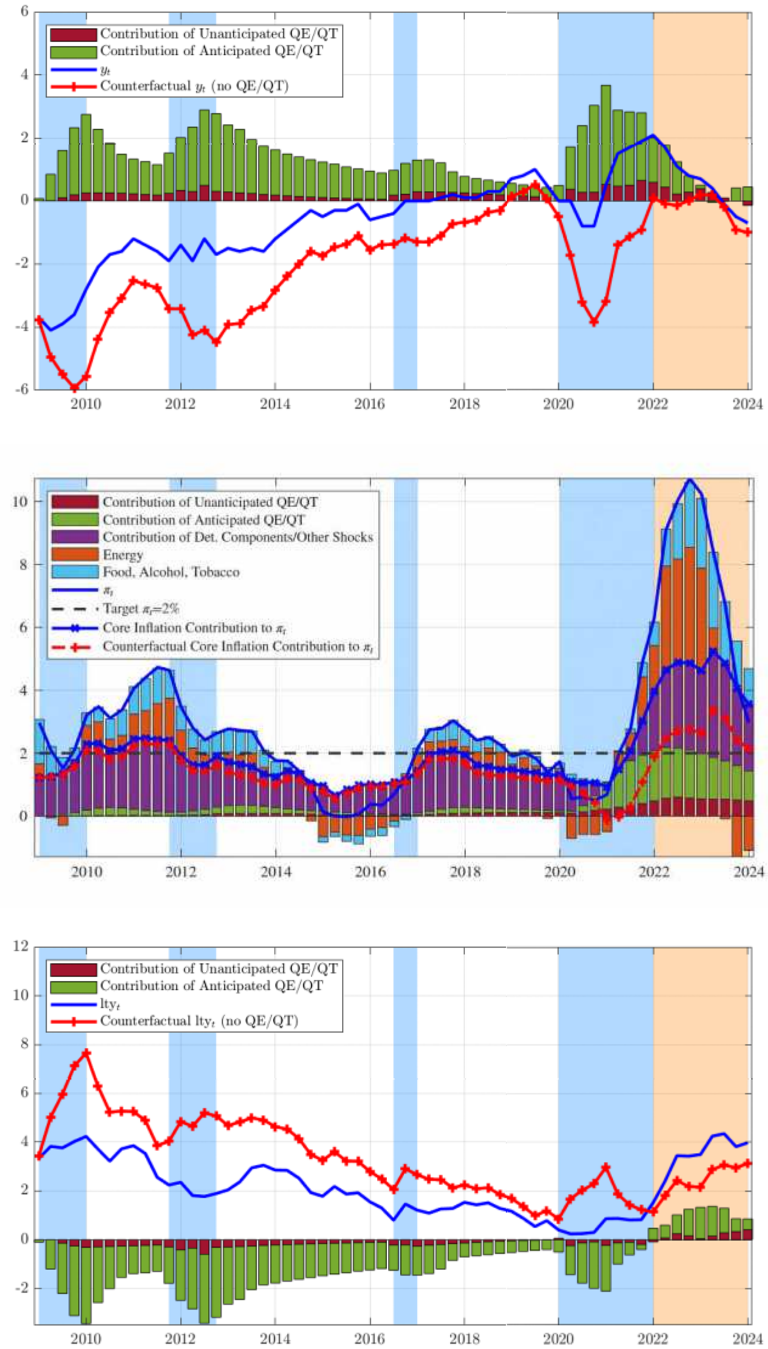

Figure 12. Counterfactual paths and contributions of components from 2009 to 2024
Notes: This figure plots the decomposition of QE/QT counterfactuals from 2009Q1–2024Q1 for the output gap, inflation, and long-term gilt yields. Bars decompose the counterfactual responses into unanticipated (burgundy bars) and anticipated components (green bars). We obtain the latter by taking the difference between the total counterfactual response minus the unanticipated component. Alongside these we plot the observable data (blue lines) and the counterfactual path (red lines) of the data in the absence of QE/QT. We obtain the latter by taking the (posterior median) counterfactual total response in all periods minus the actual data at each time t. Light blue bars refer to QE periods, Orange bars show QT periods.

The core inflation measure only excludes the direct contributions of food and energy to CPI inflation contributes a little to bringing inflation back towards target in the period up to 2024Q1.

Third, in the absence of a QE response, gilt yields would have been significantly higher than what we observe in the data and by substantially more than suggested by just the unanticipated effects in this model and those suggested by event studies. In turn, yields would have been lower during the QT period, although once we factor in commercial bank purchases of gilts, the impact is relatively small.

Fourth, from a monetarist perspective, the inflationary impact of a given boost to broad money from QE appears significantly less than proportionate. This might suggest that QE leads to offsetting effects on broad money through other credit counterparts on the balance sheet or that QE is not fully neutral in its effects, as discussed in Bridges and Thomas (2012) and Cloyne et al. (2015).

Overall, the state contingency we observe in QE/QT impacts from our analysis might recommend a gradualist approach to QE/QT policies, which is consistent with Sims and Wu (2021), who show a gradual unwinding is preferable over an immediate one, and the well-known conclusion of Brainard (1967) in the presence of parameter uncertainty. However, estimates from our VAR models show increases in the persistence of inflation over the most recent episodes of QE, and in the posterior uncertainty around the estimate; such features appear in Figure 5.[27] A general increase in inflation persistence might warrant more active monetary policy responses as in Söderström (2002) in circumstances where QE/QT is an active tool of monetary policy. Therefore, to explore the implications of state contingency for policy gradualism further, would require one to incorporate these parameter features into a broader analysis of optimal policy under uncertainty with both conventional and unconventional monetary policy instruments.[28]

7 Conclusions

This paper provides a comprehensive assessment on the state contingencies of QE and QT policies. Using UK data spanning 60 years, we construct a proxy measure of public

[27]These results are in Figure C.11.

[28]For completeness, Online Appendix E shows analogous plots to Figure 12 where we report the counterfactual path responses of the output gap, inflation and gilt yields for: alternative proxies of public sector borrowing from the banking system; alternative identification schemes; and the accounting for additional variables in the information set (i.e. all of those we consider in Section 5). These results are consistent with those from our baseline counterfactuals in Figure 12.

sector borrowing debt sales to the banking system to track the impact of unconventional monetary policies. Our measure is easy to construct for long time periods and has historical precedence as under- and over-funded deficits deliver almost identical effects on financial balance sheets as unconventional monetary policies. Using sign and zero restrictions consistent with economic rationale, we identify a novel bank funding shock to shed light on the channels that underpin such policies over time.

We show that the effects of unconventional monetary policies on the real economy are state contingent. The determinants of state contingencies depend on the interplay between policy regime, the state and structure of financial markets, and conditions in the real economy. Our appraisal of historical debt management operations and unconventional monetary policy episodes, using a range of specifications, suggests that history rhymes but does not repeat. Hence we provide substantial evidence that models with non-recurring regimes provide the flexibility to capture the stage contingent effects of QE/QT over time.

Our approach also enables us to distinguish between the unanticipated and total effects of UMP episodes on the real economy. We show that the unanticipated component contributes relatively small amounts to total movements in line with various event-study impacts. The total effect of QE/QT however appear to provides substantial support to the real economy, and aid the MPC in meeting the inflation target up until 2021. It suggests a reliance on event study impacts alone may understate the broader effects of QE.

Disentangling the transmission mechanism at work during each UMP episode reveals the impact on gilt yields for a given amount of QE/QT falls progressively after QE1. Our results suggest this is driven by a smaller signaling channel, and a negligible change in the portfolio rebalancing channel. Such a fall is offset by increases in the responsiveness of the output gap to gilt yields in later QE episodes, which may reflect an increased role for capital markets in the transmission of QE. We also detect a higher and more persistent responsive of inflation to the output gap and increased inflation. Deteriorating supply-side conditions mean that QE was a contributor to the overshoot of the inflation target in 2021. However, our results imply no contribution to the additional pick up in inflation in 2022 and 2023, with QT policies helping bring inflation towards the target at the end of the sample.

References

Alessandri, P. and H. Mumtaz (2017). Financial conditions and density forecasts for US output and inflation. *Review of Economic Dynamics 24*, 66–78.

Alessandri, P. and H. Mumtaz (2019). Financial regimes and uncertainty shocks. *Journal of Monetary Economics 101*, 31–46.

Allen, W. A. (2019). *The Bank of England and the Government Debt: Operations in the Gilt-Edged Market, 1928–1972*. Studies in Macroeconomic History. Cambridge University Press.

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2018). Inference based on Structural Vector Autoregressions Identified with Sign and Zero Restrictions: Theory and Applications. *Econometrica 86*(2), 685–720.

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2025). Uniform priors for impulse responses. *Econometrica 93*(2), 695–718.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics 131*(4), 1593–1636.

Bank of England (2021). Assessing the resilience of market-based finance. Technical report, Bank of England.

Bank of England (2024, August). Box A: Reviewing the process of quantitative tightening. *Monetary Policy Report*.

Bank of England (2025, May). Box D: Assessing recent developments in broad money. *Monetary Policy Report*.

Bank of England (2026, January). Forecast evaluation report – January 2026. Technical report, Bank of England.

Baumeister, C. and L. Benati (2013). Unconventional monetary policy and the great recession: Estimating the macroeconomic effects of a spread compression at the zero lower bound. *International Journal of Central Banking 9*(2), 165–212.

Baumeister, C. and J. D. Hamilton (2015). Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica 83*(5), 1963–1999.

Bernanke, B. and O. Blanchard (2025, July). What caused the us pandemic-era inflation? *American Economic Journal: Macroeconomics 17*(3), 1–35.

Bernanke, B. S. (2020). The new tools of monetary policy. *American Economic Review 110*(4), 943–983.

Bhattarai, S. and C. J. Neely (2022). An analysis of the literature on international unconventional monetary policy. *Journal of Economic Literature 60*(2), 527–597.

Bordo, M. D., O. Bush, and R. Thomas (2025). "Muddling Through or Tunnelling Through?" UK Monetary and Fiscal Exceptionalism and the Great Inflation. *Bank of England Working Paper No. 1135*.

Brainard, W. C. (1967). Uncertainty and the effectiveness of policy. *American Economic Review 57*(2), 411–425.

Braun, R., S. Miranda-Agrippino, and T. Saha (2025). Measuring monetary policy in the UK: The UK monetary policy event-study database. *Journal of Monetary Economics 149*, 103645.

Bridges, J. and R. Thomas (2012). The impact of QE on the UK economy–some supportive monetarist arithmetic. *Bank of England Working Paper No. 442*.

Busetto, F., M. Chavaz, M. Froemel, M. Joyce, I. Kaminska, and J. Worlidge (2022). QE at the Bank of England: a perspective on its functioning and effectiveness. *Bank of England Quarterly Bulletin*, Q1.

Canova, F. and M. Paustian (2011). Business cycle measurement with some theory. *Journal of Monetary Economics 58*(4), 345–361.

Chan, J. C. and E. Eisenstat (2018). Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *Journal of Applied Econometrics 33*(4), 509–532.

Christiano, L. J. and T. J. Fitzgerald (2003). The band pass filter. *International Economic Review 44*(2), 435–465.

Cloyne, J., R. Thomas, A. Tuckett, and S. Wills (2015). An empirical sectoral model of unconventional monetary policy: The impact of QE. *The Manchester School 83*, 51–82.

Cogley, T. and T. J. Sargent (2005). Drifts and volatilities: Monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics 8*(2), 262–302.

Congdon, T. (2025). *Money and Inflation at the Time of Covid*. Edward Elgar.

Costain, J., A. Nakov, and B. Petit (2022). Flattening of the Phillips curve with state-dependent prices and wages. *The Economic Journal 132*(642), 546–581.

Doan, T., R. Litterman, and C. Sims (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews 3*(1), 1–100.

Du, W., K. Forbes, and M. N. Luzzetti (2024). Quantitative tightening around the globe:

What have we learned? NBER Working Paper 32321, National Bureau of Economic Research, Cambridge, MA.

Ellington, M. (2022). The empirical relevance of the shadow rate and the zero lower bound. *Journal of Money, Credit and Banking 54*(6), 1605–1635.

Ellington, M., M. Michalski, and C. Milas (2023). Of votes and viruses: the uk economy and economic policy uncertainty. *European Journal of Finance 29*(16), 1849–1865.

Fatouh, M., S. Giansante, and M. Duygun (2025). The real economy effects of qe through the corporate bond market. *Bank of England Working Paper No. 1134*.

Florackis, C., G. Giorgioni, A. Kostakis, and C. Milas (2014). On stock market illiquidity and real-time GDP growth. *Journal of International Money and Finance 44*, 210–229.

Granger, C. (2008). Non-linear models: Where do we go next - time varying parameter models? *Studies in Nonlinear Dynamics & Econometrics 12*(3), 1–12.

Haldane, A., M. Roberts-Sklar, C. Young, and T. Wieladek (2016). QE: the story so far. *Bank of England Working Paper No. 624*.

Haskel, J., J. Martin, and L. Brandt (2025). What explains recent UK inflation? An application of the Bernanke-Blanchard model. Discussion Paper ESCoE DP 2025-12, Economic Statistics Centre of Excellence (ESCoE).

Hayashi, F. and J. Koeda (2019). Exiting from quantitative easing. *Quantitative Economics 10*(3), 1069–1107.

Hodrick, R. J. and E. C. Prescott (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit and Banking*, 1–16.

Hubrich, K. and R. J. Tetlow (2015). Financial stress and economic dynamics: The transmission of crises. *Journal of Monetary Economics 70*, 100–115.

Inoue, A., B. Rossi, and Y. Wang (2024). Has the Phillips Curve Flattened? *Centre for Economic Polcy Research (CEPR) Discussion Paper 18846*.

Jacquier, E., N. G. Polson, and P. E. Rossi (2002). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics 20*(1), 69–87.

Joyce, M., D. Miles, A. Scott, and D. Vayanos (2012). Quantitative easing and unconventional monetary policy–an introduction. *The Economic Journal 122*(564), F271–F288.

Joyce, M., M. Tong, and R. Woods (2011). The United Kingdom's quantitative easing policy: design, operation and impact. *Bank of England Quarterly Bulletin*, Q3.

Joyce, M. A. and M. Tong (2012). QE and the gilt market: a disaggregated analysis. *The Economic Journal 122*(564), F348–F384.

Kaminska, I. and H. Mumtaz (2022). Monetary policy transmission during QE times: role of expectations and term premia channels. *Bank of England Staff Working Paper 978*.

Kanngiesser, D. and T. Willems (2024). Forecast accuracy and efficiency at the Bank of England – and how errors can be leveraged to do better. *Bank of England Working Paper No. 1078*.

Kapetanios, G., H. Mumtaz, I. Stevens, and K. Theodoridis (2012). Assessing the economy-wide effects of quantitative easing. *The Economic Journal 122*(564), F316–F347.

King, M. A. (2024). Inflation targets: Practice ahead of theory. NBER Working Paper 32594, National Bureau of Economic Research, Cambridge, MA.

Koop, G. and D. Korobilis (2010). *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics*. Now Publishers Inc.

Koop, G., M. H. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*(1), 119–147.

Kumhof, M. and M. Salgado-Moreno (2024). Quantitative easing and quantitative tightening: the money channel. *Bank of England Working Paper No. 1090*.

Leeper, E. M. and T. Zha (2003, November). Modest policy interventions. *Journal of Monetary Economics 50*(8), 1673–1700.

Litterman, R. B. (1986). Forecasting with Bayesian Vector Autoregressions—Five Years of Experience. *Journal of Business & Economic Statistics 4*(1), 25–38.

Petrova, K. (2019). A quasi-Bayesian local likelihood approach to time varying parameter VAR models. *Journal of Econometrics 212*(1), 286–306.

Pettenuzzo, D. and A. Timmermann (2017). Forecasting macroeconomic variables under model instability. *Journal of Business & Economic Statistics 35*(2), 183–201.

Pill, H. (2022). What did the monetarists ever do for us? Speech at the Walter Eucken Institute.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies 72*(3), 821–852.

Ronicle, D. (2026). Quantitative tightening? Britain's 1980s experiment with overfunding. *Bank of England Working Paper No. 1183*.

Sims, E. and J. C. Wu (2021). Evaluating central banks' tool kit: Past, present, and future. *Journal of Monetary Economics 118*, 135–160.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(4), 583–639.

Söderström, U. (2002). Monetary policy with uncertain parameters. *The Scandinavian Journal of Economics 104*(1), 125–145.

Vayanos, D. and J.-L. Vila (2021). A preferred-habitat model of the term structure of interest rates. *Econometrica 89*(1), 77–112.

Weale, M. and T. Wieladek (2016). What are the macroeconomic effects of asset purchases? *Journal of Monetary Economics 79*, 81–93.

Wu, J. C. and F. D. Xia (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking 48*(2-3), 253–291.

Wu, J. C. and J. Zhang (2019a). Global effective lower bound and unconventional monetary policy. *Journal of International Economics 118*, 200–216.

Wu, J. C. and J. Zhang (2019b). A shadow rate new Keynesian model. *Journal of Economic Dynamics and Control 107*, 103728.

Online Appendix for

Are the Effects of Quantitative Easing & Tightening State Contingent?

Abstract

This appendix contains background historical details, prior and posterior simulation algorithm-related information and additional results not included in the main body of the paper.

A Historical debt management regimes and data

A.1 Historical debt management and the credit counterparts to broad money

Our analysis on the state contingency of UMPs covers earlier periods when the Bank of England carried out debt-management operations that implied very similar effects on financial balance sheets as current QE and QT operations. In particular, we look at earlier episodes when the commercial banking system purchased substantial amounts of government debt and also periods when more gilts were sold to non-bank investors than required to fund the deficit. These operations were known as “underfunding” and “overfunding” because they implied that the government was not financing its deficit with equivalent sales of debt to the long-term “preferred habitat” investors who normally bought government bonds (insurance companies and pension funds). When deficits were “fully funded” by sales to non-bank investors the impact on commercial bank reserves and the broad money supply (bank deposits) is neutral.[29] If however, the government found it could not raise sufficient finance from this sector it would, by residual, sell debt to the commercial banks. Relative to the fully-funded case, commercial banks end up holding more gilts at the expense of those that would have been held by the non-bank financial sector. So the marginal effect is equivalent to a commercial bank buying a government bond directly from a non-bank.

[29] In essence the government issues bonds to the non-bank private sector to finance its deficit which drains reserves and broad money. When the government then spends the money it raised by borrowing from the non-bank private sector, the equivalent amount of money floods back into the system and reserves and broad money are restored.

The data on over– and underfunding as well as QE and QT is conveniently summarized in the Bank of England's monetary statistics, which are available quarterly back to 1963. The link between funding from the banking system and broad money aggregates was a key part of the credit counterparts analysis of the money supply that developed from the late 1960s onwards, when central banks started to pay more attention to the monetary aggregates. The counterparts approach simply expresses the change in broad money - the consolidated banking system's monetary liabilities to the non-bank private sector - as an identity in terms of all the other items on the banking system's balance sheet. Over and underfunding of the deficit is captured in the "public sector contribution" to broad money growth reflecting the holdings of government debt of both the central bank and commercial banks.

The public sector contribution implicitly captures underfunding and overfunding operations as well as QE and QT. This can be expressed in both flow terms and as a contribution to broad money growth. Figure A.1 shows the data since 1963 based on contributions to M4 and, from 1997, the M4ex definition of broad money. The banking system is defined to include the building societies throughout and after 1997, excludes the deposit liabilities of intermediate other financial companies (IOFCs) - institutions recognized to be effectively part of the banking system. To complement this, we present in Table A.1 a historical summary of debt management and monetary policy regimes in the UK since 1963.

