Econometric Analysis of Monetary Policy at the Zero Lower Bound*

Daisuke Ikeda†
Bank of Japan

Shangshang Li‡
University of Oxford

Sophocles Mavroeidis §
University of Oxford and INET

Francesco Zanetti ¶
University of Oxford

1 December 2019

Preliminary Draft

Abstract

The hypothesis that the effective lower bound on short term interest rates does not constrain the ability of monetary policy to achieve its objectives ('irrelevance hypothesis') is formalized using a New Keynesian dynamic stochastic general equilibrium model with quantitative easing and forward guidance. It is formally tested empirically using a corresponding structural vector autoregressive model estimated on postwar data for the US and Japan. The irrelevance hypothesis is strongly and robustly rejected for both countries. However, even though unconventional policies are not 100% as effective as conventional policies, a comparison of the impulse responses shows that unconventional policy has been quite effective in the US.

JEL Classification: E52, E58.

Keywords: Zero lower bound, unconventional monetary policy, monetary policy shocks.

*We would like to thank Hibiki Ichiue, Moto Shintani and Nao Sudo for comments and suggestions. Views expressed in the paper are those of the authors and do not reflect the official views of the Bank of Japan.

†Bank of Japan, daisuke.ikeda@boj.or.jp.

‡University of Oxford, Department of Economics: shangshang.li@economics.ox.ac.uk

§University of Oxford, Department of Economics: sophocles.mavroeidis@economics.ox.ac.uk.

¶University of Oxford, Department of Economics: francesco.zanetti@economics.ox.ac.uk.

1 Introduction

Adjustments in the overnight nominal interest rate have been the primary tool for the implementation of monetary policy since the early 1980s. In recent years, however, the short-term nominal interest rate reached the zero lower bound (ZLB) in several countries, making the standard policy tool *de facto* ineffective. Two prominent examples are Japan that reached the ZLB since the domestic financial crisis of 1997-1998, and the United States that reached the ZLB in the aftermath of the global financial crisis of 2007-2008. The central banks in these countries countervailed the inapplicability of the standard policy tool by embarking on unconventional policies that involve central bank purchasing of government bonds, and use of forward guidance to signal future policy action.[1]

This paper studies two critical issues of monetary policy at the ZLB. First, it sheds light on the debate on whether the ZLB restricts the efficacy of monetary policy, thus representing an important constraint on what monetary policy can achieve, as argued by Eggertsson and Woodford (2003), or whether unconventional policies are fully effective in circumventing the ZLB constraint, as argued by Swanson and Williams (2014) and Debortoli et al. (2019). Second, it uses the identifying power of the ZLB, shown by Mavroeidis (2019), combined with theoretically-motivated sign restrictions on the impulse responses to monetary policy shocks, to assess the efficacy of unconventional policies.

We tackle these two issues by developing theoretical and empirical models. The theoretical model is a New Keynesian dynamic stochastic general equilibrium (DSGE) model with heterogeneous investors that sheds light on potential transmission mechanisms of unconventional policies, and motivates our empirical analysis. The model accounts for unconventional policy in the form of (i) quantitative easing (QE) implemented by a long-term government bond purchase program that directly affects long-term government bond yields when the ZLB holds, and (ii) forward guidance (FG) under which the central bank commits to keeping short-term interest rates low in the future. Under the mild assumption that the central bank continues to use inflation as the key indicator to guide the policy stance during ZLB periods, the impact of QE and FG is controlled by some key model parameters. By varying each of these parameters across admissible values, we can map the range of feasible impacts of unconventional policies. Our model shows that unconventional policies entail wide de-

[1] See Christensen and Rudebusch (2012) and Liu et al. (2019) for the US, and Ugai (2007), and Bank of Japan (2016) for Japan. Ueda (2012) provides a comparison of monetary policy between the US and Japan.

grees of effectiveness, ranging from fully effective, such that unconventional policies retain the same effectiveness as the conventional policy based on adjustments in the short-term interest rate as if there were no ZLB, to ineffective, such that the ZLB fully constrains the efficacy of monetary policy.[2] The theoretical model also provides a shadow policy rate, which can be thought of as an indicator of the desired monetary policy stance, and motivates the formal tests of the hypothesis that the ZLB is irrelevant that we use later on in our empirical analysis.

Our empirical analysis is based on the methodology developed by Mavroeidis (2019) for the identification and estimation of structural vector autoregressive (SVAR) models that include variables subject to occasionally binding constraints. This methodology provides a flexible framework to assess the overall effectiveness of unconventional policy and formally test the hypothesis that the ZLB is empirically irrelevant, as recently formulated by Debortoli et al. (2019). Specifically, it allows for a shadow rate – the short-term interest rate that the central bank would set if there were no ZLB, and identifies a measure of the overall efficacy of unconventional policy. Identification does not rely on any particular theoretical assumptions, and is therefore more agnostic than a typical DSGE model like the one we use here. This increased generality/robustness comes at the cost of being unable to disentangle the effects of different unconventional policies, such as QE versus FG. Instead, the SVAR enables us to estimate the overall effect of QE and FG. In addition, it allows us to test whether unconventional policy is fully effective, making the ZLB irrelevant from a policy perspective.

Our empirical results can be summarized as follows. We derive two alternative tests of the hypothesis that the ZLB is empirically irrelevant. In both cases, under the null hypothesis of ZLB irrelevance, the impulse responses to any shock will not depend on whether interest rates are at the ZLB or above. The first test asks whether once we include long-term interest rates and possibly measures of the money supply in a monetary policy SVAR, the short-term interest rate can be excluded from the model. This is a necessary condition for the resulting impulse responses to be constant across ZLB and non-ZLB regimes, for,

[2] The effect of QE in our model operates through the government bond purchase program, and it abstracts from several possible channels outlined in Krishnamurthy and Vissing-Jorgensen (2011). At present, there is no comprehensive model that embeds the several theoretical propagation channels of QE. As stated in Bernanke (2014): "The problem with QE is that it works in practice, but it doesn't work in theory." Thus, our approach uses the insights from theory to develop an empirical model to assess the impact of unconventional policies.

otherwise, the dynamics of the data would necessarily change when short-term rates hit their lower bound. The second test asks whether the shadow rate itself is sufficient in capturing both conventional and unconventional policy, i.e., whether, once we include shadow rates in the VAR, current and lagged values of observed short rates should become redundant. In other words, the dynamics of the economy can be characterized by a standard SVAR in the shadow rates (which are equal to observed rates above the ZLB), so the ZLB represents simply an econometric problem (pure censoring of one of the variables) and does not change the dynamics of the economy. Both of those tests can be formulated as likelihood ratio tests in the framework of Mavroeidis (2019).

We apply the above tests to SVAR models estimated on postwar data for the US and Japan. We consider many different specifications, varying the lag order of the VAR, varying the estimation sample (to account for structural change), or using different measures of the variables in the model. In all cases, the hypothesis that the ZLB has been empirically irrelevant is overwhelmingly rejected for both countries. The conclusion is therefore fairly robust: the ZLB does represent a constraint on what monetary policy can achieve in those countries.[3] However, a statistical rejection of the irrelevance hypothesis may not be particularly important economically if it turns out that unconventional policy is almost as effective as conventional policy. It also says little about the difference in the dynamic effects of conventional and unconventional policies. This is the question we turn to next.

We identify the dynamic effects of conventional and unconventional policy by combining the identifying implications of the ZLB shown by Mavroeidis (2019) with additional sign restrictions on the impulse responses to a monetary policy shock. In particular, Mavroeidis (2019) shows that the ZLB acts as a quasi-experiment that partially identifies the impulse responses to monetary policy shocks. The intuition is that if the ZLB does constrain policy, it will cause the dynamics of the data to change across regimes, and, with appropriate correction for endogeneity, this change can be used as an 'instrument' to identify the transmission mechanism of policy. The identified set based only on the ZLB turns out to be fairly wide, so we employ the following sign restrictions: a negative monetary policy shock should have a nonnegative effect on inflation and output and a nonpositive effect on the policy rate over

[3]This evidence is corroborates Eggertsson and Woodford's (2003) claim that "the zero bound does represent an important constraint on what monetary stabilization policy can achieve", and is consistent with the findings in Gust et al. (2017) and Del Negro et al. (2017), who attribute an important role to the ZLB for the decline in output during the financial crisis.

the first four quarters.[4] Use of these sign restrictions drastically sharpens the identified set of impulse responses. We find that unconventional monetary policy is roughly 75% as effective as conventional policy on impact, and remains less effective than conventional policy beyond 8 quarters, but it is as effective as conventional policy in the medium term of 2-8 quarters. We therefore conclude that, even if not 100% as effective as conventional policy, unconventional policy overall has been quite effective in the US. We do not yet have the corresponding results for Japan [in progress].

Our analysis is closely related to two strands of research. The first strand of literature pertains to theoretical studies that investigate the transmission mechanism of unconventional monetary policy. Among those, regarding QE, our theoretical model is close in spirit to Andrés et al. (2004), Chen et al. (2012), Harrison (2012), Gertler and Karadi (2013), and Liu et al. (2019). These studies use heterogeneous preferences for assets of different maturities and limit arbitrage across assets to break the irrelevance of QE, as originally discussed in Eggertsson and Woodford (2003). Regarding FG, our model follows Reifschneider and Williams (2000), and it considers this mechanism in a general equilibrium model that directly accounts for purchasing of long-term bonds. We then use the theoretical results to setup the empirical VAR model to identify the effect of unconventional policy. Our main contribution to this first strand of literature is to develop a simple theoretical framework that sheds light on an extensive range of potential effects of unconventional policy.

The second strand of literature pertains to empirical studies that estimate the effectiveness of unconventional policy. It includes Swanson and Williams (2014), Debortoli et al. (2019), who use SVARs to investigate the (ir)relevance of the ZLB constraint by comparing impulse responses to shocks between normal times and ZLB episodes. These studies do not include short-term interest rates in their SVARs. Another important study is by Inoue and Rossi (2018), who use SVAR with shocks to the entire yield curve and report evidence that unconventional monetary policy has been effective in the US. Our empirical methodology is closely related to Hayashi and Koeda (2019), who propose a SVAR model for Japan that includes short rates and takes into account the ZLB. Our methodology relaxes some of the limitations of Hayashi and Koeda (2019) which are important for our objectives. In particular, our methodology does not impose recursive identification of the SVAR that puts interest rates last in the causal ordering, thus setting the contemporaneous effect of monetary policy

[4] These restrictions were used in Debortoli et al. (2019).

shocks on inflation and output to zero, and would therefore put the empirical model at odds with any forward-looking DSGE model. It also crucially allows for lags of the shadow rate in the dynamics, which enables the model to both nest the FG rule of Reifschneider and Williams (2000) and to nest the case of the ZLB irrelevance, thus providing a simple test of that hypothesis.

The structure of the paper is as follows. Section 2 develops a simple DSGE model of unconventional monetary policy that accounts for a government long-term asset purchase program and forward guidance. Section 3 develops a corresponding structural VAR model that will be used in the empirical analysis. Section 4 describes the data. Section 5 reports the empirical results. Section 6 concludes. Appendices provide supporting material on the derivation of the DSGE model and data description.

2 A macroeconomic model of UMP

In this section, we present a simple macroeconomic model of unconventional monetary policy (UMP) to motivate our econometric methodology, which we develop in Section 3. The theoretical framework provides a key connection with the econometric methodology regarding the efficacy of UMP, and it lays out important channels for the propagation mechanism and the effectiveness of UMP. The model features both FG and QE as UMP. FG is modeled as a monetary policy rule, following Reifschneider and Williams (2000), under which the central bank commits to maintain an interest rate lower than the normal level implied by a Taylor rule. QE is activated when the economy is at the ZLB, following Chen et al. (2012) in which market segmentations add mechanisms through which a central bank's asset purchases have real effects on the economy. The central equations for our analysis are presented in subsection 2.1. In subsection 2.2, the model is simulated to study how UMP works in the model. The derivations of equations and the detail of model simulations are reported in Appendix A.

2.1 Central equations

The model assumes bond market segmentation and preferred habitat theory originally proposed by Modigliani and Sutch (1966), for which imperfect substitutability across financial assets makes yields of long-term government bonds higher than short-term government bonds. The economy consists of households, firms, and a central bank. The household sector comprises of two types of households: unrestricted households that trade short- and long-term government bonds and constrained households that can trade long-term government bonds only. The firm sector and the central bank are fairly standard as in a typical New Keynesian model.

Monetary policy rules. The nominal interest rate i_t set by the central bank is bounded below by zero,

$$i_t = \max\{i_t^*, 0\} \tag{1}$$

where i_t^* is the central banks' target rate (or the shadow rate). The central bank sets i_t^* according to the rule:

$$i_t^* = i_t^{\text{Taylor}} - \alpha Z_t, \quad Z_t = \rho_Z Z_{t-1} + \left(i_t - i_t^{\text{Taylor}}\right) \tag{2}$$

where Z_t measures (cumulative) deviations of the actual interest rate i_t from the target or desired rate i_t^*. This is a slight generalization of the FG rule in Reifschneider and Williams (2000), who set $\rho_Z = 1$. For analytical tractability, and because of the presence of the interest rate smoothing parameter ρ_i, we will set $\rho_Z = 0$ in the rest of this section (but note that we place no restrictions on the FG rule in the empirical analysis below). The shadow rate i_t^* consists of two parts: i_t^{Taylor} and αZ_t. First, i_t^{Taylor} is the Taylor-rule-based rate that responds to inflation π_t and output y_t:

$$i_t^{\text{Taylor}} - i = \rho_i \left(i_{t-1}^* - i\right) + (1-\rho_i)\left[r_\pi \log(\pi_t/\pi) + r_y \log(y_t/y)\right] + \epsilon_t^i, \tag{3}$$

where $\epsilon_t^i \sim$ i.i.d. $N(0, \sigma_i^2)$ is a monetary policy shock and variables without subscripts denote those in steady state.

Second, αZ_t in equation (2) encapsulates the strength of FG. A positive value for α will maintain the target rate i_t^* below the Talyor-rule-implied rate i_t^{Taylor}. Under the ZLB of $i_t = 0$, the more the central bank has missed to set the interest rate at its Taylor-rule-based rate, the lower the central bank sets its target rate i_t^* through equation (2) as long as the Taylor-implied rate is persistent, i.e. $\rho_i > 0$ in equation (3). Specifically, given $\rho_i > 0$, the degree of the interest rate deviated from the Taylor-rule-based rate, $Z_t = i_t - i_t^{\text{Taylor}}$, is multiplied by the parameter of FG, α, and the multiple αZ_t decreases the shadow rate today i_t^* through equation (2), which in turn decreases the shadow rate in the next period, i_{t+1}^* through the effect of the past shadow rate on the current Taylor rule rate in equation

(3). In this way the central bank commits to a longer duration of the ZLB. We focus on the parameter α as FG.[5]

The central bank activates QE – long-term government bond purchases – once the economy hits the ZLB of the interest rate i_t. As in the case of positive interest rates, under the ZLB the central bank uses the shadow rate as policy guidance. Specifically, QE linearly depends on the shadow rate, and as a result the amount of long-term government bonds $b_{L,t}$ held by the private agents is given by:

$$\widehat{b}_{L,t} = \begin{cases} 0 & \text{if } i_t^* > 0 \\ \gamma \frac{i_t^*}{1+i} & \text{if } i_t^* \leq 0 \end{cases}, \tag{5}$$

where a variable with hat denotes a deviation from steady state. This QE rule implies that asset purchases by the central bank is zero (relative to the steady state) when the zero lower bound is not binding (i.e. $i_t = i_t^* > 0$) and, given $\gamma > 0$, such purchases are positive (i.e. $\hat{b}_{L,t} < 0$) when the shadow rate goes below negative zero (i.e. $i_t^* < 0$).

Euler equation. A premium between long-term and short-term government bond yields is assumed to depend on the amount of long-term bonds held by the private agents. The bond market is segmented by preferred habitat households: restricted households trade only long-term bonds, while unrestricted households trade both short- and long-term government bonds but subject to a transaction cost for each unit of traded long-term bonds. QE can affect output through its effects on the amount of long-term bonds held by the restricted and unrestricted households, the premium, and thereby the long-term interest rate. Combining the first-order conditions of the problems of the two types of households and the good-market clearing condition yields our central equation:

$$\underbrace{\hat{y}_t = E_t \hat{y}_{t+1} - \frac{1}{\sigma}\left(\hat{\imath}_t - E_t \hat{\pi}_{t+1}\right) - \chi_z z_t^b}_{\text{Standard Euler equation}} \underbrace{-\chi_b \widehat{b}_{L,t}}_{\text{Effect of QE}}, \tag{6}$$

where $\hat{y}_t$ is output, $\hat{\imath}_t \equiv \frac{i_t - i}{1+i}$ is a short-term net nominal interest rate, $\hat{\pi}_{t+1}$ is inflation, and z_t^b is the demand (preference) shock. The coefficients, $\chi_b, \chi_z > 0$, are derived as a function

[5] Debortoli et al. (2019) consider the case of $\alpha = 0$ and interpret ρ_i – the interest rate smoothing coefficient in the Taylor rule equation (3) – as FG when i_t^* is below the ZLB. Even with $\alpha = 0$, smoothing on the shadow rate implies a longer duration of the ZLB than what would be implied by a Taylor rule that smooths on the actual interest rate:

$$i_t^{\text{Taylor}} - i = \rho_i \left(i_{t-1} - i\right) + \left(1 - \rho_i\right)\left[r_\pi \log\left(\pi_t/\pi\right) + r_y \log\left(y_t/y\right)\right] + \epsilon_t^i. \tag{4}$$

The two rules are obviously equivalent when $i_{t-1} = i_{r-1}^* > 0$.

of the structural parameters in the model. Specifically, the coefficient χ_b on the amounts of long-term government bonds $\widehat{b}_{L,t}$ depends on a parameter that governs the elasticity of the premium with respect to the amount of long-term government bonds and a fraction of restricted households among others. Equation (6) shows that unlike the standard Euler equation, QE – an exogenous purchase of the long-term bonds by a central bank (i.e., an exogenous decrease in $\hat{b}_{L,t}$) – can have a positive effect on output $\hat{y}_t$.

The Euler equation (6) can be further simplified as follows. Substituting the QE rule (5) into the Euler equation (6) yields

$$\hat{y}_t = \begin{cases} E_t\hat{y}_{t+1} - \frac{1}{\sigma}\left(\frac{i_t - i}{1+i} - E_t\hat{\pi}_{t+1}\right) - \chi_z z_t^b & \text{if } i_t^* > 0 \\ E_t\hat{y}_{t+1} - \frac{1}{\sigma}\left(\sigma\chi_b\gamma\frac{i_t^*}{1+i} - \frac{i}{1+i} - E_t\hat{\pi}_{t+1}\right) - \chi_z z_t^b & \text{if } i_t^* \leq 0 \end{cases}.$$

Because the interest rate i_t is bounded below zero as in equation (1), these two equations can be written into a single equation as:

$$\hat{y}_t = E_t\hat{y}_{t+1} - \frac{1}{\sigma}\left(\left(1-\lambda^*\right)\hat{\imath}_t + \lambda^*\hat{\imath}_t^* - E_t\hat{\pi}_{t+1}\right) - \chi_z z_t^b, \tag{7}$$

where $\lambda^* \equiv \sigma\chi_b\gamma$ and $\hat{\imath}_t^* \equiv \frac{i_t^* - i}{1+i}$. Equation (7) expresses the efficacy of QE by the single parameter λ^*, which we will use to interpret results in the empirical exercise. Importantly, λ^* is the product of χ_b, which captures the efficacy of the asset purchase policy per unit of long-term bonds purchased (in terms of its impact on $\hat{y}_t$), and γ, which captures how strongly the central bank launches the asset purchase program in response to a decrease in the shadow rate under the ZLB.

Phillips curve. Under the standard assumptions of monopolistic competition and Calvo (1983) pricing, the following Phillips curve can be derived from the problem of the firm sector:

$$\hat{\pi}_t = \beta E_t\hat{\pi}_{t+1} + \kappa\hat{y}_t - \chi_a z_t^a, \tag{8}$$

where z_t^a is a technology shock, $0 < \beta < 1$ is a preference discount factor, and parameters $\kappa, \chi_a > 0$ depend on the structural parameters of the model.

To summarize Section 2.1, given exogenous shocks z_t^a, z_t^b, and ϵ_t^i, the system of equations for this economy consists of four equations (1), (2), (7), and (8) with the same number of unknowns $\hat{\imath}_t$, $\hat{\imath}_t^*$, $\hat{y}_t$, and $\hat{\pi}_t$, where i_t^{Taylor} is substituted out using equation (3).

2.2 Model simulations

In this section, we study the transmission mechanism of QE and FG under the ZLB. Since the model cannot be solved analytically under the ZLB, we calibrate it and use the piecewise linear solution method proposed by Guerrieri and Iacoviello (2015).[6] Detailed description of the calibration and solution is reported in Appendix A.2.