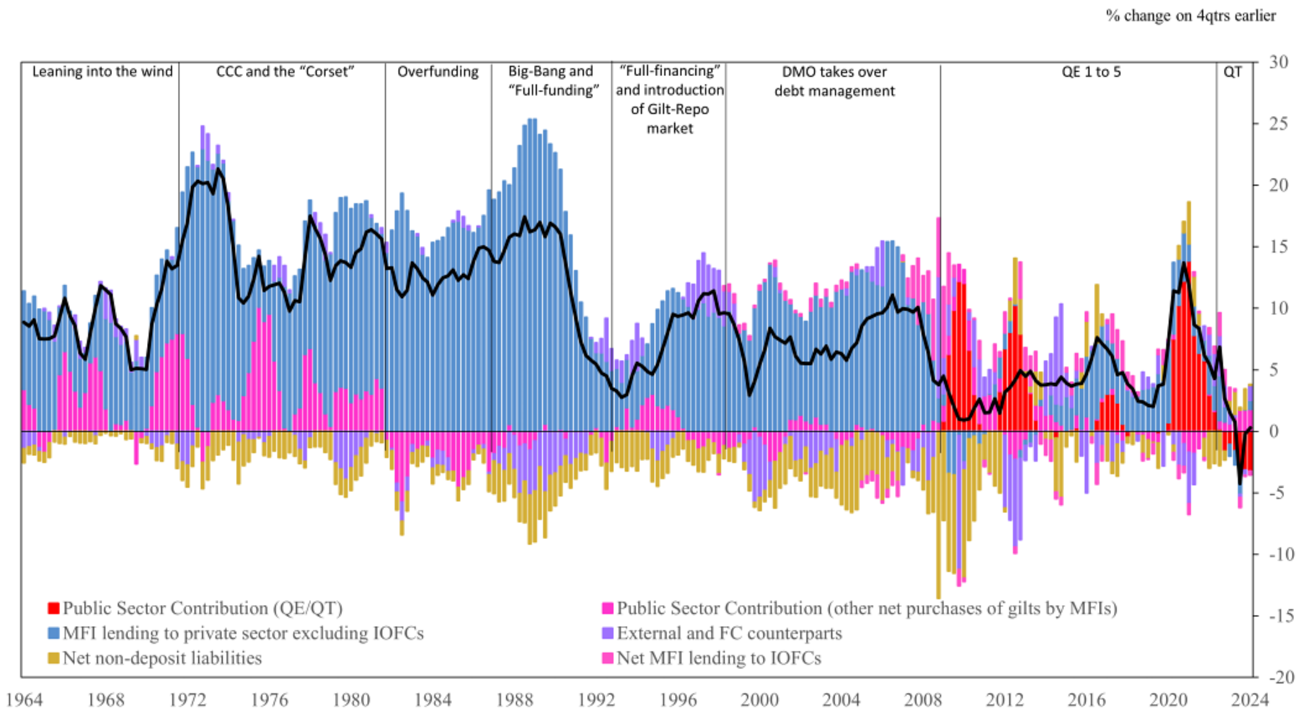

Figure A.1. UK Broad Money Decomposition from 1964 to 2024

Notes: This figure shows the credit counterparts to broad money growth since 1964. Periods of underfunding are represented by a positive public sector contribution and overfunding by a negative contribution. Different regimes of debt management and monetary policy are identified to provide some basis for the state contingency of bank funding operations. QE/QT operations from 2009 are shown in red, other purchases by the banking system (including purchases by the Bank of England prior to 2009) are shown in pink. The first episode spanning 1964Q2–1971Q3 we label "leaning against the wind", reflecting the general BoE policy to smooth long-term yield fluctuations through the purchase (sale) of government debt when gilt yields rise (fall). The 1971Q3–1981Q4 episode covers the period of Competition and Credit Control, gilt market turbulence and the beginning of monetary targeting. Throughout this period, the BoE allow for more scope for market determination of yields. As such, large, unexpected government deficits would be financed by default via the banking system if the Bank judged it too difficult for gilt sales. This occurred at various points throughout the 1970s as shown by psc_t which shows increases on a scale equivalent to the QE operations in Figure A.3. The next two episodes cover the period of explicit overfunding (1982Q1–1986Q4) and the period after the "Big Bang" reforms to the stock market when full funding was resumed (1987Q1–1992Q4). Our final two episodes span 1993Q1–1997Q2 and 1997Q3–2008Q4. The former relates to the period where deficits would be fully funded over the cycle, which occurred alongside the development of the gilt-repo market in 1996. The latter relates to the period from which the Bank of England had operational independence for monetary policy and the Debt Management Office (DMO) took over debt management. The final episodes from 2009Q1–2024Q1 cover QE and QT periods.

Table A.1. Historical Debt Management Regimes from 1963–2024

Historical Debt Management Regimes	Details
1963-1971Q3 "Leaning against the wind"	After a period of overt attempts to target long-term yields just after World War 2 ("Ultra cheap money"), the Bank's debt management operations became less ambitious and attempted to smooth fluctuations in long-term yields through intervening by purchases of debt when gilt yields were rising and sales when yields were falling. This was to protect the under-capitalised stock market "jobbers" who acted as market makers and were small partnerships vulnerable to large movements in yields Allen (2019). In this sense the Bank acted as a "jobber of last resort" since the Bank would stand ready to buy and sell government stock at given prices. That meant the Bank was directly purchasing government securities in response to macroeconomic shocks that pushed up yields or shocks to sentiment occurring in the gilt market, so there was an activate reaction function. The more general aim was to maximise the capacity for financial markets to take up government debt and so it was less attentive to how exactly deficits were funded on average. The banking system acted as the residual form of finance for the public sector. There were periods when the commercial banking system took up large amounts of gilts, such as the mid-late 1960s around the time of the devaluation. This was in an environment when the authorities were also trying to impose direct controls on bank lending to the private sector.
Competition and Credit Control, gilt strikes and beginnings of monetary targeting 1971Q3-1981Q4	In 1971 the government and the Bank's attitude to monetary control changed. Direct controls were removed and the banking system cartel was ended and competition encouraged, partly in response to the development of parallel money markets such as the Eurodollar market. But there was also growing awareness of the need for greater monetary control especially after the floatation of sterling in 1972. The Bank would intervene less in the gilt markets to smooth out fluctuations in interest rates and allow more market determination of yields. That of course put pressure on the jobbers as market makers and made the gilt market more sensitive to economic developments given the change in the Bank's role. In practice, that meant gilt sales to non-bank investors increasingly happened in opportunistic bursts when the Bank felt it could sell a lot of gilts. This meant that large, unexpected government deficits would be financed by default via the banking system during periods when the Bank judged it too difficult or expensive for gilt sales which occurred at various points throughout the 1970s, as shown by positive public sector contributions on a scale equivalent to the post-2009 QE operations.
"Big Bang" and full funding 1987Q1-1992Q4	In October 1986 the stock exchange underwent major changes with the "Big Bang" set of reforms. Out went the single capacity system of separate brokers and under-capitalised jobbers to a dual capacity system of gilt-edged market making firms (the GEMMs) that performed both roles. This invited international firms into the market and many took over the older broker and jobber partnerships. Prior to Big Bang, the capital of the gilt-edged jobbers is believed to have amounted to around £100 million. At the time of Big Bang the capital of the twenty-seven new GEMMs amounted to £595 million - a six-fold increase which greatly increased the liquidity of the market and its ability to absorb shocks. At the same time as Big Bang, disillusion with the overfunding regime led to a switch back to a "full fund" policy where the deficit would be met in full from the non-bank private sector, but that here should be no over or underfunding over the course of a financial year. The idea was to insulate monetary policy from debt management. That is reflected in little public sector contribution over this period.
Full-financing and development of gilt-repo market 1993Q1-1997Q2	Inflation targeting was introduced at the end of 1992 and 1993 saw a review of funding policy where sales of debt to banks were allowed to form part of funding. This relaxation in part reflected the early 1990s recession and the large budget deficit that emerged. Given weak lending the commercial banks' cyclical demand for gilts was high, so funding policy accommodated cyclical changes in banks' holdings. But there was an intention there should be no trend finance from the banking system in order to insulate monetary conditions from debt management and so periods of underfunding needed to be followed by overfunding. This can be observed in A.1 with periods of fluctuating over and underfunding but no trend. In 1996 the gilt repo market was launched in the UK and the Bank began repo operations directly with the market in 1997. This further increased the liquidity of the gilt market and enhanced the role of gilts as collateral.
MPC independence and Debt Management Office takes over debt management 1997Q3 – 2008Q4	The Bank gained operational independence for monetary policy in 1997 and it was decided to make a clean split between monetary and debt management policy by creating an independent debt management office (DMO). The remit of the DMO was to minimise the government funding costs given the monetary policy set by the independent MPC. Funding policy carried on as before, with periods of alternating over and underfunding, albeit on a small scale.
QE to 1-5 and QT, 2009Q1 onwards	In March 2009 the Bank of England embarked on Quantitative Easing as a means to further ease monetary policy with Bank Rate at its effective lower bound (ELB). There were five stages of QE as shown in Figure 4. But the second and third stages were relatively close together so these are often bundled together as QE2/3. At the beginning of 2022 the Bank embarked on Quantitative Tightening with the MPC decision to stop reinvesting the proceeds of maturing gilts, followed by active sales in September 2022. These were punctuated by some temporary purchases of gilts for financial stability purposes during the LDI crisis.

A.2 Scaling issues in psc_t

The choice of appropriate scaling variable for psc_t makes little difference in the most recent period, since nominal GDP (NGDP), broad money (M4ex) and the stock of government debt are of similar value. But, as we go back in time, Figure A.2 shows that the trends in velocity and the debt-to-income ratio imply that bank funding shocks as a share of nominal GDP will be smaller and the share of government debt stock larger. Given the source of the data is consistently constructed from the Bank of England's monetary statistics and because it lies in between NGDP and the stock of debt, we use money growth as the appropriate scaling variable in our baseline case.

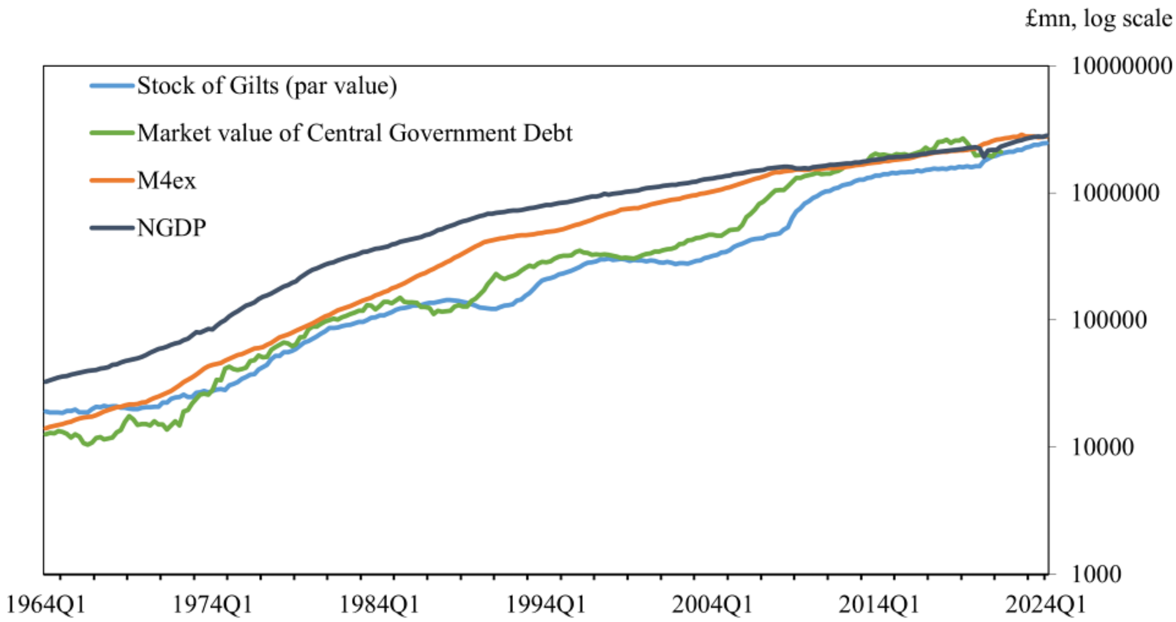

Figure A.2. UK candidate scaling variables from 1964 to 2024
Notes: This figure shows candidate scaling variables for our measure of psc_t from 1964Q1–2024Q1. We plot on a log scale, the Par value of total stock of gilts; the market value of central government debt securities; the broad money supply (M4ex), and nominal GDP.

A.3 Economic Data

Figure A.3 plots our economic data. The top left panel shows the output gap and annual CPI inflation, y_t, π_t. The top right panel shows the Bank Rate and the gilt yield, r_t, lty_t. The bottom panel plots public sector borrowing from the banking sector scaled by M4/M4ex, psc_t.

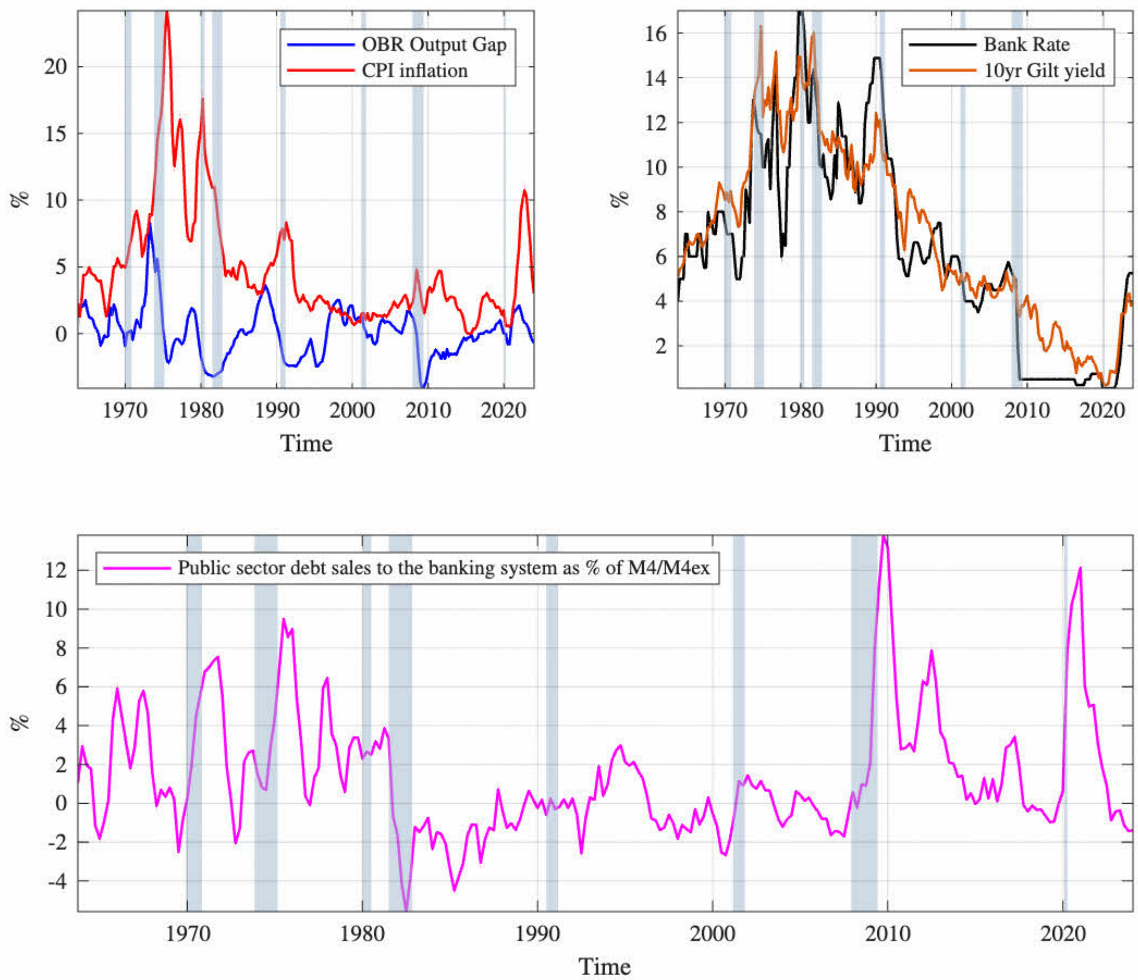

Figure A.3. UK Economic Data 1963Q3–2024Q1
Notes: This figure plots the output gap, y_t, and annual CPI inflation, π_t in the top left-hand side panel; and the Bank Rate, r_t, and the 10-year gilt yield, lty_t, in the top right-hand side panel. The bottom panel presents the public sector contribution to M4ex growth, psc_t. Grey bars indicate recessions.

B The Quasi-Bayesian local likelihood approach (Petrova, 2019): Priors and posterior simulation

We re-write Equation (1) by stacking the time-varying intercepts and autoregressive matrices in the vector ϕ_t with $Y_t = (y_1, \ldots, y_T)$, $X_t = \left(\mathbf{I}_N \otimes \left(1, x'_{t-1}, \ldots, x'_{t-p}\right)\right)$, with $\otimes$ denoting the Kronecker product:

$$Y_t = X'_t \phi_t + \mathbf{\Omega}_t^{-\frac{1}{2}} \boldsymbol{\eta}_{t,T} \tag{B.1}$$

where $\boldsymbol{\eta}_{t,T} \sim NID(0, \boldsymbol{I}_M)$. Note that all roots of the locally stationary VAR polynomial, lie outside the unit circle, and $\mathbf{\Omega}_t^{-1}$ is a positive definite time-varying covariance matrix. The local likelihood function at discrete time period s, is given by:

$$\mathrm{L}_s\left(Y_s | \phi_s, \mathbf{\Omega}_s, \bar{X}_s\right) \propto$$
$$|\mathbf{\Omega}_s|^{\mathrm{trace}(\mathbf{D}_s)/2} \exp\left\{-\frac{1}{2}(Y_s - X'_s\phi_s)' \left(\mathbf{\Omega}_s \otimes \mathbf{D}_s\right) \left(Y_s - X'_s\phi_s\right)\right\}, \tag{B.2}$$

where $\propto$ denotes "proportional to". $\mathbf{D}_s$ is a diagonal matrix whose elements hold the weights:

$$\mathbf{D}_s = \mathrm{diag}(\varrho_{s1}, \ldots, \varrho_{sT}) \tag{B.3}$$
$$\varrho_{st} = \zeta_{T,s} w_{st} / \sum_{t=1}^{T} w_{st} \tag{B.4}$$
$$w_{st} = (1/\sqrt{2\pi}) \exp((-1/2)((k-t)/W)^2), \quad \text{for } s, t \in \{1, \ldots, T\} \tag{B.5}$$
$$\zeta_{Ts} = \left(\left(\sum_{t=1}^{T} w_{st}\right)^2\right)^{-1} \tag{B.6}$$

where ϱ_{st} is a normalized kernel function. w_{st} uses a Normal kernel weighting function. ζ_{Ts} gives the rate of convergence and behaves like the bandwidth parameter W in (B.5). The kernel function assigns greater weight to observations near the parameter estimates at time s compared to those further away. Following Petrova (2019), we set $W = \sqrt{T}$.[30] We take 50,000 draws from the posterior distribution and retain every 50$^{\text{th}}$ draw; Online Appendix B provides details for the prior and posterior simulation algorithm. It is note-

[30] We also experiment with an Epanechnikov kernel; results are qualitatively similar to those we report below.

worthy to mention here that because the priors and posterior are conjugate, there is no need to use MCMC. As such, one can efficiently draw coefficient and covariance matrices from the (quasi) posterior distribution and leverage parallel computing. In addition to computational efficiency benefits, Petrova (2019) notes in simulation analysis that misspecification of the state-space invalidates inference asymptotically.[31] Online Appendix C provides additional results to our main analysis in Section 4 below.

Using a Normal-Wishart prior distribution for $\phi_s|\, \mathbf{\Omega}_s$ for $s \in \{1, \ldots, T\}$:

$$\phi_s|\mathbf{\Omega}_s \backsim \mathcal{N}\left(\phi_{0s}, (\mathbf{\Omega}_s \otimes \Xi_{0s})^{-1}\right) \tag{B.7}$$

$$\mathbf{\Omega}_s \backsim \mathcal{W}\left(\alpha_{0s}, \mathbf{\Gamma}_{0s}\right) \tag{B.8}$$

where ϕ_{0s} is a vector of prior means, Ξ_{0s} is a positive definite matrix, α_{0s} is a scale parameter of the Wishart distribution ($\mathcal{W}$), and $\mathbf{\Gamma}_{0s}$ is a positive definite matrix.

The prior and weighted likelihood function implies a Normal-Wishart quasi posterior distribution for $\phi_s|\, \mathbf{\Omega}_s$ for $s = \{1, \ldots, T\}$. Formally let $\mathbf{A} = (x_1', \ldots, x_T')'$ and $\mathbf{Y} = (y_1, \ldots, y_T)'$ then:

$$\begin{aligned} \phi_s|\mathbf{\Omega}_s, \mathbf{A}, \mathbf{Y} &\backsim \mathcal{N}\left(\tilde{\phi}_s, (\mathbf{\Omega}_s \otimes \tilde{\Xi}_{0s})^{-1}\right) && \text{(B.9)} \\ \mathbf{\Omega}_s &\backsim \mathcal{W}\left(\tilde{\alpha}_s, \tilde{\mathbf{\Gamma}}_s^{-1}\right) && \text{(B.10)} \end{aligned}$$

with quasi-posterior parameters:

$$\begin{aligned} \tilde{\phi}_s &= \left(\mathbf{I}_N \otimes \tilde{\Xi}_s^{-1}\right)\left[\left(\mathbf{I}_N \otimes \mathbf{A}'\mathbf{D}_s\mathbf{A}\right)\hat{\phi}_s + \left(\mathbf{I}_N \otimes \Xi_{0s}\right)\phi_{0s}\right] && \text{(B.11)} \\ \tilde{\Xi}_s &= \tilde{\Xi}_{0s} + \mathbf{A}'\mathbf{D}_s\mathbf{A} && \text{(B.12)} \\ \tilde{\alpha}_s &= \alpha_{0s} + \sum_{t=1}^{T} \varrho_{st} && \text{(B.13)} \\ \tilde{\mathbf{\Gamma}}_s &= \mathbf{\Gamma}_{0s} + \mathbf{Y}'\mathbf{D}_s\mathbf{Y} + \mathbf{\Phi}_{0s}\mathbf{\Gamma}_{0s}\mathbf{\Phi}_{0s}' - \tilde{\mathbf{\Phi}}_s\tilde{\Xi}_s\tilde{\mathbf{\Phi}}_s' && \text{(B.14)} \end{aligned}$$

where $\hat{\phi}_s = \left(\mathbf{I}_N \otimes \mathbf{A}'\mathbf{D}_s\mathbf{A}\right)^{-1}\left(\mathbf{I}_N \otimes \mathbf{A}'\mathbf{D}_s\right)y$ is the local likelihood estimator for ϕ_s. The matrices $\mathbf{\Phi}_{0s}$, $\tilde{\mathbf{\Phi}}_s$ are conformable matrices from the vector of prior means, ϕ_{0s}, and a draw from the quasi posterior distribution, $\tilde{\phi}_s$, respectively.