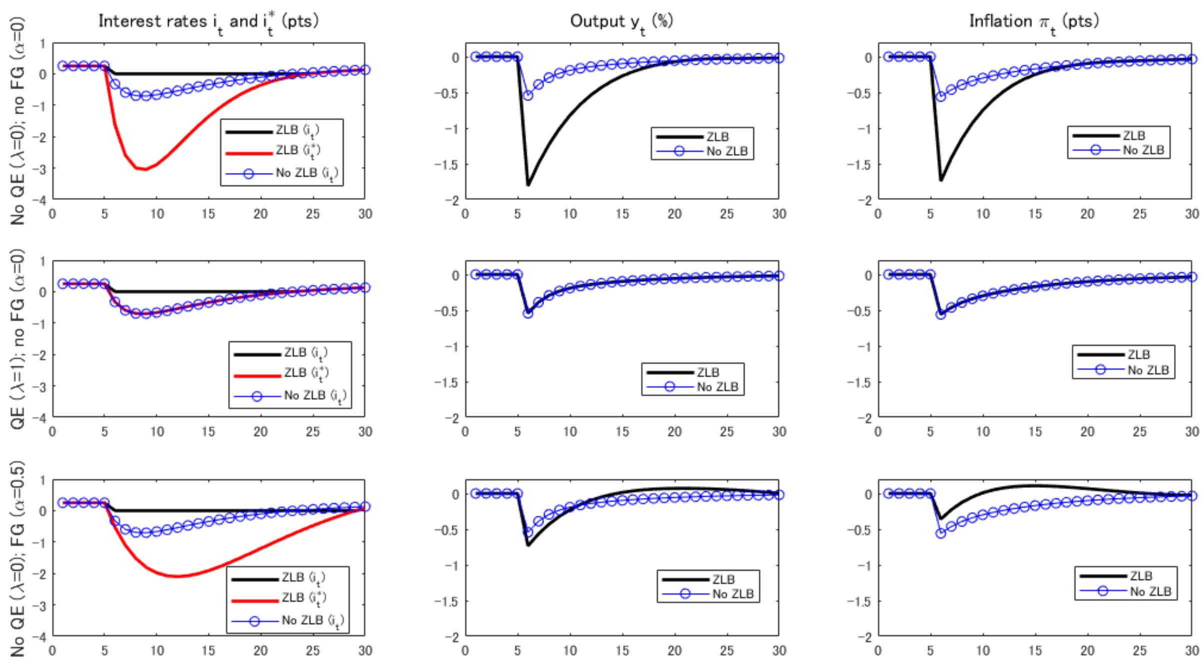
Figure 1 shows simulated paths for the model with the ZLB (denoted as "ZLB") and the model without the ZLB (denoted as "No ZLB") in which the interest rate equation (1) is replaced by $i_t = i_t^*$. For each model, the economy starts from the steady state and is hit by a severe negative demand shock in period $t = 6$. In the case of no QE ($\lambda^* = 0$) and no FG ($\alpha = 0$), which is shown in the top panels, the negative demand shock drives the economy into the ZLB and causes a severe recession by decreasing output and inflation sharply. At the ZLB the interest rate i_t cannot be lowered in response to a fall in inflation. This raises the real interest rate, decreases consumption and output, and puts further downward pressure on inflation through the Phillips curve (8). Due to this negative feedback effect, decreases in output and inflation become much severer under the ZLB (thick black lines in the top panels) than the case of no ZLB (blue circled lines).

Quantitative easing. Quantitative easing can offset the negative impact of the ZLB perfectly. When QE has a full impact, i.e. when $\lambda^* = 1$, although the interest rate i_t is stuck at zero, output and inflation follow the same paths as in the case of no ZLB, as shown in the medium panels of Figure 1. In response to a decrease in the shadow rate i_t^*, the central bank increases the purchase of long-term government bonds and by doing so it lowers the long-term government bond yield by compressing its premium, which boosts consumption and output. When $\lambda^* = 1$, this QE perfectly cancels the negative impact of the ZLB. As shown in the Euler equation (7), the interest rate i_t disappears from the Euler equation when $\lambda^* = 1$, and under such a QE policy, the economy evolves as if there were no ZLB. In other words, the ZLB becomes irrelevant when $\lambda^* = 1$.

Forward guidance. FG can be also effective to address the ZLB problem. With FG activated as $\alpha = 0.5$ but with no QE ($\lambda^* = 0$), the interest rate i_t is kept at zero longer as the shadow rate i_t^* is kept negative longer (bottom left panel of Figure 1) than the case of no

[6] It is worth noting that the simulations shown here are not intended to derive quantitative implications.

Figure 1: Quantitative easing and forward guidance



Notes: The figure shows the dynamic path of the economy where a severe demand shock hits the economy in period $t = 6$. Interest rates and inflation are shown in terms of annual percentage points (pts) and output is in terms of deviation from steady state (%). "ZLB" represents the model with the ZLB and "No ZLB" represents the model without the ZLB, where the interest rate equation (1) is replaced by $i_t = i^*_t$.

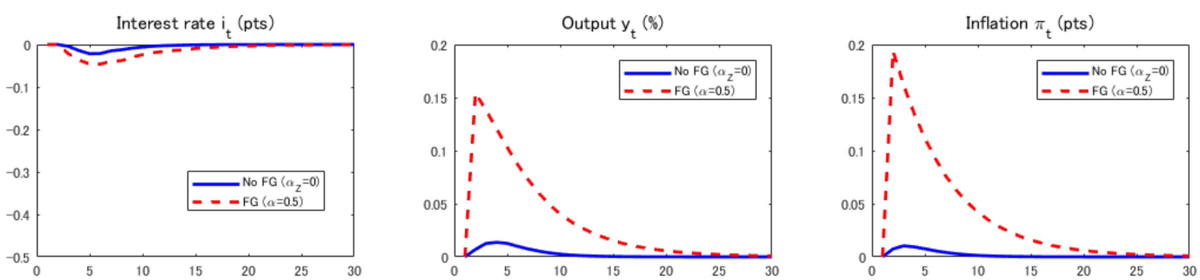
FG (top left panel of the figure). Such a credible promise stimulates output and inflation in future, which in turn affect the current output and inflation through expectations embedded in the Euler equation (7) and the Phillips curve (8). Output becomes close to and inflation becomes even higher than what would be achieved in the economy with no ZLB (bottom medium and right panels). FG appears to be extremely powerful for addressing the ZLB problem, which is typical in the standard New Keynesian model and is known as the "forward guidance puzzle."

Monetary policy shocks at the ZLB. With FG a negative monetary policy shock, $\epsilon^i_t < 0$, may affect the economy even under the ZLB. Figure 2 plots impulse responses to a 1 annual percentage point decrease in the monetary policy shock for the model with no FG ($\alpha = 0$) and the model with FG ($\alpha = 0.5$). When the monetary policy shock hits in period $t = 2$, the ZLB is already binding, which is caused by a severe negative demand shock in period $t = 1$. The impulse responses in Figure 2 are computed in a similar manner as in the econometric analysis in Section 5. That is, disturbances to the demand, technology, and monetary policy shocks are generated randomly for period $t = 2$ onward except for a monetary policy shock in period $t = 2$, and for each case with and without the monetary policy shock in period $t = 2$ the dynamic paths are computed; this process is repeated for 1000 times and the responses in Figure 2 is given by a difference between the average of the 1000 paths with the monetary shock and that without it.

Without FG the monetary policy shock has little effect on output and inflation as shown in Figure 2. This is because there is almost no room for the interest rate i_t to go down for almost all realizations of random disturbances.[7] With FG, however, the monetary policy shock has significant impacts on output and inflation, as shown in Figure 2. This is because a decrease in i_t^{Taylor} increases $Z_t = i_t - i_t^{\text{Taylor}}$ under the ZLB, which in turn decreases the target rate i_t^* and i_t^{Taylor} in the next period when FG is put in place ($\alpha > 0$). Hence, with FG a monetary policy shock acts as a shock to FG under the ZLB.

To summarize, the theoretical model shows that QE and FG can be powerful to stimulate the economy at the ZLB and, in principle, they can deliver the same effectiveness of monetary policy in periods outside the ZLB.

Figure 2: Impulse responses to a monetary policy shock at the ZLB



Notes: The impulse responses are computed as a difference between two dynamic responses: one with a 1% point decrease in the interest rate in period $t = 2$ and another with no such a shock. Each response is the average of 1000 responses with disturbances to the demand, technology, monetary policy shocks generated randomly for periods $t = 2, 3, .., 30$, except for a monetary policy shock in period $t = 2$. In period $t = 1$, there is a demand shock only; a severe negative demand shock hits and drives the economy into the ZLB.

[7] Above the ZLB, the same monetary policy shock would decrease the interest rate by more than 0.75 percentage points. The reason why it is still less than the magnitude of the monetary shock of a 1 percentage point is due to the persistence of and the endogenous response of i_t^{Taylor} to inflation and output.

3 Empirical model

Our empirical analysis will not rely on a specific DSGE model like the one introduced earlier but on an agnostic SVAR in which the short-term interest rate is subject to an effective lower bound (ELB). The relevant econometric framework is the censored and kinked SVAR (CKSVAR) developed by Mavroeidis (2019), which is described by the following equations:

$$Y_{1t} = \beta \left(\lambda Y_{2t}^{*} + (1-\lambda) Y_{2t}\right) + B_1 X_t + B_{12}^{*} X_{2t}^{*} + \varepsilon_{1t}, \tag{9}$$

$$Y_{2t}^{*} = -\alpha Y_{2t} + (1+\alpha)\left(\gamma Y_{1t} + B_2 X_t + B_{22}^{*} X_{2t}^{*} + \varepsilon_{2t}\right), \tag{10}$$

$$Y_{2t} = \max\left\{Y_{2t}^{*}, b_t\right\}, \tag{11}$$

where $Y_t = (Y_{1t}', Y_{2t})'$ are the endogenous variables, partitioned such that Y_{1t} are unconstrained and Y_{2t} is constrained, X_t comprises exogenous and predetermined variables, including lags of Y_t, X_{2t}^{*} consists of lags of Y_{2t}^{*}, ε_t are iid shocks with $\varepsilon_{1t} \perp\!\!\!\perp \varepsilon_{2t}$, b_t is an observable lower bound, and $Y_{2t}^{*} < b_t$ is unobservable. The 'latent' variable or the shadow rate Y_{2t}^{*} represents the *desired* policy stance, as opposed to the *effective* policy stance, e.g., in Wu and Xia (2016), except in the special case $\alpha = 0$ and $\lambda = 1$. When $Y_{2t}^{*} < b_t$, it represents UMP, such as QE or FG, which are not modeled explicitly.

Equation (10) nests the FG rule of Reifschneider and Williams (2000) introduced in (2) above, and the parameter α has exactly the same interpretation as in the theoretical model of the previous section: a larger α means interest rates will stay longer at the ELB. As explained in the previous section, the degree of FG is not determined solely by α. It also depends on the coefficients on the lags of the shadow rate in the policy rule, B_{22}^{*}, see also the discussion on interest rate smoothing, ρ_i, in the Taylor rule (3).[8]

The parameter λ partially characterizes the efficacy of UMP relative to conventional policy *on impact*. Specifically, from equation (9) we see that above the ZLB (i.e., when $Y_{2t} = Y_{2t}^{*} > b_t$), the contemporaneous marginal effect of a change in the short-term interest rate Y_{2t} on Y_{1t} is β, but the corresponding effect at the ZLB, driven by a change in Y_{2t}^{*}, is $\lambda\beta$. When $\lambda = 1$, the two effects are equal, while $\lambda = 0$ corresponds to the case in which UMP has no contemporaneous effect on Y_{1t}. Note that the model can also allow for policy reversal (cf. Brunnermeier and Koby, 2018) when $\lambda < 0$. Note that λ is intentionally different from the parameter λ^{*} that we introduced in the theoretical model of the previous section

[8] Note that the dynamics of the policy rule in equation (10) are completely unrestricted, in contrast to the specifications of the Taylor rule in the previous section, e.g., equations (1) and (2), where lags of Y_t are excluded. Thus, we are not relying on any short-run exclusion restrictions for identification.

to characterize the efficacy of QE only. Here, λ is intended to capture the combined impact effect of UMP, it is not restricted to QE.

However, the parameter λ does not suffice to pin down the impulse response to a monetary policy shock ε_{2t} at the ZLB. Specifically, the impulse response to a unit shock to ε_{2t} is equal to $\beta/\left(1-\gamma\beta\right)$ above the ZLB, and $\xi\beta/\left(1-\xi\gamma\beta\right)$ at ZLB, where

$$\xi = \lambda\left(1+\alpha\right). \tag{12}$$

So, it is, in fact, ξ, not λ, that measures the efficacy of an UMP shock. To see why, consider an example of a situation in which an UMP shock would have been only 50% as effective if it had the same magnitude as a conventional MP shock ($\lambda = 0.5$), but the monetary authority reacts twice as much at the ZLB than in conventional times ($\alpha = 1$), so the actual UMP shock is twice as big as the conventional shock ε_{2t}. Then the observed impact of such an UMP shock will be the same as the corresponding conventional policy shock, see Mavroeidis (2019) for further discussion.

The methodology for the identification and estimation of the CKSVAR is given in Mavroeidis (2019), where it is shown the model is generally under-identified, but that the parameter ξ, defined in (12), as well as the impulse responses to the monetary policy shock ε_{2t} are partially identified in general. To gain some intuition for this result, it is useful to write down the solution (reduced-form) of equations (9)-(11) for Y_{1t}, taken from (Mavroeidis, 2019, Proposition 2):

$$Y_{1t} = C_1 X_t + C_{12}^{*} X_{2t}^{*} + u_{1t} - \widetilde{\beta} D_t\left(C_2 X_t + C_{22}^{*} X_{2t}^{*} + u_{2t} - b_t\right) \tag{13}$$

where $D_t := 1_{\{Y_{2t}=b_t\}}$ is the indicator of the ZLB regime, the matrices $C_1, C_{12}^{*}, C_2, C_{22}^{*}$ are reduced-form coefficients, $u_t = \binom{u_{1t}}{u_{2t}}$ are reduced-form errors, and $\Omega = var\left(u_t\right)$. The reduced-form equation (13) is an 'incidentally kinked' regression, whose coefficients change across regimes. The coefficient of the kink $\widetilde{\beta}$ is identified, together with the remaining reduced-form parameters.[9] In other words, we can infer from the data whether the slope coefficients change across regimes by testing whether $\widetilde{\beta} = 0$. However, the parameter $\widetilde{\beta}$ does not have a structural interpretation but relates to the underlying structural parameters through the

[9] C_2, C_{22}^{*} and $var\left(u_{2t}\right)$ are identified from the reduced form equation for Y_{2t}, which is a dynamic Tobit regression, see (Mavroeidis, 2019, Proposition 2).

equations

$$\widetilde{\beta} = (1 - \xi)\,(I - \xi\beta\gamma)^{-1}\,\beta, \tag{14}$$

$$\gamma = (\Omega_{12}' - \Omega_{22}\beta')\,(\Omega_{11} - \Omega_{12}\beta')^{-1}\,. \tag{15}$$

Mavroeidis (2019) shows that the structural parameters ξ, β and γ are generally only partially identified, in the sense that there is a set of values of ξ, β and γ that correspond to any given value of the reduced form parameters $\widetilde{\beta}, \Omega$ – this is the set of solutions of (14) and (15). This set reduces to a point in the special case $\xi = 0$, when $\beta = \widetilde{\beta}$. In our empirical analysis below, we will use sign restrictions on the IRFs to further sharpen the identified set on the efficacy of UMP parameter ξ.

The previous discussion about the interpretation of the parameter ξ concerned the relative efficacy of UMP *on impact*. The *dynamic* effects of UMP on Y_1 are governed by the coefficients on the lags of the shadow rate B_{12}^*. For example, the case in which UMP is completely ineffective at all horizons can be represented by the restrictions $\lambda = 0$ and $B_{12}^* = 0$. A more restrictive assumption is that UMP has no effect on the conventional policy instrument Y_{2t} either, i.e., that any FG or QE is completely ineffective in changing the path of short-term interest rates as well. This can be represented by the special case $\lambda = \alpha = 0$, $B_{12}^* = 0$ and $B_{22}^* = 0$, so that the shadow rate completely drops out of the right-hand side of the equations (9)-(11). This special case is called a *kinked* SVAR (KSVAR) by (Mavroeidis, 2019, Proposition 2). The absence of latent regressors in the likelihood function makes the KSVAR much easier to estimate than the CKSVAR.

An alternative interpretation of the KSVAR is as follows. Suppose that Y_{1t} includes also variables that can capture unconventional as well as conventional monetary policy, such as money supply or long-term interest rates. If such variables can adequately characterize UMP, then the shadow rate Y_{2t}^* would be redundant in the model. Moreover, if UMP is 100% as effective as conventional policy, then the short-term interest rate, Y_{2t}, should be redundant in the subsystem of the SVAR that corresponds to Y_{1t}, for otherwise the dynamics of Y_{1t}, and hence the impulse response, would differ across regimes. In the previous section, we showed theoretically that this would be the case when $\lambda^* = 1$ and $\alpha = 0$, in the sense that the solution of the DSGE model can be equivalently written as a VAR in $\left(\pi_t, y_t, i_t^L\right)'$, where i_t^L is the yield on long-term bonds, with no role for the short rate i_t. This also underlies the approach in Swanson and Williams (2014) and Debortoli et al. (2019), whose VAR analysis included long but not short rates. We will use this insight to perform our first test of the

ZLB irrelevance hypothesis IH, i.e., test that $\widetilde{\beta} = 0$ and lags of Y_{2t} can be excluded from the Y_1 equations in a KSVAR that includes long-rates and possibly money supply in Y_1.

Another important special case discussed in (Mavroeidis, 2019, Proposition 2) is the *censored* SVAR (CSVAR). This is a standard SVAR in $(Y_{1t}', Y_{2t}^*)'$. which is obtained by excluding the current and lagged values of the *actual* interest rate Y_2 from the right-hand side of equations (9) and (10). This means the dynamics of the model, and any identified IRFs, are identical across regimes. Again, the theory of the previous section shows that when $\lambda^* = 1$ and $\alpha = 0$, the solution of the DSGE model can also be equivalently written as a VAR in $(\pi_t, y_t, i_t^*)'$, with no role for the short rate i_t. Therefore, we have second test of IH in a model that does not include long rates or money supply, so does not rely on the assumption that we have adequately measured UMP in the VAR. This is an alternative to the test proposed in the previous paragraph.

4 Data

Our empirical analysis focuses on the United States and Japan. We choose data series for the baseline specification of the SVAR model to maintain the closest specification as possible to related studies and include representative series for inflation, output and measures for short- and long-term yields. For the U.S., we use quarterly data for inflation based on the GDP deflator, the output gap measure constructed by real and potential GDP, the short-term interest rate from the Federal Funds Rate, and the 10-year government bond yields from the 10-year Treasury constant maturity rate. We also consider the different measures of money listed in Appendix B. The data are from the FRED database at the Federal Reserve Bank of St. Louis, and the Center for Financial Stability databases.[10] The estimation sample for the baseline specification is from 1960q1 to 2018q4, and the alternative specifications with money cover slightly different time periods to circumvent issue with data availability. We use the value of 0.2 as the effective lower bound on the Federal Funds Rate, such that 11% of the observations at the ZLB regime, and to be consistent with Bernanke and Reinhart (2004) who suggest that the effective lower bound on nominal interest rates may be above

[10]These data is available as follows: GDP deflator `https://fred.stlouisfed.org/series/GDPDEF` and `https://fred.stlouisfed.org/series/GDPC1`; and series to construct the output gap: `https://fred.stlouisfed.org/series/GDPPOT`; the federal funds rate `https://fred.stlouisfed.org/series/FEDFUNDS`; and the long yield `https://fred.stlouisfed.org/series/GS10`. The treatment of the data is described in Appendix B. The data for the different monetary aggregates is available at: `https://fred.stlouisfed.org/categories/24` and `http://www.centerforfinancialstability.org/amfm_data.php`.

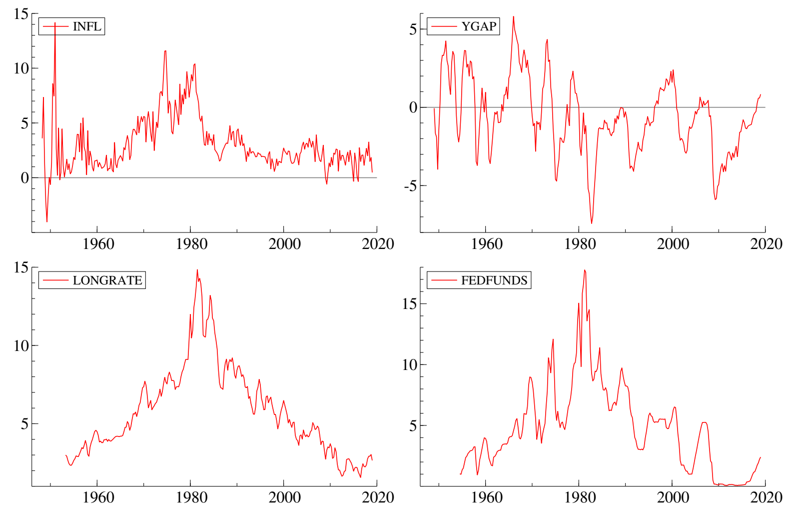


Figure 3: U.S. quarterly data on GDP deflator inflation, CBO output gap, 10 year government bond yields and the Federal funds rate. Source: FRED and CFS.

zero for institutional reasons.

For Japan, we use quarterly data for core CPI inflation, GDP growth rate, and the Call Rate. In addition, we use two alternative measures for long yields: the 9-year and the 10-year government bond yields, which are available for different sample lengths. The data source is the Bank of Japan for the Call rate, the Ministry of Finance for the 9-year and the 10-year government bond yields, the Cabinet Office for the GDP growth rate, and Statistics Bureau of Japan for core CPI inflation.[11] The estimation sample is from 1974q4 to 2019q1 if we include the 9-year government bond yields in the VAR, which is our baseline case, and from 1988q1 to 2019q1 if the 10-year yield is used instead. We use the value of 0.1 as the effective lower bound on the Call Rate, such that 36% of the observations are at the ZLB regime for the sample 1974q4-2019q1. For the sample 1988q1-2019q1, 50% of observations are at the ZLB regime.