[31] In Online Appendix G we provide additional results using an array of alternative models including standard Bayesian VAR models, TVP-VARs of Primiceri (2005), and threshold VARs as in Alessandri and Mumtaz (2017).

C Additional Results from our baseline specification

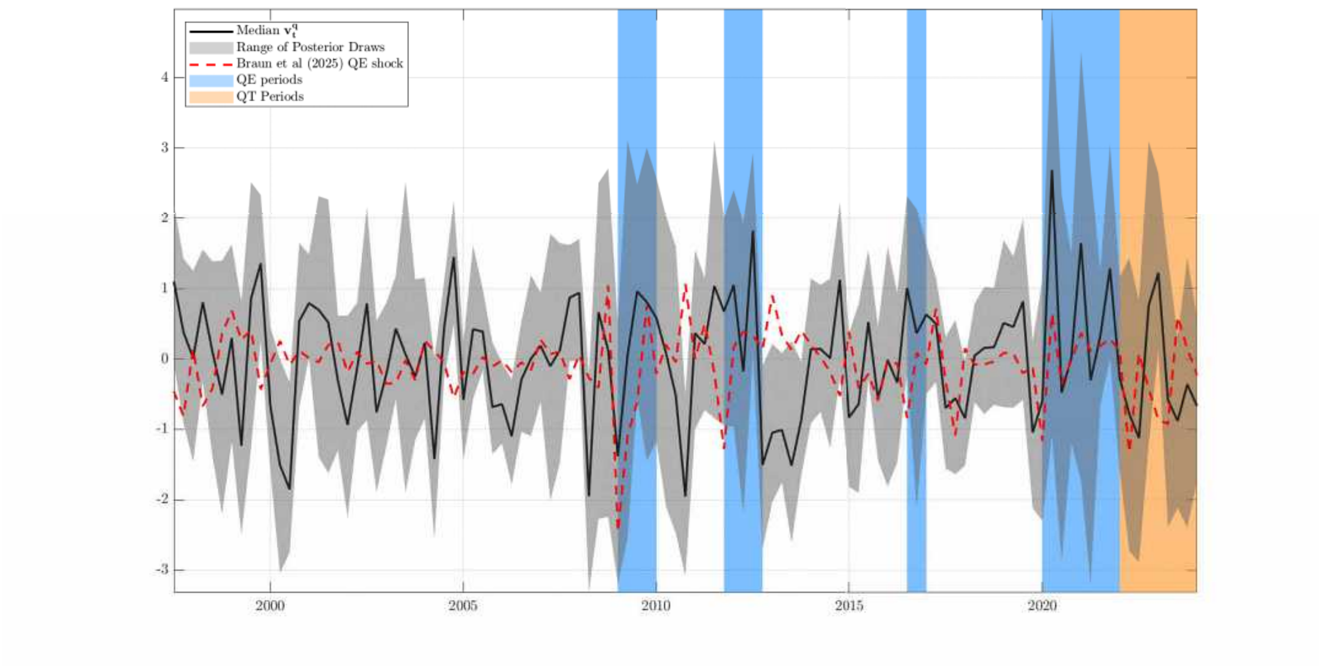

Figure C.1. Bank funding shocks, $\mathbf{v_t^q}$, and Braun et al. (2025) QE shocks from 1997Q3–2024Q1
Notes: the figure plots the posterior median (black line) and range of posterior draws (dark grey area) for structural Bank funding shocks, $\mathbf{v_t^q}$, along with the QE shocks (re-scaled for readability) in Braun et al. (2025) summed over quarters (red line). Blue vertical bars are Quantitative Easing periods, and orange bars label Quantitative tightening periods.

C.1 Peak impulse response analysis over different funding episodes

To clarify how the structure of financial markets and policy regime influence the degree of state contingency, Figure C.2 shows the peak median responses of the output gap and inflation, and average 1-year response of gilt yields over the sample. The inflation response to bank funding shocks is larger during periods of high inflation. The responsiveness of the output gap appears stronger in the 1970s, and again in the mid-1990s relative to the post-2009 period. The financial market response to bank funding shocks also differs depending upon regime. The sensitivity appears to increase throughout the 1960s, 1970s and 1980s, perhaps reflecting the large government deficits of this period and the low capitalisation and liquidity of the gilt market prior to the "Big Bang" and introduction of the gilt repo market. After these reforms, the sensitivity declines sharply up until 2009 and then falls more gradually across the QE episodes.

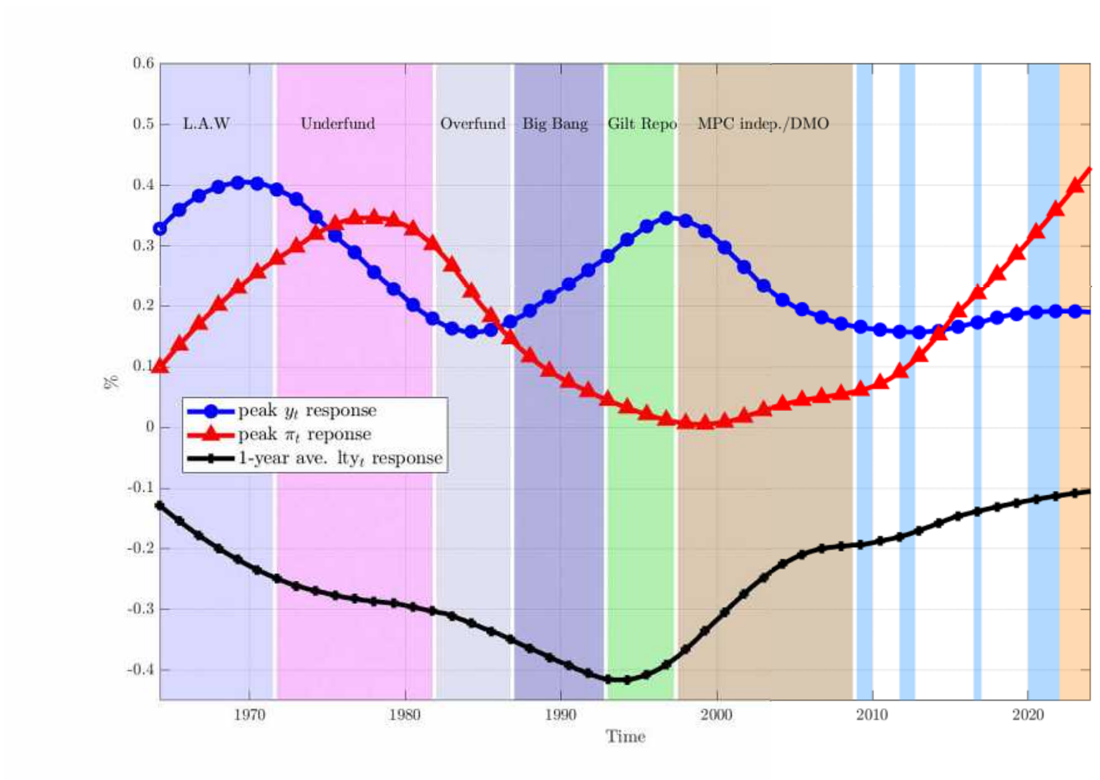

Figure C.2. Impulse response functions to bank funding shocks: Peak responses from 1964 to 2024

Notes: This figure plots the posterior median peak impulse response function of the output gap, y_t; inflation, π_t; and the long-term gilt yield, lty_t with respect to a bank funding shock, $\mathbf{v^q}_t$ from 1964Q2–2024Q1. In all cases, the peak response of the long-term gilt yield is the impact response. We normalize the distribution of impulse response functions such that a one-standard deviation bank funding shock causes psc_t to rise by 1%. We smooth peak responses for readability. Translucent bars indicate historical funding episodes and QE/QT periods. L.A.W is "Leaning Against the Wind" episode from 1963Q4–1971Q3. "Underfund", reflects the period of Competition and Credit control, the mid-1970s gilt strikes and the beginning of monetary targeting from 1971Q3–1981Q4. Overfund is the overfunding episode from 1982Q1–1986Q4, Big Bang reflects the introduction of the "Big Bang" reforms and the return to "full funding" over the period 1987Q1–1992Q4. Gilt Repo covers the more relaxed "full financing" rule and the development of the gilt repo market from 1993Q1–1997Q2. MPC indep/DMO is when the Monetary Policy Committee obtained independence and the Debt Management Office takes over debt management from 1997Q3–2008Q4. Translucent bars from 2009 – 2024 indicate the various rounds of QE/QT in the UK.

C.2 Historical decompositions

Here we report historical decompositions of our economic variables. For inflation, the Bank Rate, and long-term yields, we decompose the deterministic component of the historical decomposition into pertinent economic quantities using the VAR representation of long-run coefficients, $\lim_{h\to\infty} \mathbf{E}_t\mathbf{Y}_{t+h} = \left(\mathbf{I_N} - \beta(L)_t)^{-1}\right)\beta_{0,t}$, where $\mathbf{I_N}$, $\beta(L)_t$, $\beta_{0,t}$ are an N-dimensional identity matrix, the time t autoregressive parameters that we sum over lags

(L is the lag operator), and the time t intercepts respectively. Now let e_i be a selection vector for the relevant equation in the VAR model with $i = \{\pi, \mathrm{r}, \mathrm{lty}\}$ so that our pertinent long-run coefficients are:

$$\begin{aligned}
\pi_t^* &= \mathrm{e}_\pi \left(\mathbf{I_N} - \beta(L)_t)\right)^{-1} \beta_{0,t} \\
\mathrm{r}_t^* &= \mathrm{e}_\mathrm{r} \left(\mathbf{I_N} - \beta(L)_t)\right)^{-1} \beta_{0,t} \\
\mathrm{lty}_t^* &= \mathrm{e}_\mathrm{lty} \left(\mathbf{I_N} - \beta(L)_t)\right)^{-1} \beta_{0,t}
\end{aligned}$$

where π_t^* is core inflation, r_t^* is the natural rate of interest, and lty_t^* is the natural long-term yield. Using the Fisher relation we can express the real natural rate of interest as $(\mathrm{r}_t^* - \pi_t^*)$. We can further decompose the natural long-term yield into the real natural rate of interest, trend inflation, and the equilibrium term premium which is the difference between the natural long-term yield and the short-term natural rate of interest. Formally, $\mathrm{lty}_t^* = (\mathrm{r}_t^* - \pi_t^*) + \pi_t^* + \mathrm{tp}_t^*$, with $\mathrm{tp}_t^* = \mathrm{lty}_t^* - \mathrm{r}_t^*$. These measures allow us to understand further the assumptions our identification scheme makes on the equilibrium quantities.

Figure C.3 reports the historical decompositions of the output gap and inflation into the contribution of the different shocks. Supply shocks dominate all other shocks (i.e., monetary policy, bank funding demand or long-term yield shocks) in affecting output. Bank funding shocks contribute positively to output during the beginning of QE in 2009 and more so during the pandemic period. Monetary policy shocks have a small positive impact, between 2016 and 2021 as the Bank Rate moves away from the zero lower bound, and then weigh negatively on output following interest rate rises from late 2021. Supply shocks also dominate other shocks in affecting UK inflation. Bank funding shocks impact positively on inflation most notably since the pandemic, which is also the case for monetary policy and demand shocks. Finally, note that trend inflation accounts for a large quantity of the overall value of inflation and tends towards the BoE's target value of 2% as inflation targeting becomes the MPC's mandate.

Figure C.4 reports the historical decomposition of Bank Rate and the long-term yield into the contribution of the different shocks. Aggregate supply shocks contribute negatively to Bank Rate for most of the period since 2009; the impact of supply shocks turns positive after the pandemic. The same observation carries with reference to supply shocks for the long-term yield. For most of the post-2009 period, bank funding shocks and demand shocks also contribute negatively to the long-term yield. It is clear that the real natural rate of interest and trend inflation are the main drivers of the deterministic component of the Bank Rate. Looking at the trend components of the long-term yield, trend inflation is a strong contributor to the long-term yield. The real natural rate of interest fluctuates from contributing negatively (i.e. 1964 – 1976, 2003–2021) and positively (1977–2003, 2022–2024), with the corresponding periods equilibrium term premium having the opposing effects, to long-term yields.

For completeness, we report the historical decomposition of psc_t in Figure C.5. The only noteworthy feature to mention here is that during QE/QT episodes the contribution of bank funding shocks contributes substantially positively (negatively) to the overall value of the series.

Overall, throughout the QE (QT) episodes, bank funding shocks and monetary policy shocks always contribute to the economic variables in the same manner e.g. positive for inflation/output gap during QE periods; negative for long-term yields. This is consistent with existing theoretical and empirical results showing that at the ZLB unconventional expansionary monetary policy deliver consistently the same qualitative economic impact (in terms of sign) as conventional expansionary monetary policy (see e.g. Wu and Zhang, 2019a; Wu and Xia, 2016; Wu and Zhang, 2019b; Ellington, 2022).

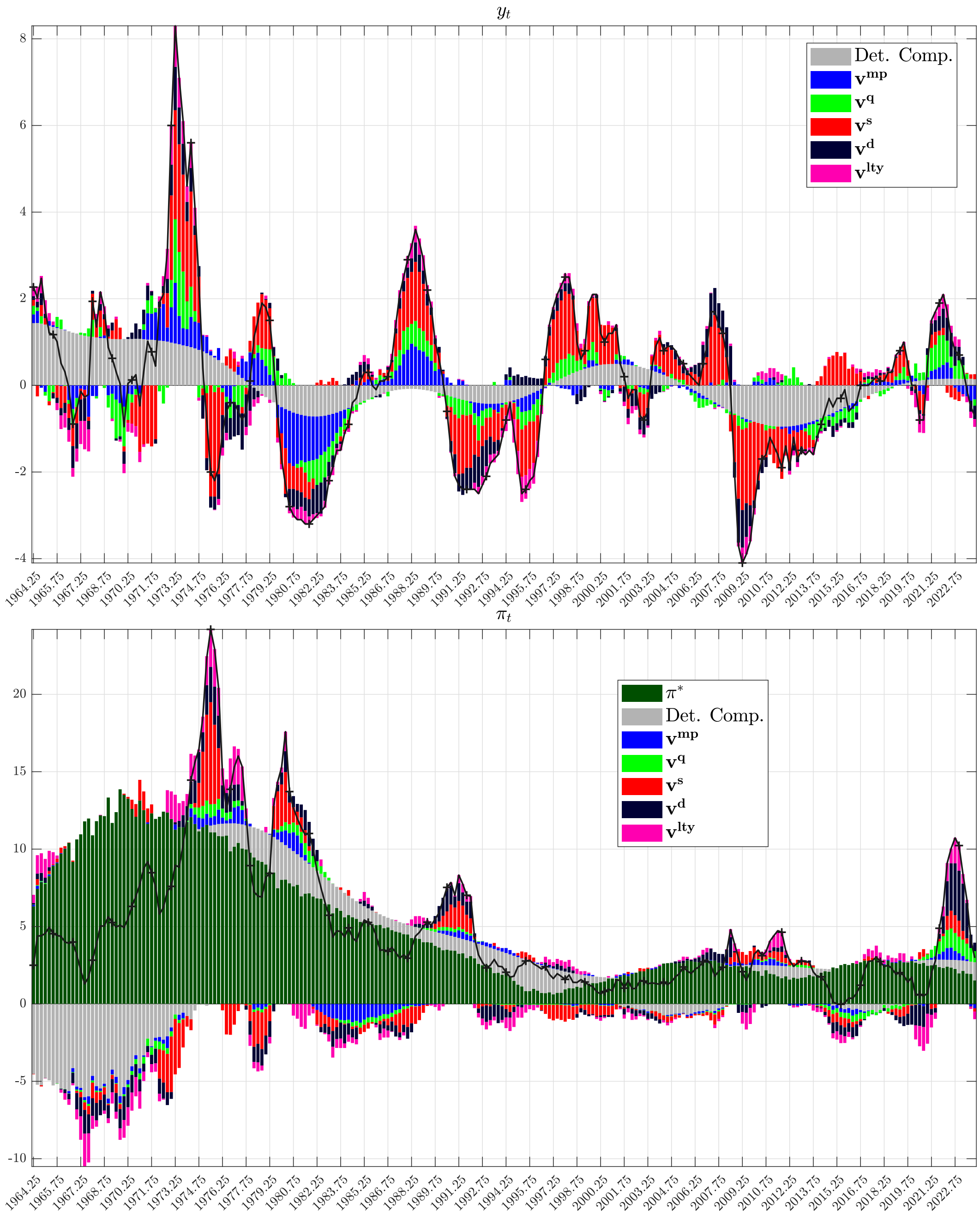

Figure C.3. Historical decomposition of the output gap and inflation from 1964Q2–2024Q1
Notes: This top and bottom panels of this figure plot the historical decomposition of the OBR output gap, y_t, and annual CPI inflation, π_t respectively from the TVP-VAR. Grey bars are the initial condition; Blue bars are monetary policy shocks; Light Green bars are bank funding shocks; Red bars are supply shocks, black bars are demand shocks; and magenta bars are long-term yield shocks. The black line is the actual time-series data we retrieve from the structural model. For CPI inflation we also report the contribution of trend inflation π^* as dark green bars which stem from the corresponding long-run coefficient VAR relationship $\lim_{h\to\infty} \mathbf{E}_t \mathbf{Y}_{t+h} = \left(\mathbf{I_N} - \beta(L)_t)^{-1}\right) \beta_{0,t}$.

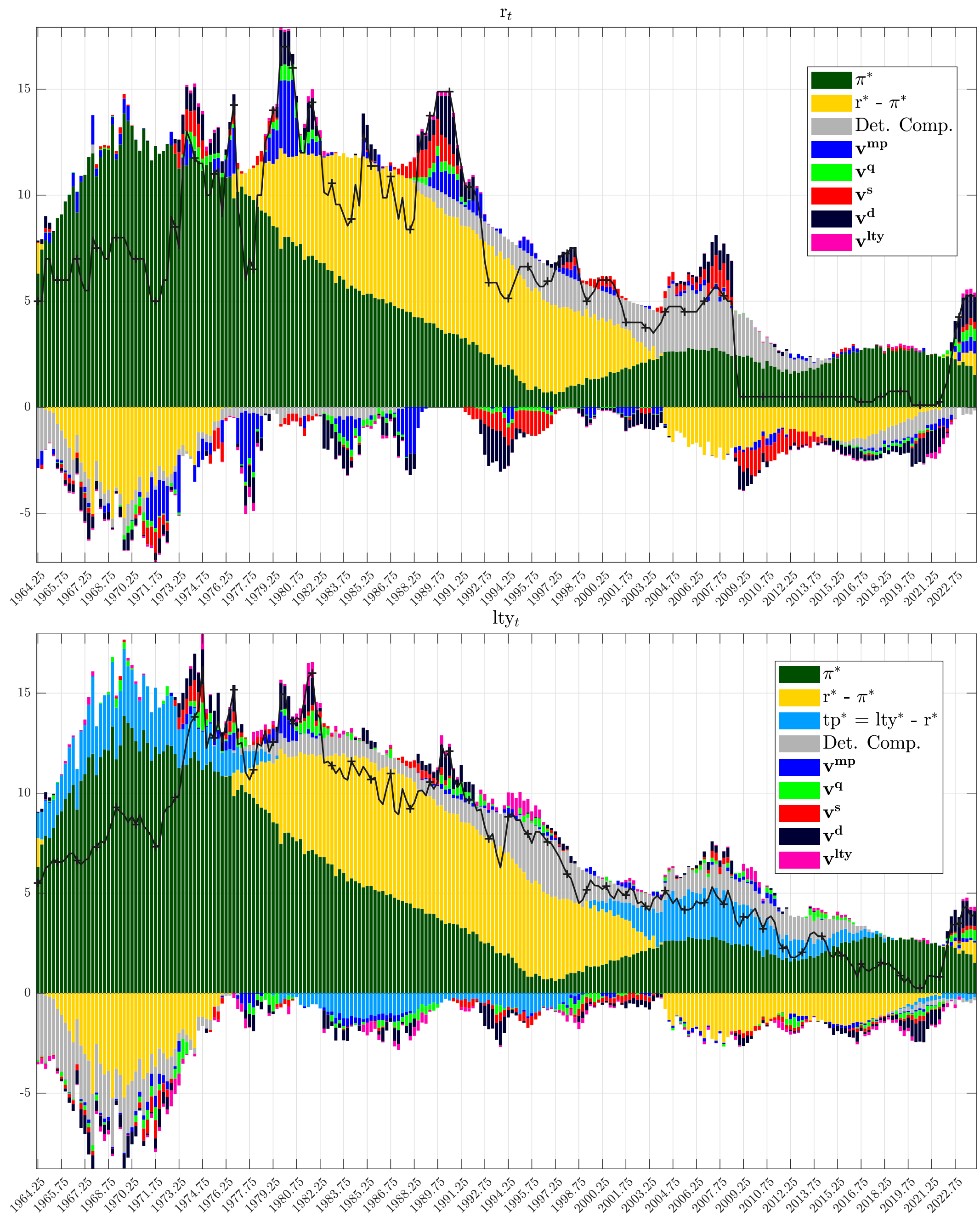

Figure C.4. Historical decomposition of the Bank Rate and long-term yields from 1964Q2–2024Q1

Notes: This top and bottom panels of this figure plot the historical decomposition of the Bank Rate, r_t, and long-term yields, lty_t respectively from the TVP-VAR. Grey bars are the initial condition; Blue bars are monetary policy shocks; Green bars are bank funding shocks; Red bars are supply shocks, black bars are demand shocks; and magenta bars are long-term yield shocks. The black line is the actual time-series data we retrieve from the structural model. For the Bank Rate we report the contributions of the natural rate of interest r^* which are violet bars. For the long-term yield we report the contributions of trend inflation, π^* (dark green bars); the real natural rate of interest, $(r_t^* - \pi_t^*)$ (yellow bars), and the equilibrium term premium, $tp_t^* = lty_t^* - r_t^*$ (light blue bars). All of which are manipulations from the long-run coefficient VAR relationship $\lim_{h\to\infty} \mathbf{E}_t \mathbf{Y}_{t+h} = \left(\mathbf{I_N} - \beta(L)_t)^{-1}\right) \beta_{0,t}$.