Figures 3 and 4 show the data series for the US and Japan, respectively. A comparison between figures shows important systematic differences in the fluctuations of the series across countries. Changes in inflation are persistent in the US and inflation fluctuates around a positive trend with no protracted episodes of deflation over the sample period, while in Japan

[11] The data is available as follows. Call Rate (Bank of Japan): http://www.stat-search.boj.or.jp/index_en.html; 9-year and 10-year government bond yields (Ministry of Finance): http://www.mof.go.jp/jgbs/reference/interest_rate/index.htm; real GDP (Cabinet Office): https://www.esri.cao.go.jp/en/sna/sokuhou/sokuhou_top.html; core CPI inflation (Statistics Bureau of Japan): https://www.e-stat.go.jp/stat-search/file-download?statInfId=000031431696&fileKind=1.

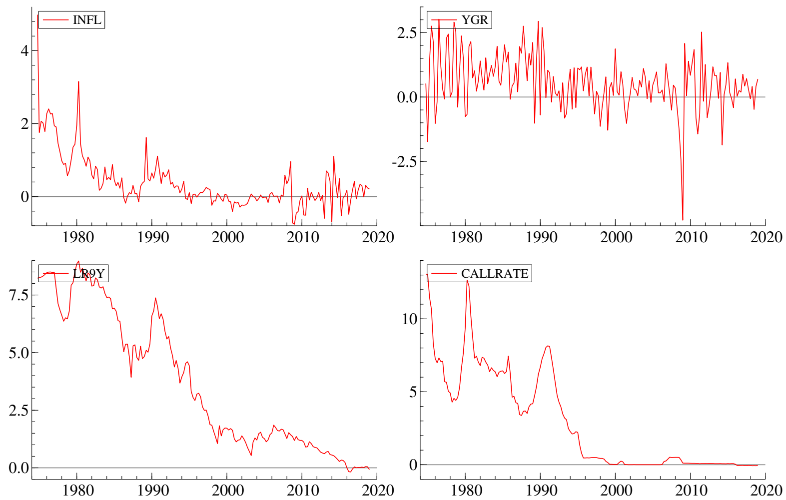


Figure 4: Japan monthly data on CPI inflation, Output gap, 10 year government bond yields and the Call rate.

inflation is much less persistent and fluctuates around zero since the 1990s.

5 Empirical results

5.1 Tests of the ZLB irrelevance hypothesis IH

Several studies assess the implications of the ZLB for the effectiveness of monetary policy by comparing the responses of key variables across ZLB and non-ZLB regimes to a monetary policy shock. If those responses are sufficiently similar across the two regimes, the ZLB is irrelevant for the effectiveness of monetary. In this section, we provide formal tests of the IH, based on the methodology discussed in Section 3.

The first approach to test the IH is motivated by Swanson (2018), and Debortoli et al. (2019), who show that monetary policy remains similarly effective across ZLB and non-ZLB regimes, and establish that long-term interest rates are a plausible indicator of the stance of monetary policy. These authors develop SVARs that include long-term, rather than short-term interest rates as indicators of monetary policy. They use such VARs to identify the impulse responses of both the macro-variables to monetary policy as well as the response of policy to economic conditions, and find that those responses are similar across ZLB and non-ZLB regimes in the US. The implicit and testable assumption that underlies their analysis is that short-terms interest rates can be excluded from the dynamics of all the other variables in the system. This hypothesis can be tested as an exclusion restriction in a SVAR that includes both the short and the long rates. Since the short rate is subject to a binding ZLB constraint, the relevant framework is the CKSVAR and the special case of KSVAR introduced in Section 3.

We first use the KSVAR model in which the shadow rate Y_{2t}^{*} does not appear on the right hand side of equations (9) and (10) and assess the exclusion restriction on the short-term nominal interest rate by testing for the joint exclusion of the current and lagged short-run yield from the Y_1 equations. For the US data, we impose the restrictions on the Federal Funds Rate for which we have 28 quarters of ZLB regimes when the series is below 20 basis points. The top panel in Table 1 shows results for the estimation of CKSVAR(p) on US data for orders p up to 5. Column (5) shows that the SVAR with three lags is the best fitting model according to the Akaike criterion, and we therefore use this specification as the benchmark model for the rest of the analysis on US data. The entries for the p-value in column (8) show that the data strongly rejects the exclusion restriction, irrespective of the number of lags in the SVAR model, thus rejecting strongly and robustly the IH. The bottom panel in the table shows results for the estimation on Japanese data with the 9-year

Table 1: Test for Excluding Short-run Rates

United States							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p	lik	par	pv-p	aic	LR	df	pval
5	-210.8	97	-	2.60	52.52	18	0.000
4	-217.9	81	0.577	2.53	49.19	15	0.000
3	-229.3	65	0.249	2.50	40.89	12	0.000
2	-262.1	49	0.000	2.66	40.55	9	0.000
1	-287.0	33	0.000	2.75	33.24	6	0.000
Japan							
p	lik	par	pv-p	aic	LR	df	pval
6	117.8	113	-	-0.06	26.72	21	0.180
5	101.7	97	0.009	-0.05	25.00	18	0.125
4	93.1	81	0.025	-0.14	23.25	15	0.079
3	85.7	65	0.058	-0.24	24.86	12	0.016
2	74.5	49	0.031	-0.30	20.25	9	0.016
1	41.1	33	0.000	-0.09	24.32	6	0.000

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood, (3) the number of parameters, (4) the p-value of the test for lag reduction, (5) the Akaike criterion for lag selection, (6) the likelihood ratio test statistics for exclusion restriction on the short-term rates, (7) the degrees of freedom for the test, and (8) the p-value for the exclusion restriction test.

yields in the model. Column (5) indicates that the best fitting model is a SVAR with two lags, and that the exclusion restriction on the Call Rate is strongly rejected. Taken together these findings suggest that the IH is consistently rejected on US and Japan data.

To ensure results are robust if we allow for alternative short-run measures on the monetary stance other then the short-run interest rate, we include the growth rate of money in the KSVAR model and test the joint exclusion restrictions on the Federal Funds Rate on US data. Table 2 shows the results for the alternative measures of the growth of money outlined in column (1). Column (2) shows that the data prefers specification of the SVAR model with 3 or 4 lags, consistent with the benchmark model. Columns (3) and (5) reports the likelihood ratio test statistics for the joint exclusion hypothesis and the corresponding p-values, respectively. These results show that the data strongly reject the joint exclusion restrictions on the Federal Funds Rate across all the alternative specifications for all measures of money supply, which corroborates the findings in the benchmark model, pointing to a strong rejection to the IH.

Next we test the same exclusion restrictions using the more flexible CKSVAR specifica-

Table 2: Test on joint exclusion of the Federal Funds Rate and Monetary Aggregates

(1) Mon. Aggr.	(2) Lags	(3) Test Stat	(4) df	(5) p-value
MB	3	52.16	16	0.0000
M1	3	54.84	16	0.0000
M2	3	54.19	16	0.0000
M2M	4	72.65	20	0.0000
MZM	4	78.04	20	0.0000
DIVM1	4	83.38	20	0.0000
DIVM2	4	114.59	20	0.0000
DIVM2M	4	113.52	20	0.0000
DIVMZM	4	108.41	20	0.0000
DIVM4	4	139.63	20	0.0000
SHO	3	91.65	16	0.0000

Note. Each column reports the following: (1) Measure of money aggregate, (2) number of lags in the SVAR model, (3) χ^2 test statistics for the joint exclusion restrictions test, (4) the degrees of freedom for the test, and (5) the p-values for the exclusion restrictions test.

tion. Recall that the key difference of the CKSVAR relative to the KSVAR model is that the latent shadow rate Y_{2t}^* appears in equation (10) (i.e. the policy rule) as well as equation (9). Specifically we test $\lambda = 0$, $B_{12} = 0$ and $B_{12}^* = 0$ in equation (9) without imposing any restrictions on equation (10). This corresponds to testing, equivalently, that $\widetilde{\beta} = 0$, $C_{12} = 0$ and $C_{12}^* = 0$ in the reduced-form equation (13). The motivation for using the CKSVAR model is to avoid any misspecification in the dynamics of the short-term interest (by incorrectly setting $B_{22}^* = 0$) rate which would affect the asymptotic validity of the test of IH. Table 3 shows the results of this test for the US (top panel) and Japan (bottom panel). Column (5) shows that the number of lags to best fit the data is 3 for the US and 2 for Japan, respectively. As in the benchmark specification, the data strongly reject the IH in both countries, irrespective of the number of lags in the model.

The second test of the IH is a likelihood ratio test of the CSVAR model, according to which policy is as effective at the ZLB as it is above the ZLB, against the CKSVAR model, which allows for a wide range of degrees of effectiveness of UMP at the ZLB. As explained in Section 3, this model does not rely on any direct measures of UMP, but rather models UMP through the shadow rate Y_{2t}^*. So, the CKSVAR includes only three variables, inflation, output gap or growth and the policy rate.

Table 4 shows the results for the US (top entries) and Japan (bottom entries). We include

Table 3: Exclusion of Short Rate and Shadow Rate, Incorporating Forward Guidance

United States							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p	lik	par	pv-p	aic	LR	df	pval
5	-191.7	117	-	2.60	76.82	33	0.000
4	-200.4	97	0.628	2.52	69.96	27	0.000
3	-212.6	77	0.395	2.46	60.78	21	0.000
2	-252.0	57	0.000	2.64	47.96	15	0.000
1	-279.9	37	0.000	2.72	37.60	9	0.000
Japan							
p	lik	par	pv-p	aic	LR	df	pval
6	143.6	137	-	-0.08	58.13	39	0.025
5	122.8	117	0.003	-0.07	48.50	33	0.040
4	111.6	97	0.010	-0.17	41.12	27	0.040
3	102.4	77	0.030	-0.30	39.40	21	0.009
2	87.4	57	0.010	-0.35	28.32	15	0.020
1	51.2	37	0.000	-0.16	23.97	9	0.004

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood, (3) the number of parameters, (4) the p-value of the test for lag reduction, (5) the Akaike criterion for lag selection, (6) the likelihood ratio test statistics for exclusion restriction on the short-term rates, (7) the degrees of freedom for the test, and (8) the p-value for the exclusion restriction test.

Table 4: Testing CSVAR against CKSVAR

Country	p	LR	df	pval
US	3	25.63	15	0.042
Japan	2	24.43	11	0.011

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood ratio, (3) the degrees of freedom for the test, and (4) the p-value for the test.

Table 5: Test for Excluding Short-run Rates, for the subsample 1984q1-2019q1

United States							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p	lik	par	pv-p	aic	LR	df	pval
5	103.4	97	-	-0.09	27.78	18	0.066
4	96.09	81	0.600	-0.22	25.76	15	0.040
3	94.12	65	0.971	-0.42		12	0.099
2	69.77	49	0.034	-0.30	40.55	9	0.003
1	40.05	33	0.000	-0.10	33.24	6	0.152

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood, (3) the number of parameters, (4) the p-value of the test for lag reduction, (5) the Akaike criterion for lag selection, (6) the likelihood ratio test statistics for exclusion restriction on the short-term rates, (7) the degrees of freedom for the test, and (8) the p-value for the exclusion restriction test.