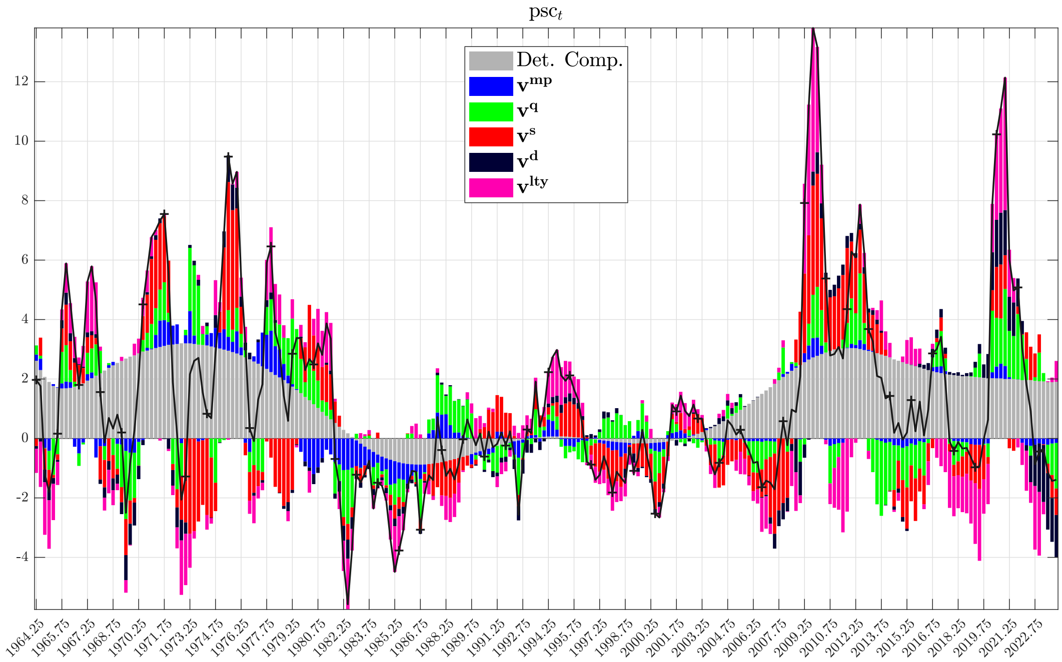

Figure C.5. Historical Decomposition of Public Sector Contribution to M4ex from 1964Q2–2024Q1

Notes: This top and bottom panels of this figure plot the historical decomposition of public sector contributions as a proportion of M4 from the TVP-VAR. Grey bars are the initial condition; Blue bars are monetary policy shocks; Green bars are bank funding shocks; Red bars are supply shocks, black bars are demand shocks; and magenta bars are long-term yield shocks. The black line is the actual time-series data we retrieve from the structural model.

C.3 Counterfactual experiments: full plots

Figures C.6 and C.7 show posterior median and posterior credible sets for our two counterfactual experiments analyzing the paths of the output gap, inflation and long-term yields as a result of QE (QT) episodes.

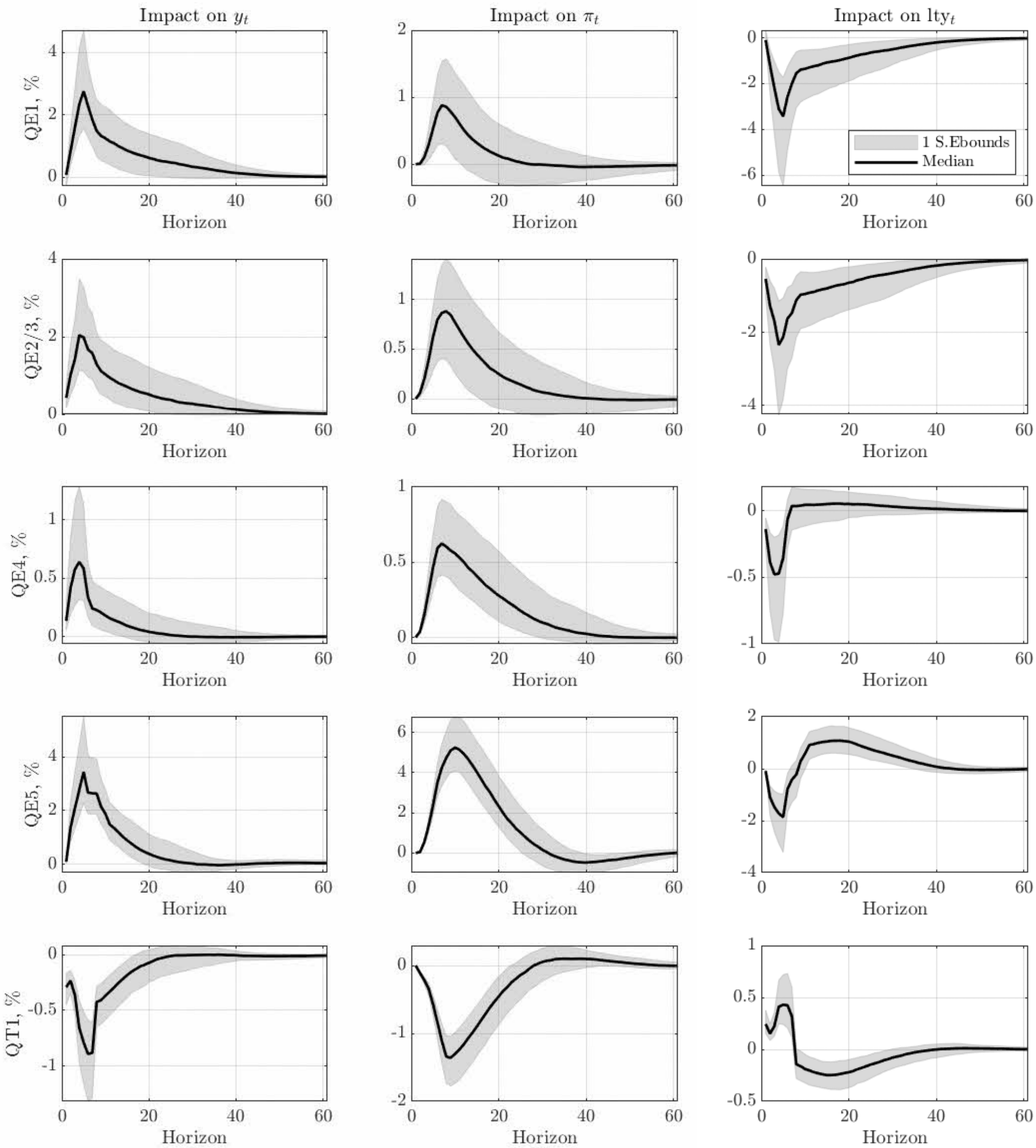

Figure C.6. Counterfactual Experiments: Quantitative Easing/Tightening Episodes
Notes: This figure plots the posterior median and 68% posterior credible intervals for counterfactual paths of the output gap, y_t; CPI inflation, π_t; and long-term yields, lty_t across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. The counterfactual shows the impact that each round of QE (QT) has in isolation of all other shocks.

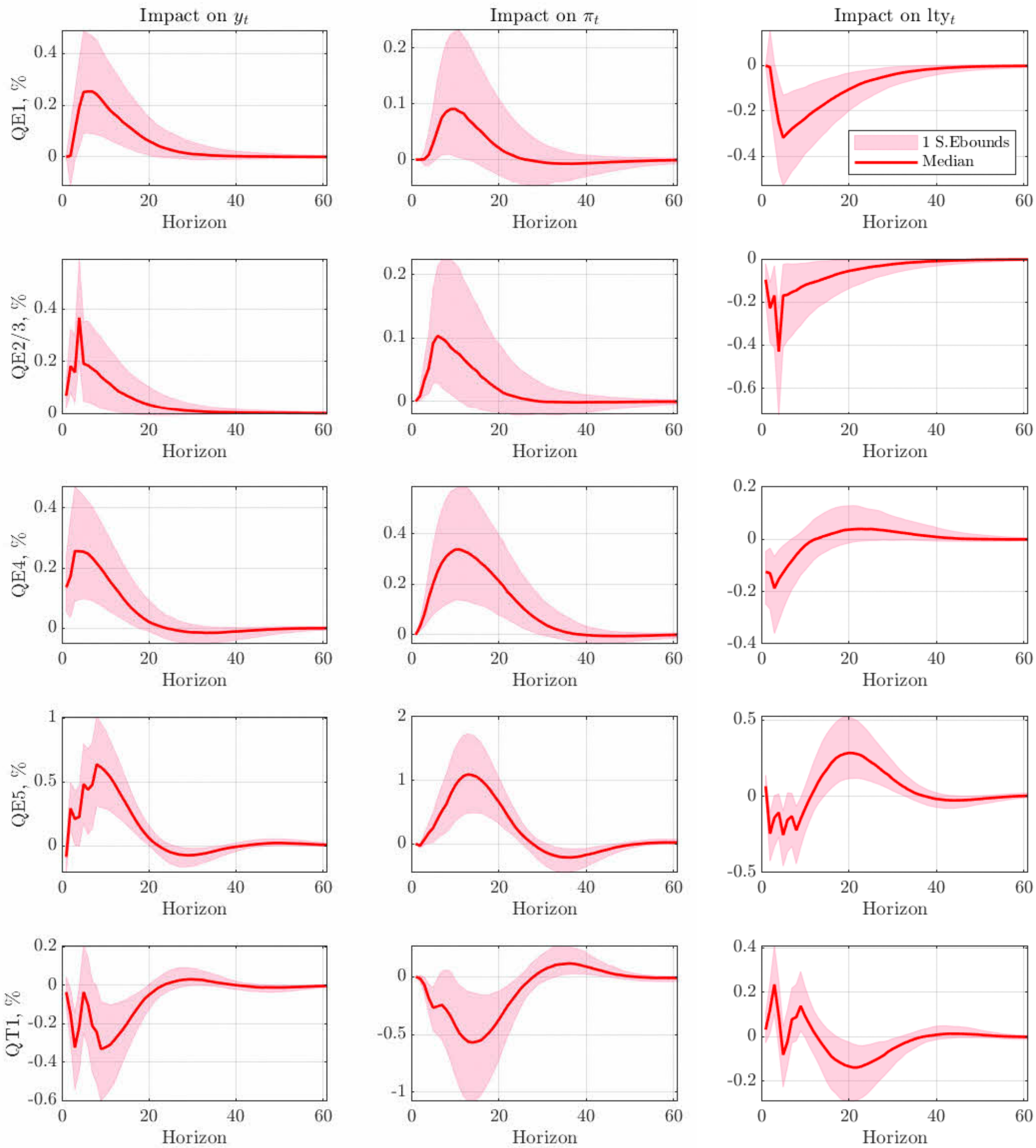

Figure C.7. Counterfactual Experiments: Quantitative Easing/Tightening Episodes
Notes: This figure plots the posterior median and 68% posterior credible intervals for counterfactual paths of the output gap, y_t; CPI inflation, π_t; and long-term yields, lty$_t$ across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. The counterfactual shows the shows us the unanticipated proportion of QE (QT) for economic variables.

C.4 Additional transmission mechanism plots

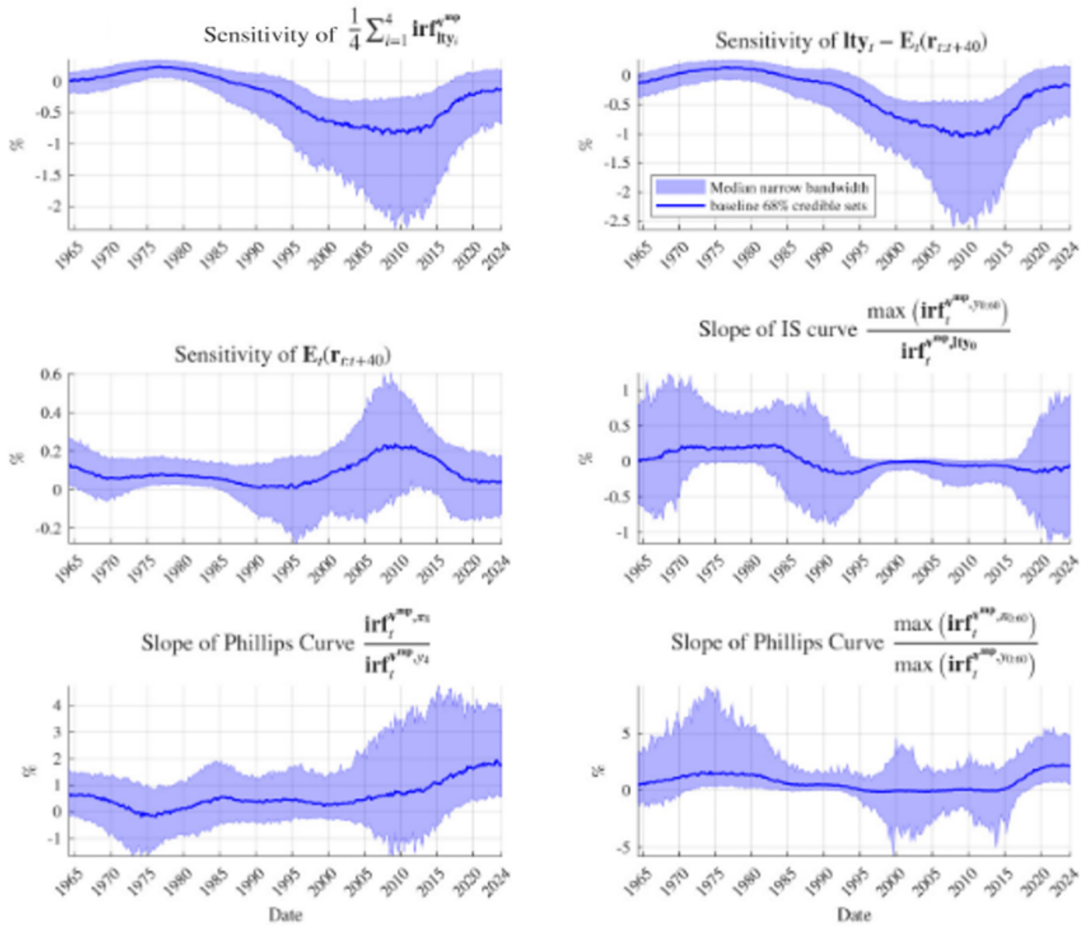

Figure C.8. Transmission Mechanism breakdown of monetary policy shocks ($\mathbf{v^{mp}}$) shocks
Notes: The top left panel of this figure shows the sensitivity of the term premium which we proxy as the 1-year average of the current 10-year yield rate, lty_t, minus the time t expectation of the average level of Bank rate over the next 40 quarters. We back out the expectation from the impulse response functions of r_t with respect to a monetary policy shock. The top right panel shows the 1-year average of the time t expectation of the average level of Bank rate over the next 40 quarters. The bottom right panel shows the slope of the IS curve which represents the sensitivity of demand to yields. We proxy this as the ratio between time t peak horizon response of the impulse response function of the output gap to monetary policy shocks, divided by the impact response of lty_t to a monetary policy shock. The bottom right panel shows the slope of the Phillips curve which we proxy as the ratio of the peak horizon impulse response of inflation at time t with respect to bank funding shocks and the corresponding peak horizon response at time t of the output gap.

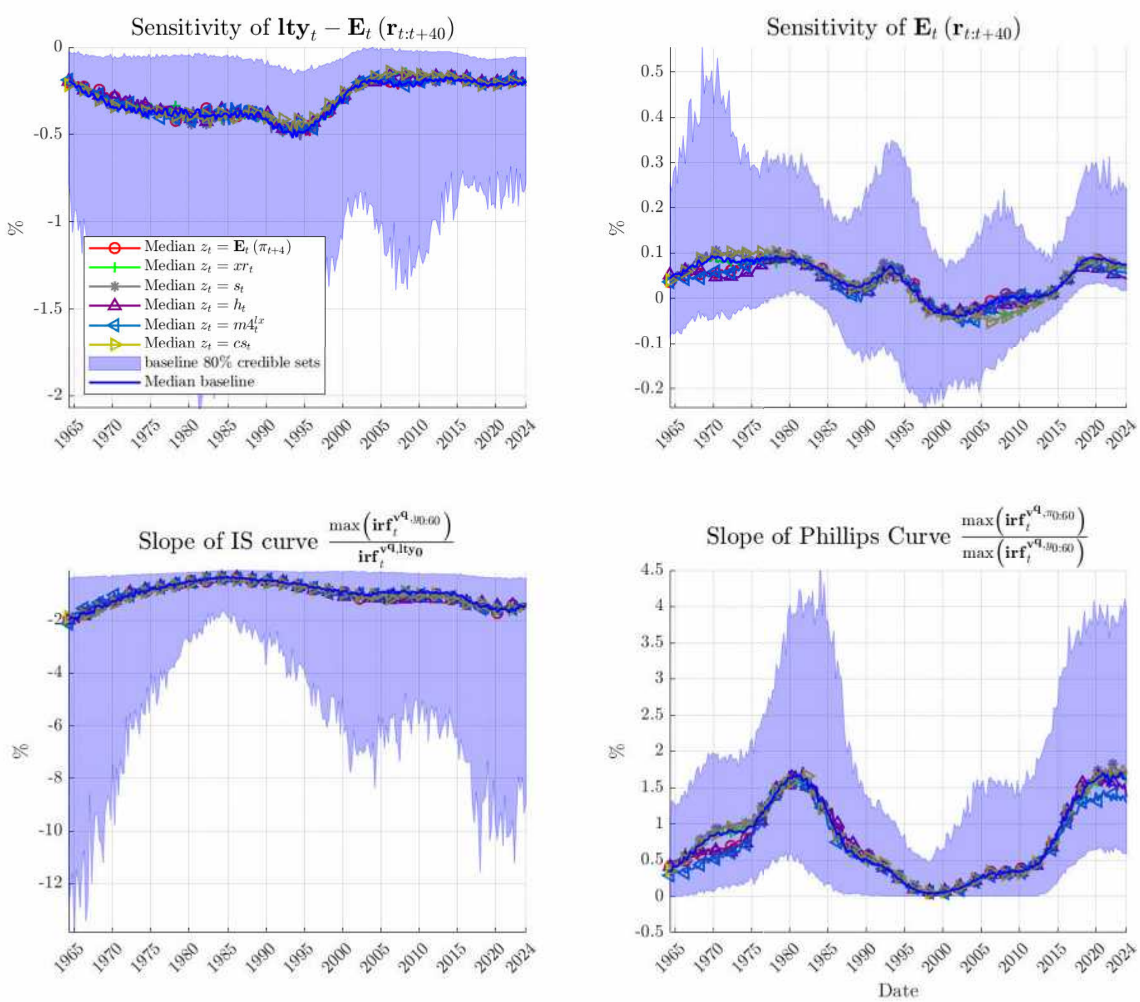

Figure C.9. Transmission Mechanism of bank funding shocks (v^q) with additional variables added to the VAR

Notes: The top left panel of this figure shows the sensitivity of the term premium which we proxy as the 1-year average of the current 10-year yield rate, lty_t, minus the time t expectation of the average level of Bank rate over the next 40 quarters. We back out the expectation from the impulse response functions of r_t with respect to a bank funding shock. This represents the asset substitutability channel throughout QE (QT) and other unconventional monetary policy episodes. The top right panel shows the 1-year average of the time t expectation of the average level of Bank rate over the next 40 quarters, which represents the signaling channel throughout QE (QT) and other unconventional monetary policy episodes. The bottom right panel shows the slope of the IS curve which represents the sensitivity of demand to yields. We proxy this as the ratio between time t peak horizon response of the impulse response function of the output gap to bank funding shocks, divided by the impact response of lty_t to the bank funding shock. The bottom right panel shows the slope of the Phillips curve which we proxy as the ratio of the peak horizon impulse response of inflation at time t with respect to bank funding shocks and the corresponding peak horizon response at time t of the output gap. Solid blue and shaded areas represent results from our baseline specification. The red, green, gray and purple lines show results that add $z_t = \{\mathbf{E}_t\left(\pi_{t+4}\right), xr_t, s_t, h_t, cs_t, M4Lx_t\}$ to the information set. These are: are time t one-year ahead inflation expectations, the effective exchange rate, annualized stock returns, and annual house price growth, investment-grade corporate bond spreads and bank and building society lending growth to the real economy respectively.