3 lags for the US and 2 lags for Japan, according to the Akaike information criteria.[12] These results clearly indicate that the IH is rejected both for the US and Japan at the 5% level of significance.

Robustness to changes in volatility Our econometric analysis is fully parametric and crucially relies on the assumption that the model is correctly specified. This requires in particular that the variance of the errors is constant over the sample, and therefore, the results are subject to a possible misspecification arising from the 'Great Moderation', a drop in US macroeconomic volatility in the mid-1980s. Therefore, we assess the robustness of our results by estimating the model and performing the above tests of IH over the sub-sample which starts at 1984q1. Tables 6 - 7 show the results of the test of the IH for this sub-sample. The test for lag reduction and the AIC still support choosing 3 lags. Various tests for IH demonstrate that our results largely hold for the Great Moderation sub-sample. Neither the CSVAR model, where the unconventional policy is fully effective, nor the model without the short-term interest rate will fully capture the policy effects at the ZLB periods.

Similarly, we test the robustness of our results for the Japanese data by using the 10-year yields in the model instead. This shortens the available sample for estimation to 1988q1 to 2019q1. Tables 8 - 10 report test statistics for the 3 types of tests for the IH hypothesis. From Tables 8 and 9, the IH is rejected across all lags. For the CKSVAR alternative, we select 3 lags based in the Akaike criterion. Then Table 10 also suggests the rejection of the IH.

[12]Table 3 reports the AIC for several alternative specifications of the model.

Table 6: Exclusion of Short Rate and Shadow Rate, Incorporating Forward Guidance, for the subsample 1984q1-2019q1

United States							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p	lik	par	pv-p	aic	LR	df	pval
5	129.4	117	-	-0.18	70.71	33	0.000
4	117.0	97	0.214	-0.28	60.79	27	0.000
3	111.2	77	0.635	-0.49	44.97	21	0.002
2	89.6	57	0.046	-0.47	55.92	15	0.000
1	46.5	37	0.000	-0.14	20.24	9	0.016

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood, (3) the number of parameters, (4) the p-value of the test for lag reduction, (5) the Akaike criterion for lag selection, (6) the likelihood ratio test statistics for exclusion restriction on the short-term rates, (7) the degrees of freedom for the test, and (8) the p-value for the exclusion restriction test.

Table 7: Testing CSVAR against CKSVAR, for the subsample 1984q1-2019q1

Country	p	LR	df	pval
US	3	24.19	15	0.062

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood ratio, (3) the degrees of freedom for the test, and (4) the p-value for the test.

Table 8: Test for Excluding Short-run Rates

Japan							
p	lik	par	pv-p	aic	LR	df	pval
6	210.1	113	-	-1.55	56.12	21	0.000
5	188.9	97	0.000	-1.45	43.79	18	0.001
4	182.1	81	0.005	-1.62	40.37	15	0.000
3	177.3	65	0.047	-1.80	38.49	12	0.000
2	162.0	49	0.006	-1.81	38.99	9	0.001
1	113.4	33	0.000	-1.29	30.79	6	0.000

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood, (3) the number of parameters, (4) the p-value of the test for lag reduction, (5) the Akaike criterion for lag selection, (6) the likelihood ratio test statistics for exclusion restriction on the short-term rates, (7) the degrees of freedom for the test, and (8) the p-value for the exclusion restriction test.

Table 9: Exclusion of Short Rate and Shadow Rate, Incorporating Forward Guidance

Japan							
p	lik	par	pv-p	aic	LR	df	pval
6	247.4	137	-	-1.77	112.97	39	0.000
5	225.3	117	0.001	-1.73	102.10	33	0.000
4	211.2	97	0.001	-1.83	85.42	27	0.000
3	196.8	77	0.001	-1.92	63.84	21	0.000
2	175.7	57	0.000	-1.90	51.82	15	0.002
1	124.3	37	0.000	-1.40	36.79	9	0.000

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood, (3) the number of parameters, (4) the p-value of the test for lag reduction, (5) the Akaike criterion for lag selection, (6) the likelihood ratio test statistics for exclusion restriction on the short-term rates, (7) the degrees of freedom for the test, and (8) the p-value for the exclusion restriction test.

Table 10: Testing CSVAR against CKSVAR

Country	p	LR	df	pval
Japan	3	61.53	15	0.000

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood ratio, (3) the degrees of freedom for the test, and (4) the p-value for the test.

5.2 Tests of the (ir)relevance of the long-term rate

The statistical tests in the previous section reject the IH of the ZLB, and thus the possibility of excluding short yields by controlling for long yields. We now assess whether movements in the short-term interest rate are sufficient to encapsulate the effect of both conventional and unconventional monetary policy, and test the exclusion restriction on long-term rates from the SVAR models. Following our benchmark specification, we test for excluding long-run rates in the KSVAR and the CKSVAR models.

Table 11 shows results of this test for the US (top panel) and Japan (bottom panel) in the KSVAR specification. Column (5) shows that the number of lags that best fit the data is 3 for the US and 2 for Japan, and column (8) shows that the null hypothesis of excluding long-term yields cannot be rejected in the SVAR with the preferred lags specification for the US. However, this evidence is less strong for Japan since the hypothesis is rejected in the SVAR across all lags. Overall, the analysis suggests that long-term yields are not necessary to determine the effect of monetary policy on US data, whereas this does not apply for Japan.

We further check the robustness of this result by doing the same test for an unrestricted

Table 11: Exclusion of Long-term Rates

USA							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
p	lik	par	pv-p	AIC	LR	df	pval
5	-210.8	97	-	2.597	9.632	10	0.473
4	-217.9	81	0.577	2.533	8.026	8	0.431
3	-229.3	65	0.249	2.504	5.145	6	0.525
2	-262.1	49	0.000	2.659	11.326	4	0.023
1	-287.0	33	0.000	2.747	14.727	2	0.001
Japan							
p	lik	par	pv-p	aic	LR	df	pval
6	117.8	113	-	-0.06	28.75	12	0.004
5	101.7	97	0.009	-0.05	29.23	10	0.001
4	93.1	81	0.025	-0.14	21.97	8	0.005
3	85.7	65	0.058	-0.24	18.34	6	0.005
2	74.5	49	0.031	-0.30	20.70	4	0.000
1	41.1	33	0.000	-0.09	19.16	2	0.000

Note. Each column reports the following: (1) number of lags in the SVAR, (2) the likelihood, (3) the number of parameters, (4) the p-value of the test for lag reduction, (5) the Akaike criterion for lag selection, (6) the likelihood ratio test statistics for exclusion restriction on the short-term rates, (7) the degrees of freedom for the test, and (8) the p-value for the exclusion restriction test.

Country	p	LR	df	asymptotic pval	bootstrapped pval
US	3	4.671	6	0.587	
Japan	2	15.375	4	0.004	

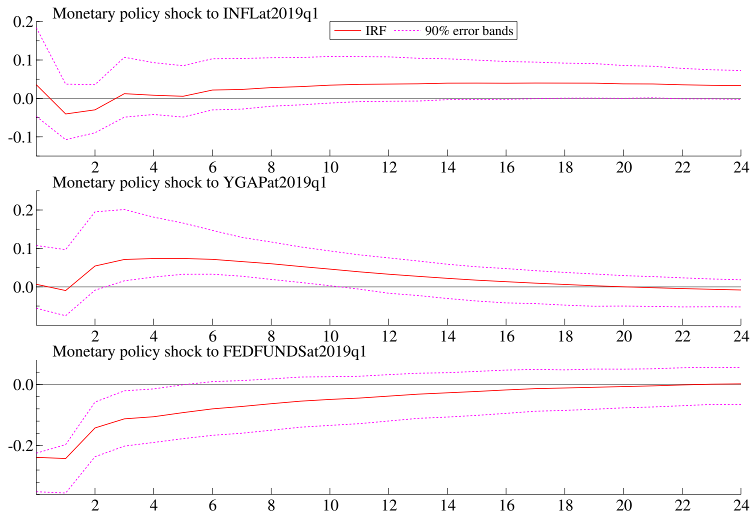
Table 12: Exclusion of Long Rate in a CKSVAR model

CKSVAR model. The test results are reported in Table 12. We still observe that long-term yields can be excluded for the US data, but not for Japan.

5.3 Impact of monetary policy in the US

Our test results in the previous section show that the dynamic responses of inflation and output to a monetary policy shock can be evaluated within a 3-equation CKSVAR(3) model with short-term interest rate for the US data. In this section, we show the effects of monetary policy by identifying and depicting the corresponding impulse response functions (IRFs) to monetary policy shocks within this model. Since the model is nonlinear, the IRFs are state-dependent. There is no unique way of defining the IRFs in this case, so, we will use the

Figure 5: IRFs to -25bp monetary policy shock in 2019q1 for the US based on a CKSVAR(3) estimated over the period 1960q1-2018q4, under the assumption that UMP has no impact effect on inflation and output. 90% error bands are computed with a parametric bootstrap using 1000 replications.



Notes. Add here description of IRFs.

IRF definition in Koop et al. (1996), which is the most commonly used in the literature. According to this approach, the IRFs to a monetary policy shock of magnitude ς are given by:

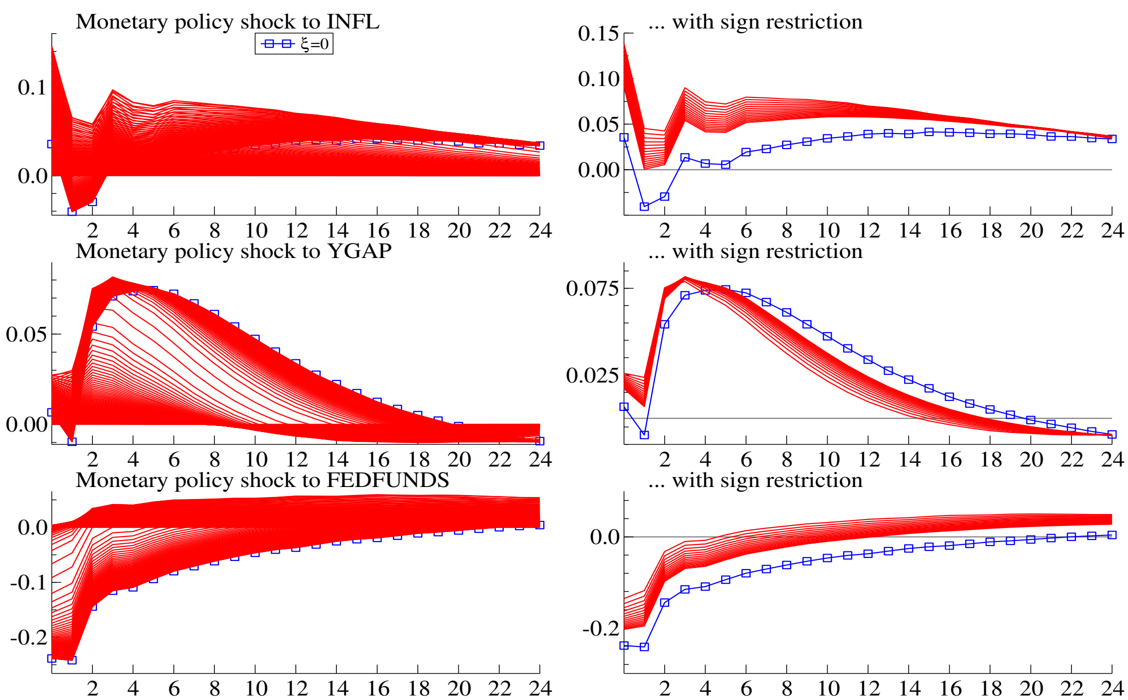
$$IRF_{h,t}\left(\varsigma, X_t, \overline{X}_t^*\right) = E\left(Y_{t+h}|\varepsilon_{2t} = \varsigma, X_t, \overline{X}_t^*\right) - E\left(Y_{t+h}|\varepsilon_{2t} = 0, X_t, \overline{X}_t^*\right), \tag{16}$$

where $\overline{X}_{t,j}^* = \min\left(Y_{2t-j}^* - b_t, 0\right)$ for $j = 1, ..., p$, where p is the order of the VAR.