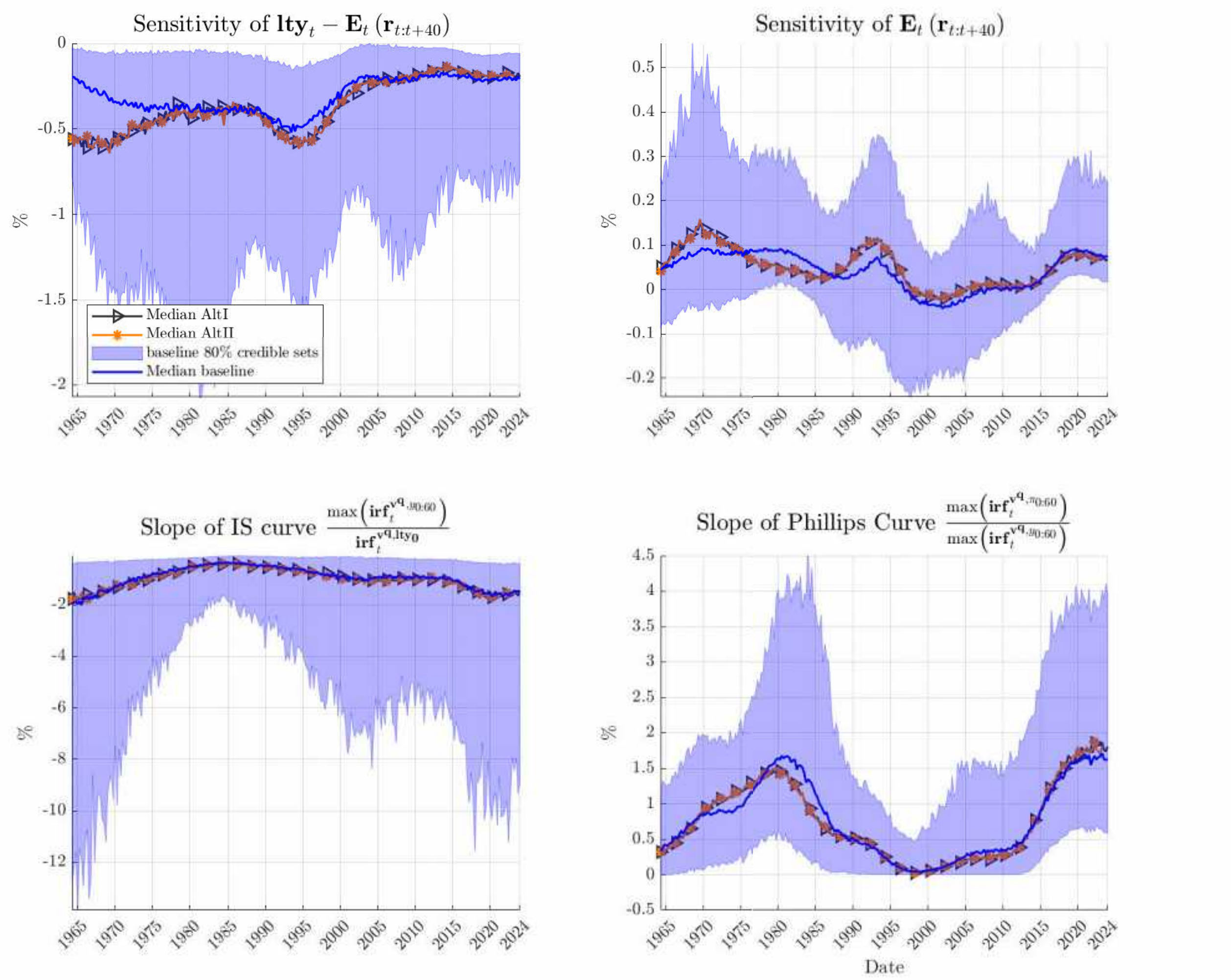

Figure C.10. Transmission Mechanism of bank funding (v^q) shocks under alternative proxies for public sector funding from banks

Notes: The top left panel of this figure shows the sensitivity of the term premium which we proxy as the 1-year average of the current 10-year yield rate, lty_t, minus the time t expectation of the average level of Bank rate over the next 40 quarters. We back out the expectation from the impulse response functions of r_t with respect to a bank funding shock. This represents the asset substitutability channel throughout QE (QT) and other unconventional monetary policy episodes. The top right panel shows the 1-year average of the time t expectation of the average level of Bank rate over the next 40 quarters, which represents the signaling channel throughout QE (QT) and other unconventional monetary policy episodes. The bottom right panel shows the slope of the IS curve which represents the sensitivity of demand to yields. We proxy this as the ratio between time t peak horizon response of the impulse response function of the output gap to bank funding shocks, divided by the impact response of lty_t to the bank funding shock. The bottom right panel shows the slope of the Phillips curve which we proxy as the ratio of the peak horizon impulse response of inflation at time t with respect to bank funding shocks and the corresponding peak horizon response at time t of the output gap. Solid blue and shaded areas represent results from our baseline specification. Solid blue and shaded areas represent results from our baseline specification, the black and orange lines are posterior median estimates using alternative proxies for psc_t.

C.5 Inflation persistence

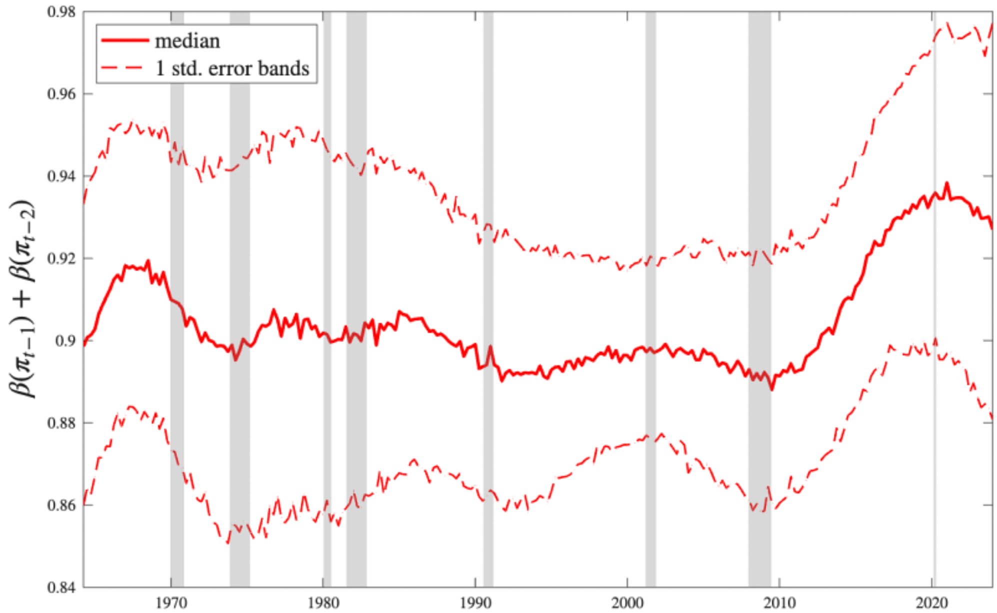

Figure C.11. Inflation Persistence: Reduced form autoregressive coefficient from 1964Q2–2024Q1
Notes: The figure plots the reduced form inflation autoregressive coefficients summed over lags. The solid line is the posterior median, the dashed lines are the one standard deviation bounds corresponding to the 16th and 84th percentiles of the posterior distribution. Gray bars show NBER recession dates.

D Replacing the output gap with real GDP growth

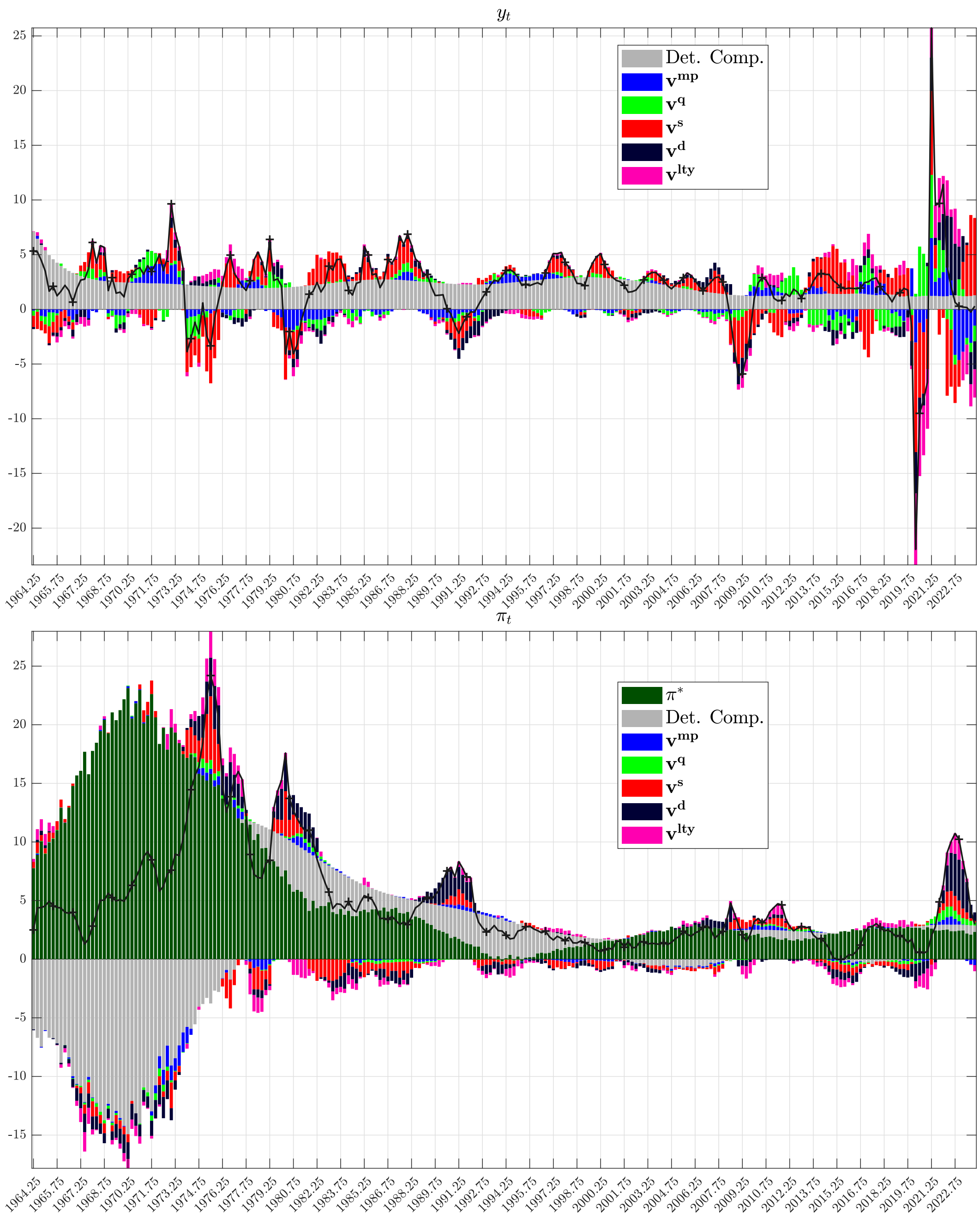

Figure D.1. Historical decomposition of the output gap and inflation from 1964Q2–2024Q1
Notes: This top and bottom panels of this figure plot the historical decomposition of output growth, y_t, and annual CPI inflation, π_t respectively from the TVP-VAR. Grey bars are the initial condition; Blue bars are monetary policy shocks; Light Green bars are bank funding shocks; Red bars are supply shocks, black bars are demand shocks; and magenta bars are long-term yield shocks. The black line is the actual time-series data we retrieve from the structural model. For CPI inflation we also report the contribution of trend inflation π^* as dark green bars which stem from the corresponding long-run coefficient VAR relationship $\lim_{h\to\infty} \mathbf{E}_t \mathbf{Y}_{t+h} = \left(\mathbf{I_N} - \beta(L)_t)^{-1}\right) \beta_{0,t}$.

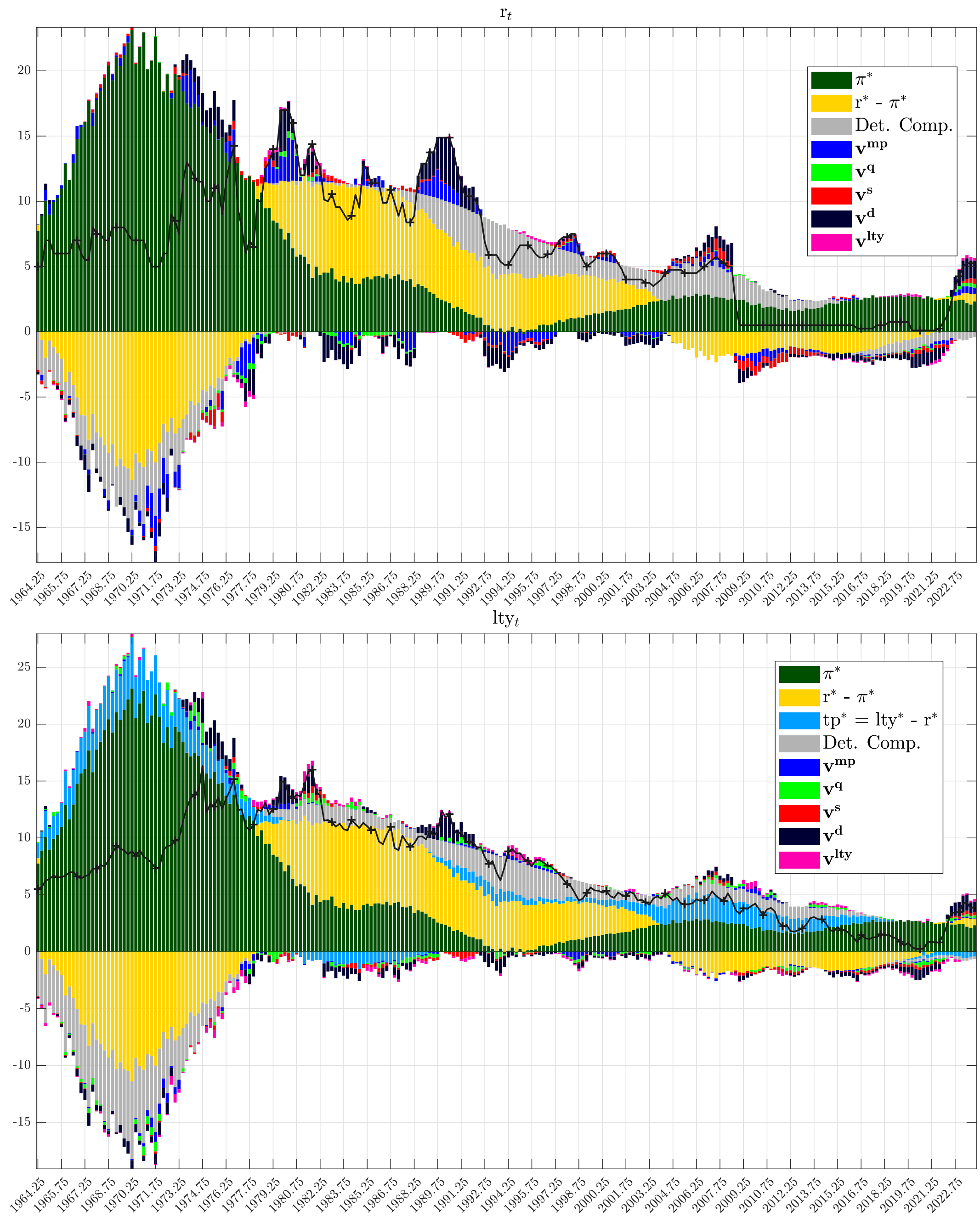

Figure D.2. Historical decomposition of the Bank Rate and long-term yields from 1964Q2–2024Q1

Notes: This top and bottom panels of this figure plot the historical decomposition of the Bank Rate, r_t, and long-term yields, lty_t respectively from the TVP-VAR. Grey bars are the initial condition; Blue bars are monetary policy shocks; Green bars are bank funding shocks; Red bars are supply shocks, black bars are demand shocks; and magenta bars are long-term yield shocks. The black line is the actual time-series data we retrieve from the structural model. For the Bank Rate we report the contributions of the natural rate of interest r^* which are violet bars. For the long-term yield we report the contributions of trend inflation, π^* (dark green bars); the real natural rate of interest, $(r^*_t - \pi^*_t)$ (yellow bars), and the equilibrium term premium, $tp^*_t = lty^*_t - r^*_t$ (light blue bars). All of which are manipulations from the long-run coefficient VAR relationship $\lim_{h\to\infty} \mathbf{E}_t \mathbf{Y}_{t+h} = \left(\mathbf{I_N} - \beta(L)_t)^{-1}\right) \beta_{0,t}$.

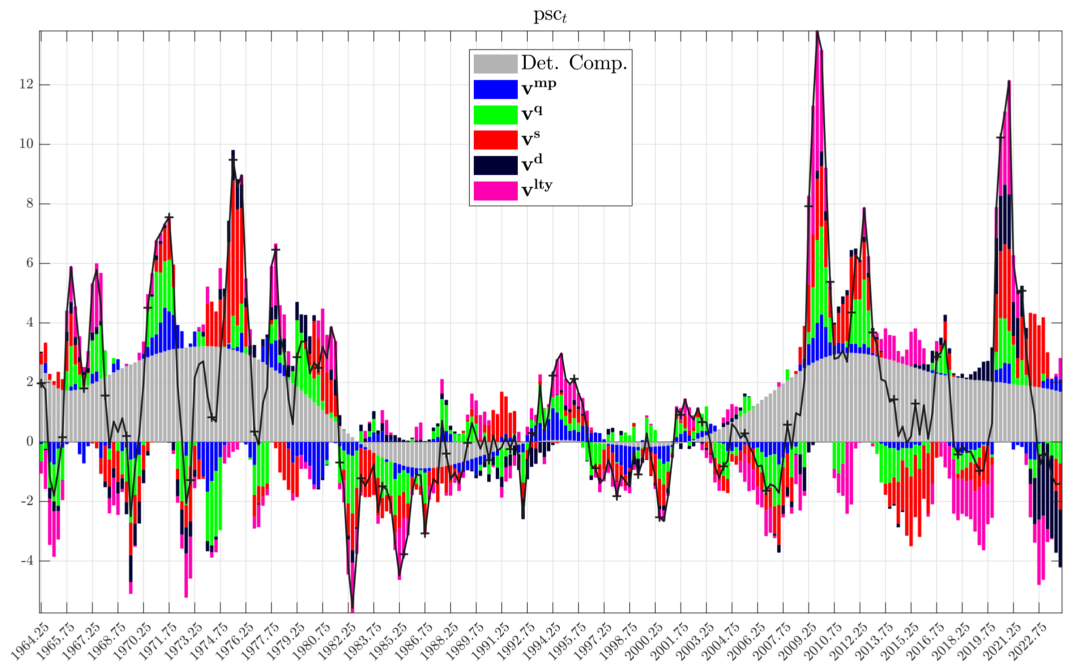

Figure D.3. Historical Decomposition of Public Sector Contribution to M4ex from 1964Q2–2024Q1

Notes: This top and bottom panels of this figure plot the historical decomposition of public sector contributions as a proportion of M4 from the TVP-VAR. Grey bars are the initial condition; Blue bars are monetary policy shocks; Green bars are bank funding shocks; Red bars are supply shocks, black bars are demand shocks; and magenta bars are long-term yield shocks. The black line is the actual time-series data we retrieve from the structural model.

E Additional counterfactual plots

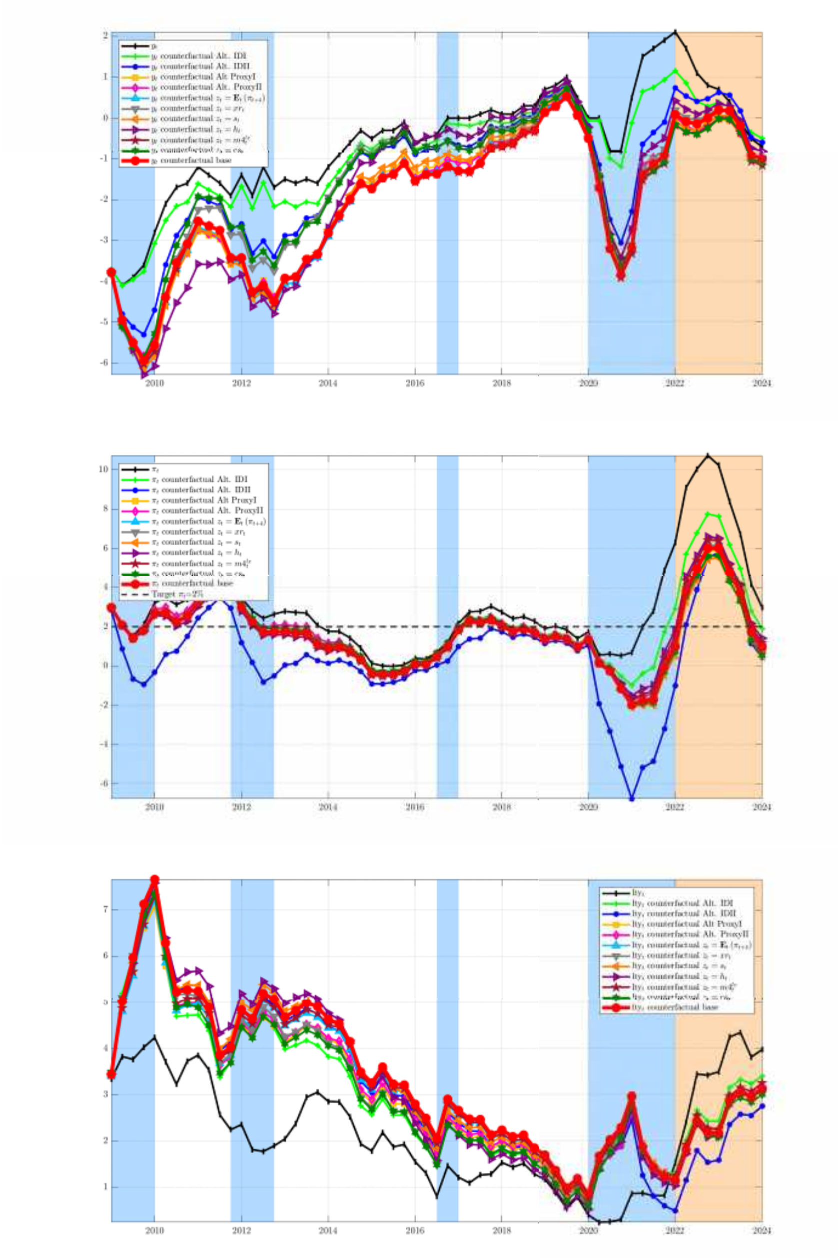

Figure E.4. Counterfactual paths and contributions of components from 2009 to 2024
Notes: This figure plots the counterfactual paths in the absence of QE/QT policies for the output gap, inflation, and long-term gilt yields, respectively. We show such paths for our baseline results and those considering: i) alternative identification schemes for bank funding shocks, Alt IDI imposes a zero impact response of y_t, Alt IDII imposes a impact positive response on y_t, π_t; two alternative proxies of psc_t, AltI (AltII) changes the numerator (denominator) to public sector contributions held by the Bank (the par-value of yields outstanding); Additional variables we add to the information set are $z_t = \{\mathbf{E}_t(\pi_{t+4}), xr_t, s_t, h_t, cs_t, M4Lx_t\}$, these are inflation expectations, exchange rate growth, stock returns, and house price growth, investment-grade corporate bond spreads and bank and building society lending growth to the real economy respectively. Alongside the counterfactual paths, we plot the observable data (black lines).

F Alternative identification schemes

Table F.1. Contemporaneous Sign Restrictions: Alternative Identification Schemes
Notes: This table reports two alternative identification schemes for the contemporaneous responses of: the Bank Rate, r_t; public sector contributions to the banking sector as a proportion of broad money stock, psc_t; the long-term (i.e., 10-year gilt) gilt yield, lty_t; inflation, π_t; and output gap, y_t to structural shocks. $\mathbf{v}_t^{\mathbf{mp}}$ is a monetary policy shock; $\mathbf{v}_t^{\mathbf{q}}$ is a bank funding shock; $\mathbf{v}_t^{\mathbf{s}}$ is a supply shock; $\mathbf{v}_t^{\mathbf{d}}$ is a demand non-policy shock; and $\mathbf{v}_t^{\mathbf{lty}}$ is a long-term yield shock, in panels A and B respectively. 0 denotes a zero response on impact, "x" denotes no restriction.

	A: Alternative I						**B:** Alternative II				
	$\mathbf{v}_t^{\mathbf{mp}}$	$\mathbf{v}_t^{\mathbf{q}}$	$\mathbf{v}_t^{\mathbf{s}}$	$\mathbf{v}_t^{\mathbf{d}}$	$\mathbf{v}_t^{\mathbf{lty}}$		$\mathbf{v}_t^{\mathbf{mp}}$	$\mathbf{v}_t^{\mathbf{q}}$	$\mathbf{v}_t^{\mathbf{s}}$	$\mathbf{v}_t^{\mathbf{d}}$	$\mathbf{v}_t^{\mathbf{lty}}$
r_t	$\geq$	0	x	$\geq$	0	r_t	$\geq$	0	x	$\geq$	0
psc_t	x	$\geq$	x	x	$\geq$	psc_t	x	$\geq$	x	x	$\geq$
lty_t	x	$\leq$	x	$\geq$	$\geq$	lty_t	x	$\leq$	x	$\geq$	$\geq$
π_t	$\leq$	0	$\leq$	$\geq$	$\leq$	π_t	$\leq$	$\geq$	$\leq$	$\geq$	$\leq$
y_t	$\leq$	0	$\geq$	$\geq$	$\leq$	y_t	$\leq$	$\geq$	$\geq$	$\geq$	$\leq$

Here we consider the robustness of our identification of bank funding shocks. In Table F.1 we present two alternative combinations of sign and zero restrictions for bank funding shocks. Panel A shows our first alternative, where we impose a zero impact response of the output gap with respect to funding shocks; all other responses remain the same as our baseline. Panel B shows our second alternative, where we allow inflation to increase on impact of bank funding shocks; holding all other responses as in our baseline scheme.

Figures F.7 and F.8 plot analogous results using our respective alternative identification schemes for our counterfactual experiments in Figure 7. Considering results from our first alternative identification scheme, observe the responses are all smaller in magnitude for our first counterfactual experiment. For example QE5 has a peak impact of 0.9 and 3.9 percentage points on the output gap and inflation, respectively; the response of long-term yields remain similar. Regarding the unanticipated effect of QE (QT), the responses are all similar in the bottom row of Figure F.7 relative to those in the bottom row of Figure 7.

Regarding our second alternative identification scheme, the impact of each round of QE (QT) for the output gap and inflation is relatively larger than those in the top row of Figure 7; with those from quantifying the unanticipated effect of QE (QT) remaining similar. We also plot the posterior median transmission mechanisms for each alternative identification scheme along with those from our baseline identification scheme in Figure

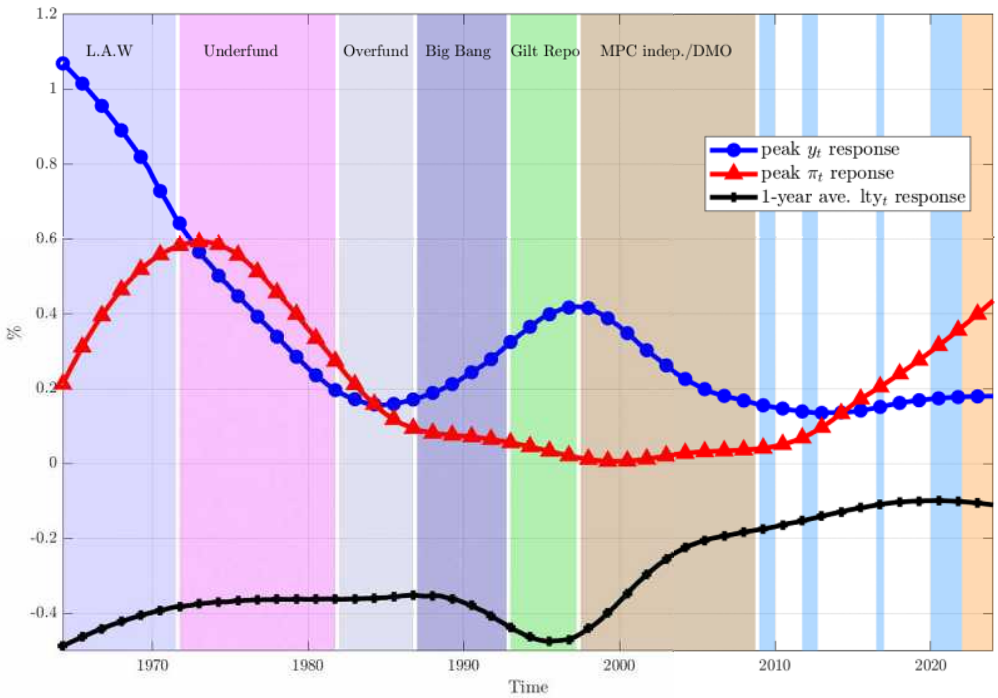

Figure F.5. Impulse response functions to bank funding shocks: Peak responses from 1964 to 2024: Alternative identification I

Notes: This figure plots the posterior median peak impulse response function of the output gap, y_t; inflation, π_t; and the long-term gilt yield, lty_t with respect to a bank funding shock, $\mathbf{v^q}_t$ from 1964Q2–2024Q1. We use the identification scheme in Panel A of Table F.1. In all cases, the peak response of the long-term gilt yield is the impact response. We normalize the distribution of impulse response functions such that a one-standard deviation bank funding shock causes psc_t to rise by 1%. We smooth peak responses for readability. Translucent bars indicate historical funding episodes and QE/QT periods. L.A.W is "Leaning Against the Wind" episode from 1963Q4–1971Q3. "Underfund", reflects the period of Competition and Credit control, the mid-1970s gilt strikes and the beginning of monetary targeting from 1971Q3–1981Q4. Overfund is the overfunding episode from 1982Q1–1986Q4, Big Bang reflects the introduction of the "Big Bang" reforms and the return to "full funding" over the period 1987Q1–1992Q4. Gilt Repo covers the more relaxed "full financing" rule and the development of the gilt repo market from 1993Q1–1997Q2. MPC indep/DMO is when the Monetary Policy Committee obtained independence and the Debt Management Office takes over debt management from 1997Q3–2008Q4. Translucent bars from 2009 – 2024 indicate the various rounds of QE/QT in the UK.

F.9. In comparing each panel of with those in Figure 6 it is clear there are no major differences to those in our main results.

In general, our baseline identification scheme delivers results that are most consistent with existing literature. Notably posterior credible intervals from each of our counterfactuals and transmission mechanism plots for each identification scheme overlap. Therefore although median estimates differ, the set of structural models we identify are consistent

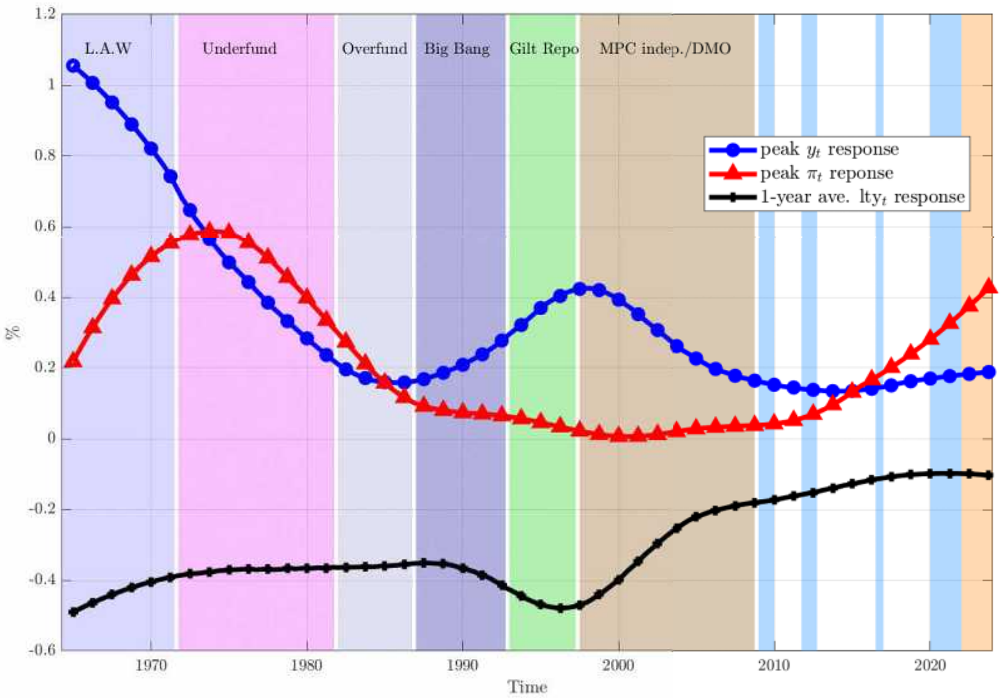

Figure F.6. Impulse response functions to bank funding shocks: Peak responses from 1964 to 2024: Alternative identification II

Notes: This figure plots the posterior median peak impulse response function of the output gap, y_t; inflation, π_t; and the long-term gilt yield, lty_t with respect to a bank funding shock, $\mathbf{v^q}_t$ from 1964Q2–2024Q1. We use the identification scheme in Panel B of Table F.1. In all cases, the peak response of the long-term gilt yield is the impact response. We normalize the distribution of impulse response functions such that a one-standard deviation bank funding shock causes psc_t to rise by 1%. We smooth peak responses for readability. Translucent bars indicate historical funding episodes and QE/QT periods. L.A.W is "Leaning Against the Wind" episode from 1963Q4–1971Q3. "Underfund", reflects the period of Competition and Credit control, the mid-1970s gilt strikes and the beginning of monetary targeting from 1971Q3–1981Q4. Overfund is the overfunding episode from 1982Q1–1986Q4, Big Bang reflects the introduction of the "Big Bang" reforms and the return to "full funding" over the period 1987Q1–1992Q4. Gilt Repo covers the more relaxed "full financing" rule and the development of the gilt repo market from 1993Q1–1997Q2. MPC indep/DMO is when the Monetary Policy Committee obtained independence and the Debt Management Office takes over debt management from 1997Q3–2008Q4. Translucent bars from 2009 – 2024 indicate the various rounds of QE/QT in the UK.

with one another from different assumptions around how bank funding shocks track and capture the effects of unconventional monetary policy.

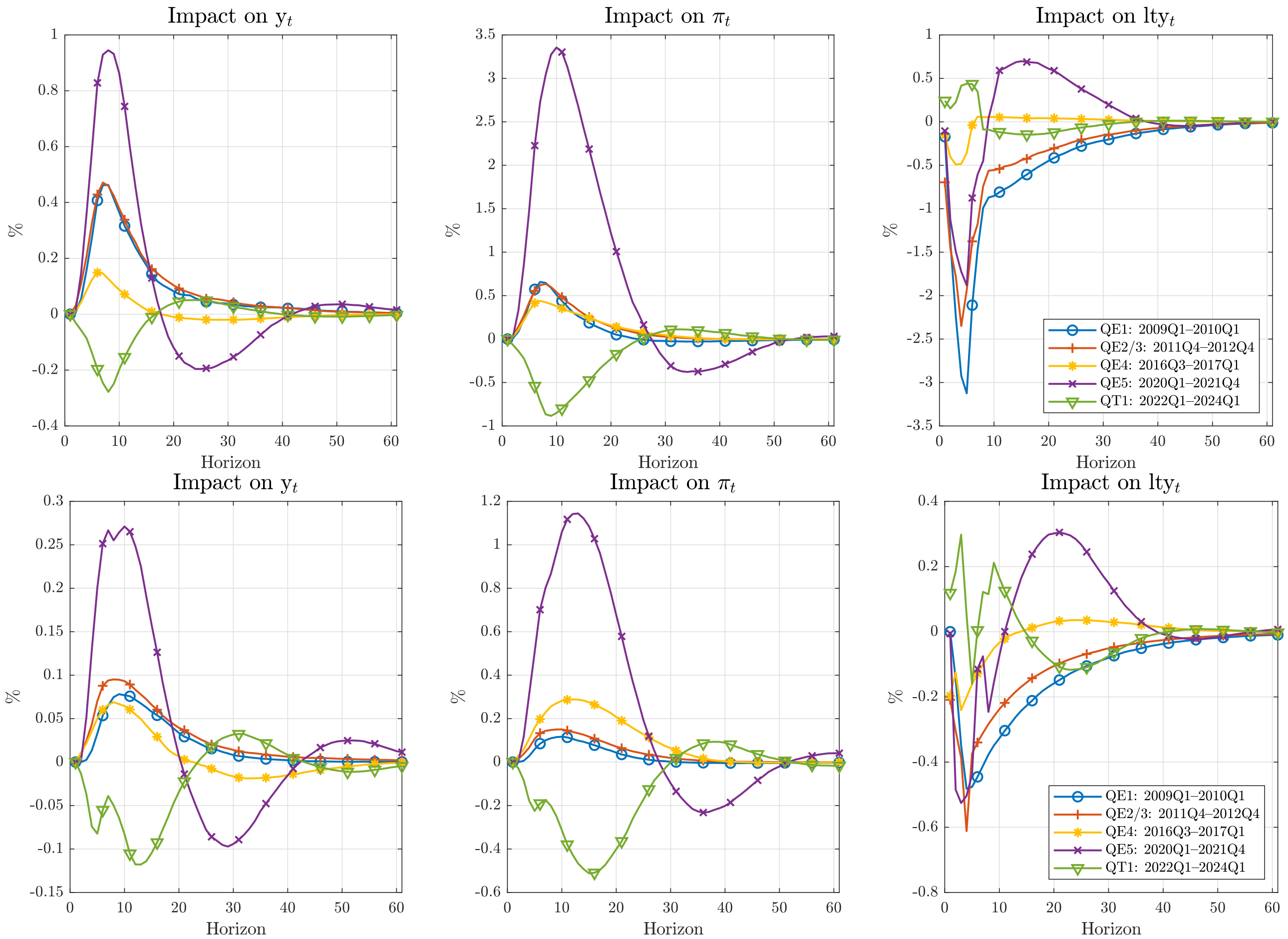

Figure F.7. Alternative I: Counterfactual experiments; Quantitative Easing/Tightening episodes
Notes: This figure plots the posterior median counterfactual paths of the output gap, y_t; CPI inflation, π_t; and long-term yields, lty_t across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. The counterfactual in the first row shows the impact that each round of QE (QT) has in isolation of all other shocks. This tells us the model-implied impact of each QE/QT episode. The counterfactual in the second row shows us the unanticipated proportion of QE (QT). We identify the bank funding shock as having a zero impact response on y_t, π_t, r_t and a positive impact on lty_t.

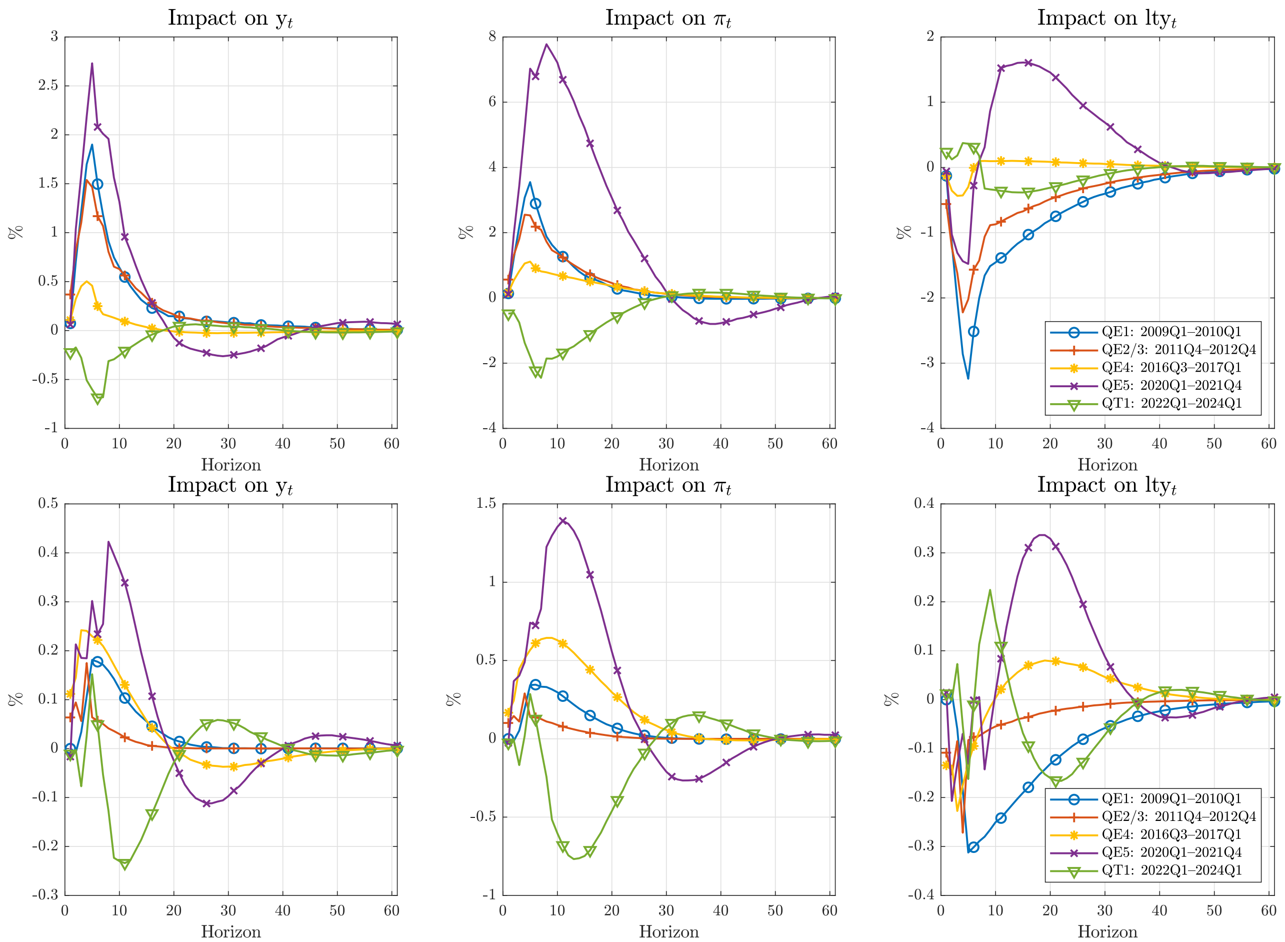

Figure F.8. Alternative II: Counterfactual experiments; Quantitative Easing/Tightening episodes
Notes: This figure plots the posterior median counterfactual paths of the output gap, y_t; CPI inflation, π_t; and long-term yields, lty_t across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. The counterfactual in the first row shows the impact that each round of QE (QT) has in isolation of all other shocks. This tells us the model-implied impact of each QE/QT episode. The counterfactual in the second row shows us the unanticipated proportion of QE (QT). We identify the bank funding shock as having a zero impact response on r_t and a positive impact on y_t, π_t, and lty_t.