Following the identification strategy of Mavroeidis (2019), the IRFs are point-identified if we assume that there is no contemporaneous effect of UMP on Y_1, which corresponds to $\xi = 0$ in the CKSVAR model. The IRFs of this case for a $-25bp$ monetary policy shock in the US in 2019q1 are shown in Figure 5, together with 90% error bands obtained using a parametric bootstrap with 1000 replications. The graph shows that the federal funds rate falls while inflation and the output gap increase following an expansionary monetary policy shock.

If we relax the assumption that the unconventional monetary policy is completely ineffective contemporaneously at the ZLB, the IRFs are no longer point-identified. We can get the identified set on ξ, β and γ and hence the IRFs, by solving equations (14) and (15) at the estimated values of $\widetilde{\beta}$ and Ω, as explained in Section 3 above, see the discussion below equations (14) and (15). The algorithm for obtaining the identified set is given in Mavroeidis (2019). When we restrict the range of ξ to $[0, 1]$ but impose no further identifying restrictions, the identified set for ξ is $[0, 0.78]$ for the US data. The corresponding IRFs to a $-25bp$ monetary policy shock are shown on the left graphs in Figure 6. The identified sets are quite wide but their signs mostly conform to the theoretical IRFs from the DSGE model of section 2.

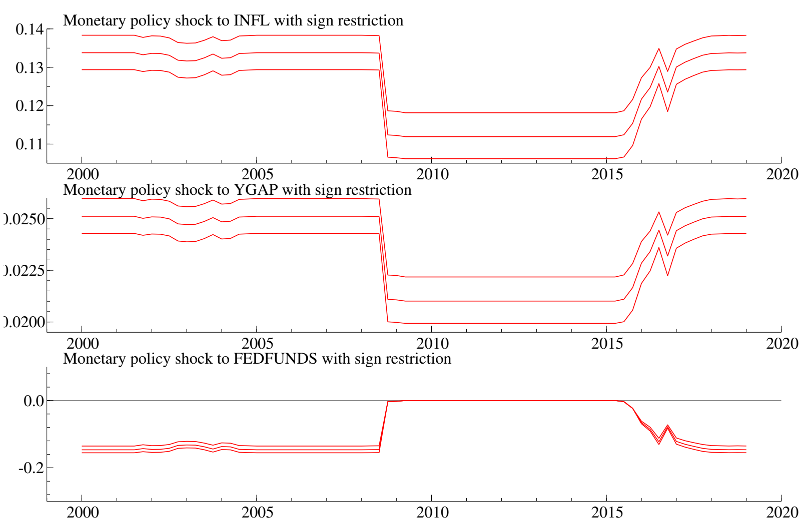
Figure 6: Identified sets of IRFs in 2019q1 to -25bps monetary policy shock in the CKSVAR(3) model for the US, estimated over the period 1960q1-2018q4. On the left side, $\xi \in [0, 0.78]$ without any additional restrictions, on the right graphs, $\xi \in [0.74, 0.76]$ by the additional sign restrictions that the shock has nonnegative effects on inflation and output and nonpositive effects on the Fed funds rate up to 4 quarters.



The identified set can be further sharpened by using sign restrictions for identification. For comparability, we follow Debortoli et al. (2019) and impose the restrictions that a negative interest rate shock should have a nonnegative effect on inflation and output, and nonpositive effect on interest rates for the first 4 quarters following the shock. For the US data, the set of ξ that is consistent with this set of sign restrictions over the entire sample (recall that IRFs are state dependent, so the sign restrictions need to be imposed in every period) is narrowed down to $[0.74, 0.76]$. This is a dramatic reduction in uncertainty, but note that this set still does not incorporate sampling uncertainty in the estimation of the reduced-form parameters $\widetilde{\beta}$ and Ω. Sampling uncertainty can be measured using the frequentist confidence bands recently proposed by Granziera Moon and Schorfheide (2018) [still in progress]. The corresponding set for the IRFs is given by the right graphs in Figure 6. These now accord even more with the IRFs in the calibrated DSGE model of Section 2.

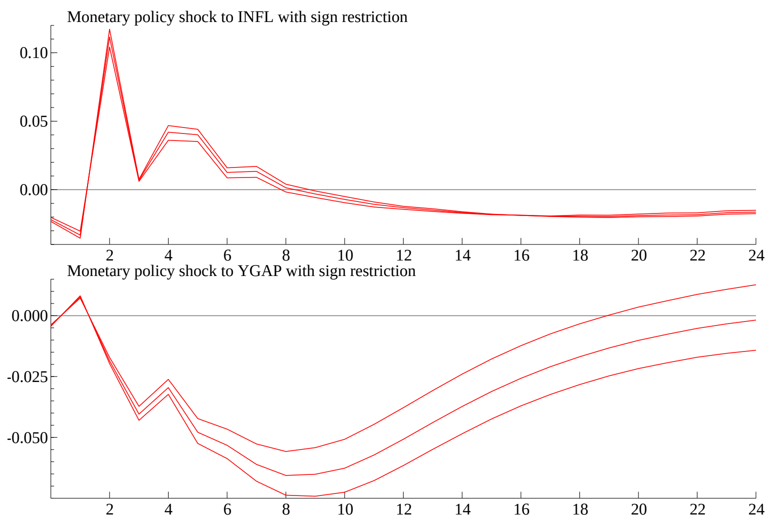
Our aforementioned estimates of ξ suggest that UMP has been roughly 75% as effective as conventional monetary policy in the US, on impact. To further demonstrate this point, Figure 7 plots the impact effect of a -25bp monetary policy shock on all variables over the period 2000-2019 for $\xi \in [0.74, 0.76]$. It is evident that the effects of monetary easing on inflation and output gap are reduced during 2009-2015, which corresponds to the ZLB periods in the sample.

Figure 7: Impact effect of -25bp monetary policy shock over the period 2000-2019. The identified set on ξ is [0.74,0.76], based on a CKSVAR(3) model for the US, estimated over the period 1960q1-2018q4 with the sign restrictions that the shock has nonnegative effect on inflation and output and nonpositive effect on the Fed funds rate up to 4 quarters.



These comparisons nevertheless only capture the relative efficacy of the unconventional policy on impact. It is also interesting to see the relative efficacy of the UMP at longer horizons. Figure 8 illustrates the difference between the identified set of IRFs at a typical ZLB period, which is the average of the observations at the ZLB periods, and the IRFs at a typical non-ZLB period, which is chosen to be 2000q1. The identified sets with sign restrictions for ξ for both starting points are $[0.62, 0.76]$. The figures indicates that conventional monetary policy is more effective than UMP especially over longer horizons. However, UMP appears to have been more effective than conventional policy within 2-8 quarters.

Figure 8: Differences of IRFs to a -25bp monetary policy shock during a typical ZLB period and a typical non-ZLB period, from a CKSVAR(3) model for the US, estimated over the period 1960q1-2018q4. The identified set on ξ is [0.74,0.76] obtained under the sign restrictions that the shock has nonnegative effect on inflation and output and nonpositive effect on the Fed funds rate up to 4 quarters.



6 Conclusion

The paper develops theoretical and empirical models to study the effectiveness and transmission mechanism of unconventional policies. The theoretical model allows the degree of effectiveness of unconventional policy to range from fully effective to completely ineffective. Our empirical analysis is based on a more agnostic a structural VAR model that accounts for the effective lower bound on the policy rate. Our analysis shows that there is strong evidence against the hypothesis that the ZLB is empirically irrelevant, which implies that the ZLB has been an important constraint on monetary policy both in the US and Japan, but also shows that unconventional monetary policy has been quite effective in the US.

References

Andrés, J., D. López-Salido, and E. Nelson (2004). Tobin's imperfect asset substitution in optimizing general equilibrium. *Journal of Money, Credit and Banking 36*(4), 665–690.

Bernanke, B. (2014). A conversation: the fed yesterday, today and tomorrow. A discussion with Federal Reserve Chairman Ben Bernanke on the Fed's 100th Anniversary (Excerpt).

Bernanke, B. S. and V. R. Reinhart (2004, May). Conducting monetary policy at very low short-term interest rates. *American Economic Review 94*(2), 85–90.

Calvo, G. A. (1983). Staggered prices in a utility maximizing framework. *Journal of Monetary Economics 12*, 383–398.

Chen, H., V. Cúrdia, and A. Ferrero (2012, November). The macroeconomic effects of large-scale asset purchase programmes. *The Economic Journal 122*, F289–F315.

Christensen, J. H. E. and G. D. Rudebusch (2012, 10). The Response of Interest Rates to US and UK Quantitative Easing. *The Economic Journal 122*(564), F385–F414.

Debortoli, D., J. Galí, and L. Gambetti (2019, April). *On the Empirical (Ir)Relevance of the Zero Lower Bound Constraint.* University of Chicago Press.

Del Negro, M., G. Eggertsson, A. Ferrero, and N. Kiyotaki (2017, March). The great escape? a quantitative evaluation of the fed's liquidity facilities. *American Economic Review 107*(3), 824–57.

Eggertsson, G. B. and M. Woodford (2003). Zero bound on interest rates and optimal monetary policy. *Brookings papers on economic activity 2003*(1), 139–233.

Gertler, M. and P. Karadi (2013). Qe 1 vs. 2 vs. 3...: A framework for analyzing large-scale asset purchases as a monetary policy tool. *International Journal of Central Banking 9*(1), 5–53.

Guerrieri, L. and M. Iacoviello (2015). Occbin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics 70*, 22–38.

Gust, C., E. Herbst, D. Lpez-Salido, and M. E. Smith (2017, July). The empirical implications of the interest-rate lower bound. *American Economic Review 107*(7), 1971–2006.

Harrison, R. (2012). Asset purchase policy at the effective lower bound for interest rates. Working Paper 444, Bank of England.