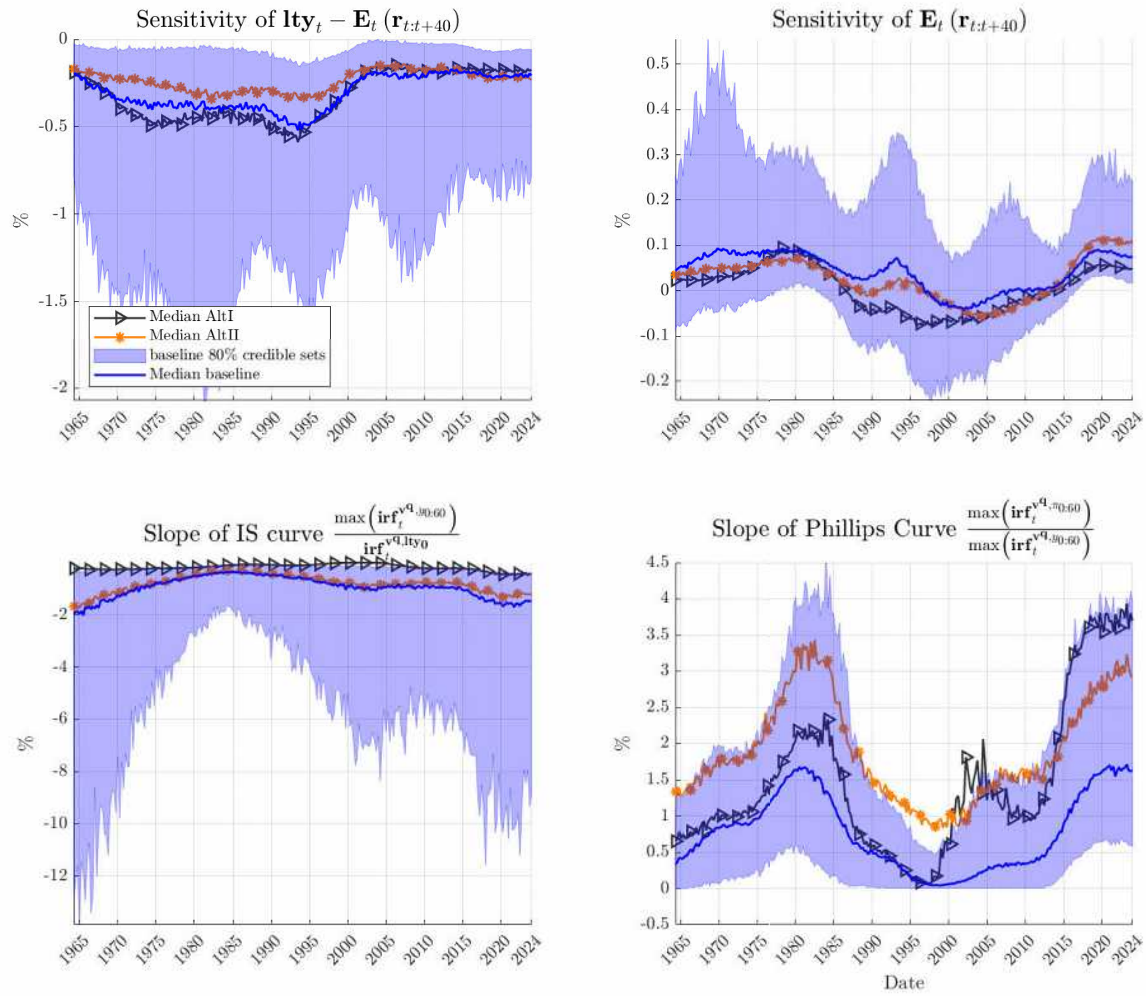

Figure F.9. Transmission Mechanism of bank funding (v^q) shocks under alternative identification schemes

Notes: The top left panel of this figure shows the sensitivity of the term premium which we proxy as the 1-year average of the current 10-year yield rate, lty_t, minus the time t expectation of the average level of Bank rate over the next 40 quarters. We back out the expectation from the impulse response functions of lty_t with respect to a bank funding shock. This represents the asset substitutability channel throughout QE (QT) and other unconventional monetary policy episodes. The top right panel shows the 1-year average of the time t expectation of the average level of Bank rate over the next 40 quarters, which represents the signaling channel throughout QE (QT) and other unconventional monetary policy episodes. The bottom right panel shows the slope of the IS curve which represents the sensitivity of demand to yields. We proxy this as the ratio between time t peak horizon response of the impulse response function of the output gap to bank funding shocks, divided by the impact response of lty_t to the bank funding shock. The bottom right panel shows the slope of the Phillips curve which we proxy as the ratio of the peak horizon impulse response of inflation at time t with respect to bank funding shocks and the corresponding peak horizon response at time t of the output gap. For ALTI, we identify the bank funding shock as having a zero impact response on y_t, π_t, r_t and a positive impact on lty_t. For ALTII we identify the bank funding shock as having a zero impact response on r_t and a positive impact on y_t, π_t, lty_t, respectively.

G Alternative Models: Priors, Posterior Simulation, and Additional Results

G.1 Linear Bayesian VAR

We first consider a simple linear Bayesian VAR model using a standard Minnesota prior with p=2 lags and N=5. The BVAR is given by

$$\mathbf{Y}_t = \mathbf{C}_0 + \sum_{p=1}^{P} \mathbf{C}_p \mathbf{Y}_{t-p} + \mathbf{e}_t, \quad \mathbf{e}_t \backsim N\left(0, \boldsymbol{\Omega}\right) \tag{G.1}$$

where $\mathbf{Y}_t \equiv [\mathrm{r}_t, \mathrm{psc}_t, \mathrm{lty}_t, \pi_t, y_t,]'$, with r_t being the Bank Rate, psc_t is public sector borrowings from the banking sector as a proportion of broad money stock (i.e., M4/M4ex), lty_t is the 10-year gilt yield, π_t is the annual rate of consumer price inflation, and y_t is annual real GDP growth. $\mathbf{C}_0$ is an $N \times 1$ vector of intercepts and $\mathbf{C}_p$ are $N \times N$ coefficient matrices. We estimate this model using a standard Gibbs sampler taking 20,000 draws from the posterior distribution and saving every 20$^{\text{th}}$ in order to reduce autocorrelation. We outline the priors in the Appendix; they are the same for those we use to train the TVP VAR model in the next Subsection.

G.2 The Time-varying Parameter VAR with Stochastic Volatility (Primiceri, 2005)

Next, we consider time-varying coefficient VAR models with stochastic volatility (TVP-VAR). We write the TVP-VAR with p=2 lags and N=5 variables as

$$\mathbf{Y}_t = \beta_{0,t} + \beta_{1,t}\mathbf{Y}_{t-1} + \beta_{2,t}\mathbf{Y}_{t-2} + \mathbf{e}_t \equiv \mathbf{X}_t'\mathbf{B}_t + \mathbf{e}_t \tag{G.2}$$

where $\mathbf{Y}_t \equiv [\mathrm{r}_t, \mathrm{psc}_t, \mathrm{lty}_t, \pi_t, y_t,]'$ is the same as in Equation G.1, and $\mathbf{X}_t'$ contains lags of $\mathbf{Y}_t$ and a constant. The VAR's time-varying coefficients are collected in the vector $\mathbf{B}_t$, and conditional on the roots of the VAR polynomial lying outside the unit circle evolve as a driftless random walk.

$$\mathbf{B}_t = \mathbf{B}_{t-1} + u_t, \quad u_t \backsim N(0, \bar{\mathbf{Q}}) \tag{G.3}$$

where $\bar{\mathbf{Q}}$ is a full matrix allowing parameters across equations to be correlated. Note that when $\bar{\mathbf{Q}}$=0, the model reduced to a constant parameter VAR with a stochastic volatility structure.

The innovations of the measurement equation, $\mathbf{e}_t$ are Normal with zero mean and time-varying covariance matrix $\mathbf{\Omega}_t$ which we factor as

$$\mathbf{\Omega}_t = \mathbf{A}_t^{-1}\mathbf{H}_t(\mathbf{A}_t^{-1})' \tag{G.4}$$

here $\mathbf{A}_t$ is a lower triangular matrix with unit diagonal containing the contemporaneous relations between variables in the model, and $\mathbf{H}_t$ is a diagonal matrix containing the reduced-form stochastic volatility innovations. Collecting the non-unit non-zero elements of $\mathbf{A}_t$ and the diagonal elements of $\mathbf{H}_t$ in the vectors, $\mathbf{h}_t \equiv [\mathbf{h}_{1,t}, \mathbf{h}_{2,t}, \mathbf{h}_{3,t}\mathbf{h}_{4,t}, \mathbf{h}_{h,t}]'$ and $\mathbf{a}_t \equiv [\mathbf{a}_{21,t}, \mathbf{a}_{31,t}, \ldots, \mathbf{a}_{54,t}]'$, they evolve as a geometric random walk and random walk respectively

$$\ln \mathbf{h}_{i,t} = \ln \mathbf{h}_{i,t-1} + \eta_t, \quad \eta_t \backsim N(0, \mathbf{Z_h}) \tag{G.5}$$

$$\mathbf{a}_t = \mathbf{a}_{t-1} + \zeta_t, \quad \zeta_t \backsim N(0, \mathbf{S}) \tag{G.6}$$

The innovations in the model, collected in the diagonal matrix $\mathbf{V}$, are jointly Normal and the structural shocks, $\mathbf{v}_t$ are such that, $\mathbf{e}_t \equiv \mathbf{A}_t^{-1}\mathbf{H}_t^{\frac{1}{2}}\mathbf{v}_t$. $\mathbf{S}$ is a block diagonal matrix, which implies that the non–zero and non–unit elements of $\mathbf{A}_t$ that belong to different rows evolve independently. This is a simplifying assumption that permits estimation of $\mathbf{A}_t$ equation by equation. We estimate this model using Bayesian methods. We obtain the deterministic component using the posterior mean of the coefficients and covariance matrix from a linear BVAR using the first 12 years of data. The Appendix provides specific information with regards to our choice of priors and an outline of the Markov-Chain Monte Carlo (MCMC) posterior simulation algorithm.

G.3 Prior for the Linear BVAR and initialising the Primiceri (2005) TVP-VAR

Our prior specification involves estimating a Bayesian fixed coefficient VAR (BVAR) model over the training sample. The priors imposed on this BVAR model combine the traditional Minnesota prior of Doan et al. (1984) and Litterman (1986) on the coefficient matrices with an inverse-Wishart prior on the BVAR's covariance matrix. In our specification, the prior mean on the coefficient matrix sets all elements equal zero, except those corresponding to the own first lag of each dependent variable which are set to 0.9. This imposes the prior belief that our variables exhibit persistence whilst simultaneously ensuring shrinkage of the other VAR coefficients to zero. The prior variance of the coefficient matrix is set similar to Litterman (1986). Our prior for the BVAR's covariance matrix follows an inverse-Wishart distribution with the prior scale matrix and degrees of freedom set to an N-dimensional identity matrix and 1+N respectively.

We estimate the BVAR using a standard Gibbs sampler. For the sake of brevity, we do not explicitly outline our algorithm since it is well documented; see e.g. Koop and Korobilis (2010). Our alternative prior specification essentially replaces the conventional Cogley and Sargent (2005) prior with the posterior means from the draws of an estimated BVAR over the training sample

$$\bar{\mathbf{B}}_{\text{BVAR}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{B}_i, \tag{G.7}$$

$$\overline{\text{V}(\mathbf{B})}_{\text{BVAR}} = \frac{1}{M}\sum_{i=1}^{M}\text{V}(\mathbf{B}_i), \tag{G.8}$$

$$\bar{\Sigma}_{\text{BVAR}} = \frac{1}{M}\sum_{i=1}^{M}\Sigma_i \tag{G.9}$$

respectively. Here M denotes the number of saved draws from the estimated BVAR which we set to 20,000. $\mathbf{B}_i$ and $\text{V}(\mathbf{B}_i)$ denote the ith draw of the coefficient matrix and the variance of the coefficient matrix respectively. Σ_i denotes the ith draw of the BVAR's covariance matrix. From these estimates, the deterministic component of the time-varying coefficient models, $\mathbf{B}_0$, $\mathbf{a}_0$, $\mathbf{h}_0$ are Normal and independent of one another, and the distributions of the hyperparameters. We set

$$\mathbf{B}_0 \backsim N\left[\bar{\mathbf{B}}_{\text{BVAR}}, 4 \cdot \overline{\text{V}(\mathbf{B})}_{\text{BVAR}}\right] \tag{G.10}$$

G.4 Additional priors for the TVP-VAR

for $\mathbf{a}_0$, $\mathbf{h}_0$, let $\bar{\Sigma}_{\text{BVAR}}$ be the posterior median of the estimated covariance matrix of the residuals from the time–invariant BVAR. Let C be the lower–triangular Choleski factor such that $CC' = \bar{\Sigma}_{\text{BVAR}}$. We then set

$$\ln \mathbf{h}_0 \backsim N(\ln \mu_0, 10 \times I_3) \tag{G.11}$$

where μ_0 collects the logarithms of the squared elements along the diagonal of C. We divide each column of C by the corresponding element on the diagonal; call this matrix $\tilde{C}$.

We then set

$$\mathbf{a}_0 \backsim N\left[\tilde{\mathbf{a}}_0, \tilde{\mathrm{V}}(\tilde{\mathbf{a}}_0)\right] \tag{G.12}$$

with $\tilde{\mathbf{a}}_0 \equiv [\tilde{\mathbf{a}}_{0,11}, \tilde{\mathbf{a}}_{0,21}, \tilde{\mathbf{a}}_{0,31}]'$ which is a vector collecting all the elements below the diagonal of $\tilde{C}^{-1}$. We assume $\tilde{\mathrm{V}}(\tilde{\mathbf{a}}_0)$ is diagonal with each element equal to 10 times the absolute value of the corresponding element of $\tilde{\mathbf{a}}_0$. This is an arbitrary prior but correctly scales the variance of each element of $\mathbf{a}_0$ to account for their respective magnitudes.

$\bar{\mathbf{Q}}$ is set to follow an inverse–Wishart distribution,

$$\bar{\mathbf{Q}} \backsim IW(\underline{\mathbf{Q}}^{-1}, \mathrm{T}_0) \tag{G.13}$$

where $\underline{\mathbf{Q}} = (1 + \dim(\mathbf{B}_t)) \cdot \overline{\mathrm{V}}(\bar{\mathbf{B}}_{\mathrm{BVAR}}) \cdot 3.4 \times 10^{-4}$. The prior degrees of freedom, $(1 + \dim(\mathbf{B}_t))$, are the minimum allowed for the prior to be proper. Our choice of scaling parameter of 3.4×10^{-4} is consistent with Cogley and Sargent (2005). We have also estimated our models using different priors, we allowed for a more restrictive scaling parameter of 1.0×10^{-4} and have also set the degrees of freedom to be the length of the training sample; in our case this is 80. The scaling parameter essentially sets the amount of drift within the $\mathbf{B}_t$ matrices. The results and conclusions presented within the main body are robust to changing the value of the scaling parameter, and the prior degrees of freedom imposed.

The blocks of $\mathbf{S}$ are also assumed to follow inverse–Wishart distributions with prior degrees of freedom equal to the minimum allowed (i.e. $1 + \dim(\mathbf{S}_i)$).

$$\mathbf{S}_1 \backsim IW(\underline{\mathbf{S}}_1^{-1}, 2) \tag{G.14}$$

$$\mathbf{S}_2 \backsim IW(\underline{\mathbf{S}}_2^{-1}, 3) \tag{G.15}$$

$$\mathbf{S}_3 \backsim IW(\underline{\mathbf{S}}_3^{-1}, 4) \tag{G.16}$$

$$\mathbf{S}_4 \backsim IW(\underline{\mathbf{S}}_4^{-1}, 5) \tag{G.17}$$

we set $S_1, \ldots, S_5$ in accordance with $\tilde{\mathbf{a}}_0$ such that our calibration is consistent with setting $\mathbf{S}_1, \ldots, \mathbf{S}_5$ to 10^{-4} times the corresponding diagonal block of $\tilde{\mathrm{V}}(\tilde{\mathbf{a}}_0)$. The variances for the stochastic volatility innovations, as in Cogley and Sargent (2005), follow an inverse–

Gamma distribution for the elements of $\mathbf{Z_h}$,

$$\mathbf{Z}_{\mathbf{h},i,i} \backsim IG(\frac{10^{-4}}{2}, \frac{1}{2}) \tag{G.18}$$

G.5 Posterior Simulation of the Primiceri (2005) TVP-VAR

In order to simulate the posterior distribution of the hyperparameters and states, conditional on the data, we implement the following MCMC that combines elements from Primiceri (2005) and Cogley and Sargent (2005).

1) *Draw elements of* $\mathbf{B}_t$ Conditional on $\mathbf{Y}^T$, $\mathbf{a}^T$ and $\mathbf{H}^T$, the observation equation (1) is linear with Gaussian innovations with a known covariance matrix. Factoring the density of $\mathbf{B}_t$, $p(\mathbf{B}_t)$ in the following manner

$$p(\mathbf{B}^T|\mathbf{y}^T, \mathbf{A}^T, \mathbf{H}^T, V) = p(\mathbf{B}_T|\mathbf{Y}^T, \mathbf{A}^T, \mathbf{H}^T, \mathbf{V}) \prod_{t=1}^{T-1} p(\mathbf{B}_t|\mathbf{B}_{t+1}, \mathbf{Y}^t, \mathbf{A}^T, \mathbf{H}^T, \mathbf{V}) \tag{G.19}$$

the Kalman filter recursions pin down the first element on the right hand side of the above; $p(\mathbf{B}_T|Y^T, \mathbf{A}^T, \mathbf{H}^T, \mathbf{V}) \backsim N(\mathbf{B}_T, P_T)$, with P_T being the precision matrix of $\mathbf{B}_T$ from the Kalman filter. The remaining elements in the factorisation are obtained via backward recursions as in Cogley and Sargent (2005). Since $\mathbf{B}_t$ is conditionally Normal

$$\mathbf{B}_{t|t+1} = P_{t|t}P_{t+1|t}^{-1}(\mathbf{B}_{t+1} - \mathbf{B}_t) \tag{G.20}$$

$$P_{t|t+1} = P_{t|t} - P_{t|t}P_{t+1|t}^{-1}P_{t|t} \tag{G.21}$$

which yields, for every t from $T-1$ to 1, the remaining elements in the observation equation (1). More precisely, the backward recursion begins with a draw, $\tilde{\mathbf{B}}_T$ from $N(\mathbf{B}_T, P_T)$. Conditional on $\tilde{\mathbf{B}}_T$, the above produces $\mathbf{B}_{T-1|T}$ and $P_{T-1|T}$. This permits drawing $\tilde{\mathbf{B}}_{T-1}$ from $N(\mathbf{B}_{T-1|T}, P_{T-1|T})$ until t=1.

2) *Drawing elements of* $\mathbf{a}_t$ Conditional on $\mathbf{Y}^T$, $\mathbf{B}^T$ and $\mathbf{H}^T$ we follow Primiceri (2005) and note that (1) can be written as

$$\mathbf{A}_t\tilde{\mathbf{Y}}_t \equiv \mathbf{A}_t(\mathbf{Y}_t - \mathbf{X}_t'\mathbf{B}_t) = \mathbf{A}_t\mathbf{e}_t \equiv \mathbf{v}_t \tag{G.22}$$

$$\text{Var}(\mathbf{v}_t) = \mathbf{H}_t \tag{G.23}$$

The corresponding observation equations and the state equation permit drawing the elements of $\mathbf{a}_t$ equation by equation using the same algorithm as above; assuming S is block diagonal.

3) *Drawing elements of* $\mathbf{H}_t$ Conditional on $\mathbf{Y}^T$, $\mathbf{B}^T$ and $\mathbf{a}^T$, the orthogonal innovations u_t, $\text{Var}(\mathbf{e}_t) = \mathbf{H}_t$ are observable. Following Jacquier et al. (2002) the stochastic volatilities, $\mathbf{h}_{i,t}$'s, are sampled element by element; Cogley and Sargent (2005) provide details in Appendix B.2.5 of their paper.

4) *Drawing the hyperparameters* Conditional on Y^T, $\mathbf{B}^T$, $\mathbf{H}_t$ and $\mathbf{a}^T$, the innovations in $\mathbf{B}_t$, $\mathbf{a}_t$ and $\mathbf{h}_{i,t}$'s are observable, which allows one to draw the elements of $\bar{\mathbf{Q}}$, $\mathbf{S}_1, \ldots, \mathbf{S}_5$ and the $\mathbf{Z}_{\mathbf{h},i,i}$ from their respective distributions.

G.6 Additional Results from Alternative Models

Here we present results from alternative models. First we present those from a linear BVAR. Second we examine results from the TVP-VAR with stochastic volatility as in Primiceri (2005). Finally, we show a snap shot of results from a battery of TVAR models we estimate in the spirit of Alessandri and Mumtaz (2017).

G.6.1 Linear Bayesian VAR Model

Figure G.1 shows the posterior median and one-standard error credible intervals of impulse response functions of the output gap, inflation, the Bank Rate, and the long-term yield with respect to bank funding shocks and monetary policy shocks. Historical decompositions are available upon request.

G.6.2 Primiceri (2005) Results

Figure G.2 presents the counterfactual results under the TVP–VAR model analogous to those we report in the top row of Figure 7. Historical decompositions are available upon request.

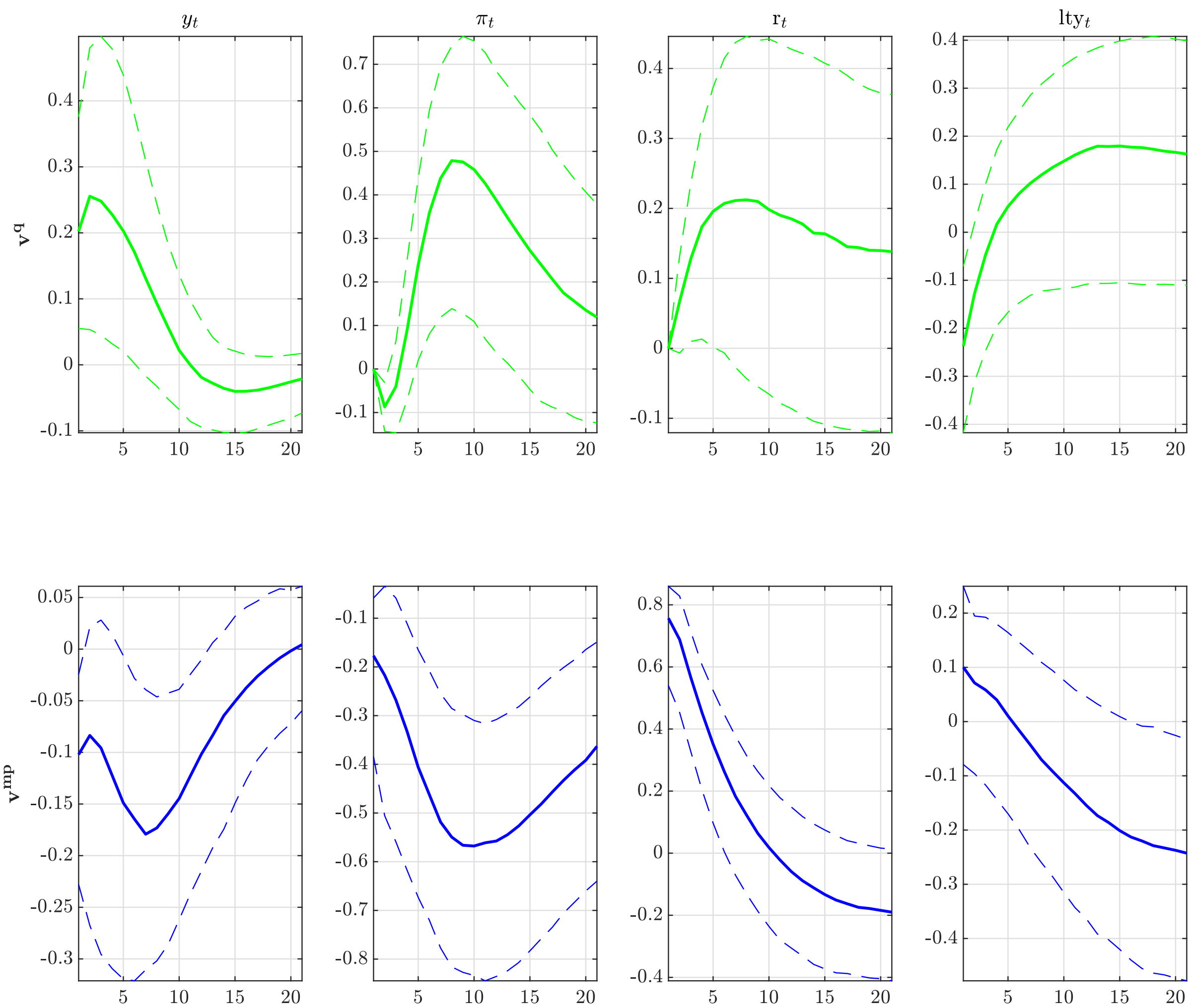

Figure G.1. Impulse Response functions for Bank Funding and Monetary Policy Shocks
Notes: This top and bottom panels of this figure plot the impulse response functions of macroeconomic variables with respect to quantitative easing and monetary policy shocks, respectively. y_t is the output gap; π_t is inflation; r_t is the Bank Rate; and lty_t is the long-term yield.

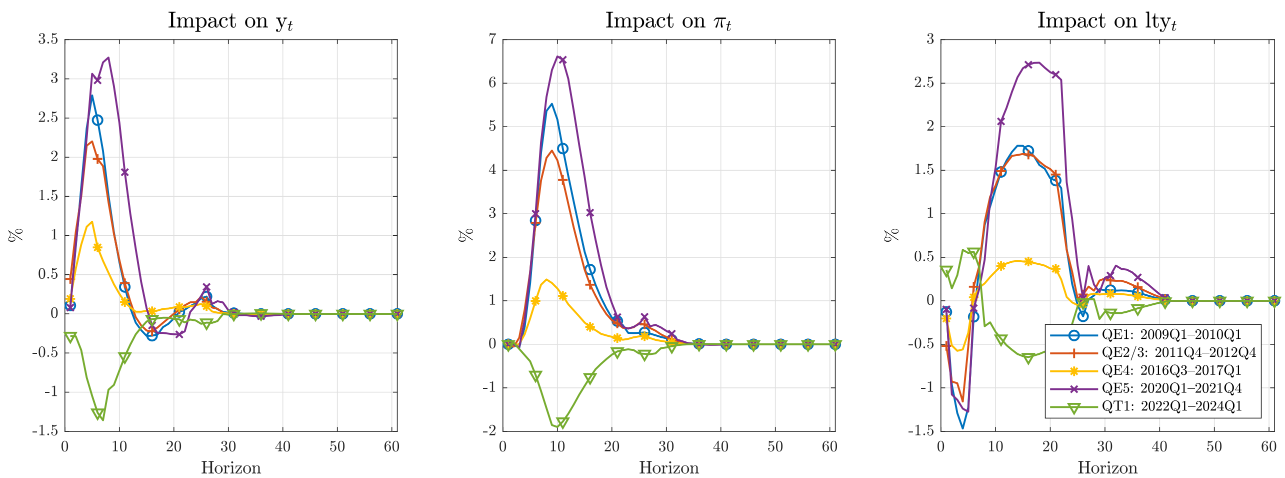

Figure G.2. Counterfactual experiments: Quantitative Easing/Tightening episodes
Notes: This figure plots the posterior median counterfactual paths of the output gap, y_t; CPI inflation, π_t; and long-term yields, lty_t across the four quantitative easing (QE) episodes, and the quantitative tightening (QT) episode in our sample. The counterfactual shows the impact that each round of QE (QT) has in isolation of all other shocks. This tells us the model-implied impact of each QE/QT episode. These are directly comparable to Figure 7 in the main text.

G.7 Threshold VAR Models

Here, we outline an $s = 2$ regime threshold VAR (TVAR) model using a standard Minnesota prior with p=2 lags and N=5. The TVAR is given by

$$\mathbf{Y}_t = \left\{ \mathbf{C}_{0,1} + \sum_{p=1}^{P} \mathbf{C}_{p,1}\mathbf{Y}_{t-p} + \mathbf{\Omega}_1^{\frac{1}{2}}\mathbf{e}_{t,1} \right\} \delta + \left\{ \mathbf{C}_{0,2} + \sum_{p=1}^{P} \mathbf{C}_{p,2}\mathbf{Y}_{t-p} + \mathbf{\Omega}_s^{\frac{1}{2}}\mathbf{e}_{t,2} \right\} (1-\delta) \tag{G.24}$$

$$\delta = 1, \text{ if } \mathbf{Y}_{i,t-d} \leq \gamma \tag{G.25}$$

where $\mathbf{Y}_t \equiv [\text{r}_t, \text{psc}_t, \text{lty}_t, \pi_t, y_t,]'$, is the same as in Equation G.1. $\mathbf{C}_0$ is an $N \times 1$ vector of intercepts and $\mathbf{C}_p$ are $N \times N$ coefficient matrices. δ is an indicator function which is equal to 1 when the threshold variable $\mathbf{Y}_{i,t-d} \leq \gamma$. We estimate both the delay parameter d, and the threshold value γ. We follow the estimation algorithm in Alessandri and Mumtaz (2017) to obtain draws from the posterior distributions of the coefficients and covariance matrices, and refer the reader to their paper for details.

G.7.1 TVAR Generalized Impulse Response Computation

To compute generalized impulse responses, we follow an algorithm similar to Koop et al. (1996). The impulse response function is defined as the difference between two conditional expectations, with and without exogenous shocks

$$\mathbf{IRF}_{t+j} = \mathbb{E}\left[y_{t+j}|\mathbf{v}_t,\, \mathcal{F}_t\right] - \mathbb{E}\left[y_{t+j}|\mathcal{F}_t\right] \tag{G.26}$$

where y_{t+j} contains forecasts of the endogenous variables at horizon $j = 1, ..., 20$, $\mathcal{F}_t$ represents the current information set and $\mathbf{v}_t$ is a vector of current disturbance terms. The information set with which the forecasts are conditioned on, contains the actual values of the lagged variables and a draw from the posterior distribution. We obtain the structural impact matrix using Algorithm 1 in Arias et al. (2018) using the same combination of sign and zero restrictions in Table 1. Generalized impulse response functions in the TVAR model allow for asymmetries in the size and sign of shock; as well as allowing for the possibility of movement from one regime to another.

G.7.2 Investigating Asymmetries from Generalized Impulse Response Functions using Threshold VAR Models

Table G.1. Threshold estimates from TVAR models (Alessandri and Mumtaz, 2017).
Notes: This table presents the threshold estimates for TVAR models. We consider each variable within the information set as the variable that governs the regime change; namely the Bank Rate, r_t; public sector funding from banks, psc_t; the long-term yield, lty_t; inflation, π_t; and the output gap, y_t. We also consider four exogenous variables to govern regime changes. These are: economic policy uncertainty (Baker et al., 2016), epu_t; financial stress, fsi_t; the budget surplus, surp_t, and the annualized volatility of quarterly stock returns, σ_t^R. Along side the posterior median threshold estimates, γ, we report the sample mean for each transition variable below.

$\mathbf{Y}_{i,t-d}$ =	r_{t-1}	psc_{t-1}	lty_{t-1}	π_{t-1}	y_{t-1}
γ=	8.21	1.51	6.71	5.83	0.34
Sample mean	6.38	1.35	7.08	4.93	0.08
$\mathbf{Y}_{i,t-d}$ =	epu_{t-1}	fsi_{t-1}	surp_{t-1}	σ_{t-1}^R	
γ=	87.04	0.15	1.10	48.88	
Sample Mean	86.97	0.14	-3.50	35.31	

Table G.1 reports the threshold estimates for TVAR models. We consider each variable within the information set as the variable that governs the regime change; namely the Bank Rate, r_t; public sector funding from banks, psc_t; the long-term yield, lty_t; inflation, π_t; and the output gap, y_t. We also consider four exogenous variables to govern regime changes. These are: economic policy uncertainty (Baker et al., 2016), epu_t (available from the economic policy website); financial stress, fsi_t (available from the Statistical Data Warehouse of the European Central Bank); the budget surplus, surp_t (available from the Bank of England's Millennium of Macroeconomic Database), and the annualized volatility of quarterly stock returns, σ_t^R (also available from the Bank of England's Millennium of Macroeconomic Database). Along side the posterior median threshold estimates, γ, we report the sample mean for each transition variable below. Overall, the majority of variables have threshold estimates close to their means. The only variables with an estimate of γ that deviate substantially from their means are the surplus, and annualized stock return volatility.

Figures G.3–G.4 show the regime specific generalized impulse response functions for economic variables with respect to a positive one-standard deviation bank funding shock and monetary policy shock, respectively. We report the posterior median impulse response functions and consider all endogenous variables as those that govern regime changes.

In addition, we also consider four exogenous variables as those that govern the regime changes. These are: economic policy uncertainty (Baker et al., 2016), epu_t; financial stress, fsi_t; the budget surplus, surp_t, and the annualized volatility of quarterly stock returns, σ_t^R. The plots provide little evidence in favor of asymmetries in the response of variables to such shocks across different regimes. We also investigate the sizes and signs of shocks. Our findings indicate that changing the size of the shock tends to scale up the impulse responses, and changing the sign of the shocks results in no difference in the absolute values of the impulse response functions; these are available upon request.

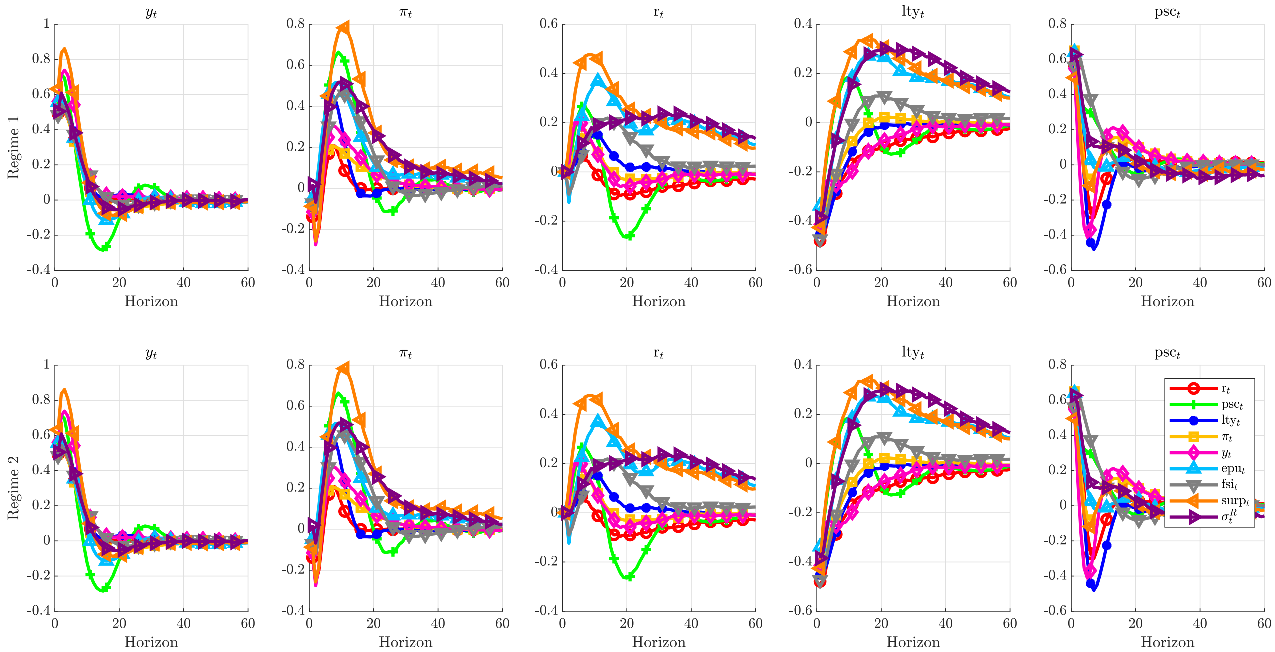

Figure G.3. Regime-specific impulse response functions for bank funding shocks $\mathbf{v^q}$

Notes: This figure reports impulse response functions of economic variables with respect to bank funding shocks. We consider each variable within the information set as the variable that governs the regime change; namely the Bank Rate, r_t; public sector funding from banks, psc_t; the long-term yield, lty_t; inflation, π_t; and the output gap, y_t. We also consider four exogenous variables to govern regime changes. These are: economic policy uncertainty (Baker et al., 2016), epu_t; financial stress, fsi_t; the budget surplus, surp_t, and the annualized volatility of quarterly stock returns, σ_t^R. y_t is the output gap π_t is inflation; r_t is the Bank Rate; and lty_t is the long-term yield. The first row corresponds to Regime 1 when $\mathbf{Y}_{i,t-d} \leq \gamma$ and the second row corresponds to Regime 2 when $\mathbf{Y}_{i,t-d} > \gamma$.

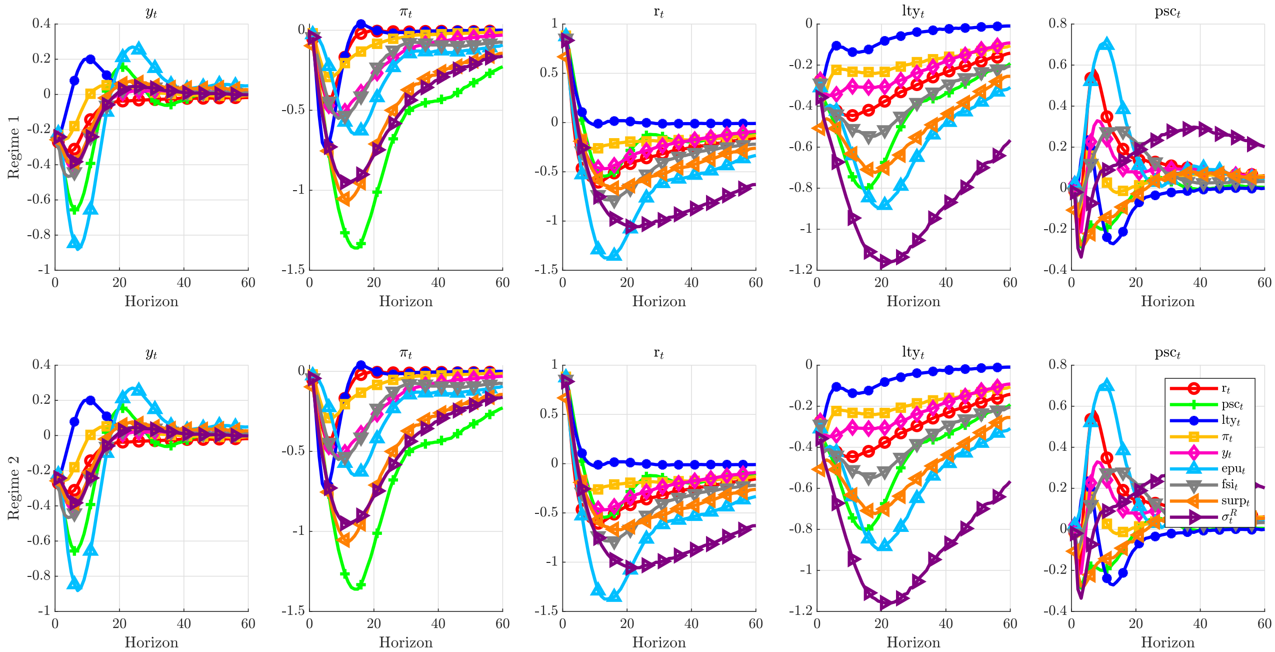

Figure G.4. Regime-specific impulse response functions for monetary policy shocks v^{mp}

Notes: This figure reports impulse response functions of economic variables with respect to monetary policy shocks. We consider each variable within the information set as the variable that governs the regime change; namely the Bank Rate, r_t; public sector funding from banks, psc_t; the long-term yield, lty_t; inflation, π_t; and the output gap, y_t. We also consider four exogenous variables to govern regime changes. These are: economic policy uncertainty (Baker et al., 2016), epu_t; financial stress, fsi_t; the budget surplus, $surp_t$, and the annualized volatility of quarterly stock returns, σ_t^R. y_t is the output gap π_t is inflation; r_t is the Bank Rate; and lty_t is the long-term yield. The first row corresponds to Regime 1 when $\mathbf{Y}_{i,t-d} \leq \gamma$ and the second row corresponds to Regime 2 when $\mathbf{Y}_{i,t-d} > \gamma$.