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

Inoue, A. and B. Rossi (2018). The effects of conventional and unconventional monetary policy: A new approach. Technical report.

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

Krishnamurthy, A. and A. Vissing-Jorgensen (2011). The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy. *Brookings Papers on Economic Activity 42*(2 (Fall)), 215–287.

Liu, P., K. Theodoridis, H. Mumtaz, and F. Zanetti (2019). Changing macroeconomic dynamics at the zero lower bound. *Journal of Business & Economic Statistics 37*(3), 391–404.

Mavroeidis, S. (2019). Identification at the Zero Lower Bound. mimeo.

Reifschneider, D. and J. C. Williams (2000). Three lessons for monetary policy in a low-inflation era. *Journal of Money, Credit and Banking 32*(4), 936–966.

Bank of Japan (2016). Comprehensive assessment: Developments in economic activity and prices as well as policy effects since the introduction of quantitative and qualitative easing (qqe), the background. Released on September 21, 2016 as Attachment 1 to the statement on monetary policy.

Swanson, E. T. (2018). The federal reserve is not very constrained by the lower bound on nominal interest rates. Technical report, National Bureau of Economic Research.

Swanson, E. T. and J. C. Williams (2014). Measuring the effect of the zero lower bound on medium- and longer-term interest rates. *American Economic Review 104*(10), 3154–3185.

Ueda, K. (2012). Deleveraging and monetary policy: Japan since the 1990s and the United States since 2007. *Journal of Economic Perspectives 26*(3), 177–202.

Ugai, H. (2007). Effects of the quantitative easing policy: A survey of empirical analyses. *Monetary and Economic Studies 25*(1), 1–48.

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.

A Theoretical Model

We present a simple New Keynesian model with a ZLB and UMP. The model is a simplified version of Chen, Curdia, and Ferrero (2012). To keep the analysis focused on the salient features of the transmission mechanisms of UMP, the model abstracts from capital accumulation, consumption habit formation, and various shocks. The model combines two types of unconventional monetary policies: quantitative easing (QE) – a central bank's purchase of long-term government bonds, activated when the economy hits the ZLB; and forward guidance (FG). There are three types of shocks: a demand (preference) shock, a supply (TFP) shock, and a monetary policy shock.

A.1 Model building blocks

A.1.1 Long-term bonds

We start by describing long-term government bonds. There is a consol bond. The consol bond issued at time t yields κ^{j-1} dollars at time $t+j$ over time. Let $R_{L,t+1}$ denote the gross nominal rate from time t to $t+1$. The period-t price of the bond issued at time t, $P_{L,t}$, is defined as

$$\begin{aligned} P_{L,t} &= E_t\left(\frac{1}{R_{L,t+1}} + \frac{\kappa}{R_{L,t+1}R_{L,t+2}} + \frac{\kappa^2}{R_{L,t+1}R_{L,t+2}R_{L,t+3}} + ...\right) \\ &= E_t\left(\frac{1}{R_{L,t+1}} + \frac{\kappa}{R_{L,t+1}}P_{L,t+1}\right). \end{aligned} \tag{17}$$

The gross yield to maturity (or the long-term interest rate) at time t, $\bar{R}_{L,t}$ is defined as

$$E_t\left(\frac{1}{\bar{R}_{L,t}} + \frac{\kappa}{\left(\bar{R}_{L,t}\right)^2} + \frac{\kappa^2}{\left(\bar{R}_{L,t}\right)^3} + ...\right) = P_{L,t},$$

or

$$P_{L,t} = \frac{1}{\bar{R}_{L,t} - \kappa}. \tag{18}$$

Let $B_{L,t|t-s}$ denote period-t bond holdings issued at time $t-s$. Suppose that a household owns $B_{L,t|t-s}$ for $s = 1, 2, ...$ in the beginning of period t. The total amount of dividends the household receives in period t is

$$\sum_{s=1}^{\infty} \kappa^{s-1} B_{L,t|t-s}.$$

Note that having one unit of $B_{L,t|t-s}$ is equivalent to having κ^{s-1} units of $B_{L,t|t-1}$ because the both yield κ^{s-1} dollars. The total amount of dividends then can be expressed in terms of $B_{L,t|t-1}$ as

$$\sum_{s=1}^{\infty} \kappa^{s-1} B_{L,t|t-s} \equiv B_{L,t-1},$$

where $B_{L,t-1}$ denotes the amount of bonds in units of the bonds issued at time $t-1$, held by the household in the beginning of period t. Let $P_{L,t|t-s}$ denote the time-t price of the bond issued at time $t-s$. Then, the value of all the bonds at time t is

$$\sum_{s=1}^{\infty} P_{L,t|t-s} B_{L,t|t-s}$$

Each price satisfies

$$\begin{aligned} P_{L,t|t-s} &= E_t \left(\frac{\kappa^s}{R_{L,t+1}} + \frac{\kappa^{s+1}}{R_{L,t+1} R_{L,t+2}} + \frac{\kappa^{s+2}}{R_{L,t+1} R_{L,t+2} R_{L,t+3}} + \dots \right) \\ &= \kappa^s P_{L,t} \end{aligned}$$

Then the value of all the bonds at time t is

$$\sum_{s=1}^{\infty} P_{L,t|t-s} B_{L,t|t-s} = P_{L,t} \kappa \sum_{s=1}^{\infty} \kappa^{s-1} B_{L,t|t-s} = \kappa P_{L,t} B_{L,t-1}.$$

So the return of holding $B_{L,t-1}$ is given by the sum of dividends and the value of all the bonds as:

$$B_{L,t-1} + \kappa P_{L,t} B_{L,t-1} = (1 + \kappa P_{L,t}) B_{L,t-1} = P_{L,t} \bar{R}_{L,t} B_{L,t-1} = \frac{\bar{R}_{L,t}}{\bar{R}_{L,t} - \kappa} B_{L,t-1}.$$

A.1.2 Households

There are two types of households: unrestricted households (U-households) and restricted households (R-households). U-households, with population ω_u, can trade both short-term and long-term government bonds subject to a transaction cost ζ_t per unit of long-term bonds purchased. R-households, with population $\omega_r = 1 - \omega_u$, can trade only long-term government bonds. For $j = u, r$, each household chooses consumption c_t^j, hours worked h_t^j, government bond holdings $B_{L,t}^j$ and B_t^j tomaximize utility,

$$\sum_{t=0}^{\infty} \beta_j^t d_t \left[\frac{\left(c_t^j\right)^{1-\sigma}}{1-\sigma} - \psi \frac{\left(h_t^j\right)^{1+1/\nu}}{1+1/\nu} \right],$$

subject to: for a U-household,

$$P_t c_t^u + B_t^u + (1 + \zeta_t) P_{L,t} B_{L,t}^u = (1 + i_{t-1}) B_{t-1}^u + P_{L,t} \bar{R}_{L,t} B_{L,t-1}^u + W_t h_t^u - T_t^u + \Pi_t^u,$$

and for a R-household,

$$P_t c_t^r + P_{L,t} B_{L,t}^r = P_{L,t} \bar{R}_{L,t} B_{L,t-1}^r + W_t h_t^r - T_t^r + \Pi_t^r,$$

where P_t is the price level and i_t is the short-term interest rate. In addition $\bar{R}_{L,t}$ denotes the gross yield to maturity at time t on the long-term bond

$$\bar{R}_{L,t} = \frac{1}{P_{L,t}} + \kappa, \quad 0 < \kappa \leq 1.$$

The average duration of the bond is given by $\bar{R}_{L,t} / \left(\bar{R}_{L,t} - \kappa\right)$. There is a shock d_t to the preference, and it is given by:

$$d_t = \begin{cases} e^{z_1^b} e^{z_2^b} ... e^{z_t^b} & \text{for } t \geq 1 \\ 1 & \text{for } t = 0 \end{cases},$$

where z_t^b is the preference (demand) shock, which is assumed to follow the AR(1) process

$$z_t^b = \rho_b z_{t-1}^b + \epsilon_t^b,$$

with $\epsilon_t^b \sim$ i.i.d. $N\left(0, \sigma_b^2\right)$.

We assume that the transaction cost of trading long-term bonds for the U-households is collected by financial frims and redistributed as a lump-sum profits to the U-households. Under the assumption, the transaction cost does not appear in the good market clearing condition, which is given by:

$$y_t = \omega_u c_t^u + (1 - \omega_u) c_t^r. \tag{19}$$

Arranging the first-order conditions of the U-household's problem yields the following optimality conditions:

$$w_t = \psi \left(c_t^u\right)^\sigma \left(h_t^u\right)^{1/\nu}, \tag{20}$$

$$1 = E_t \beta_u e^{z_{t+1}^b} \left(\frac{c_{t+1}^u}{c_t^u}\right)^{-\sigma} \frac{1 + i_t}{\pi_{t+1}}, \tag{21}$$

$$1 + \zeta_t = E_t \beta_u e^{z_{t+1}^b} \left(\frac{c_{t+1}^u}{c_t^u}\right)^{-\sigma} \frac{R_{L,t+1}}{\pi_{t+1}}, \tag{22}$$

where $w_t \equiv W_t / P_t$ denotes the real wage, $\pi_t \equiv P_t / P_{t-1}$ denotes the inflation rate, and $R_{L,t+1}$ denotes the annual yield of the long-term bond between periods t and $t+1$, given by

$$R_{L,t+1} \equiv \frac{P_{L,t+1}}{P_{L,t}} \bar{R}_{L,t+1} = \frac{P_{L,t+1}}{P_{L,t}} \left(\frac{1}{P_{L,t+1}} + \kappa\right) = \frac{1 + \kappa P_{L,t+1}}{P_{L,t}}.$$

Similarly, arranging the first-order conditions of the R-household's problem yields

$$w_t = \psi \left(c_t^r\right)^\sigma \left(h_t^r\right)^{1/\nu}, \tag{23}$$

$$1 = E_t \beta e^{z_{t+1}^b} \left(\frac{c_{t+1}^r}{c_t^r}\right)^{-\sigma} \frac{R_{L,t+1}}{\pi_{t+1}}, \tag{24}$$

A.1.3 Firms

The firm sector consist of two types of firms: final good firms and intermediate goods firms. The problem of these firms is standard except that the average discount rate between U-households and R-households is used in discounting the profits of these firms. The profits need to be derived explicitly because one of the two households' budget constraints constitutes an equilibrium condition as well as a good market clearing condition.

Competitive final good firms combine intermediate goods $\{Y_t(i)\}_{i=0}^{1}$ and produce the final good Y_t according to

$$Y_t = \left[\int_0^1 Y_t(i)^{\frac{1}{\lambda_p}}\, di\right]^{\lambda_p},\ \lambda_p > 1.$$

The demand function for the i-th intermediate good is given by

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{\frac{\lambda_p}{1-\lambda_p}} Y_t.$$

Intermediate goods firms use labor and produce intermediate goods according to

$$Y_t(i) = e^{z_t^a} h_t(i)^{\theta},\ 0 < \theta \leq 1.$$

where z_t^a denotes a technology shock, which is assumed to follow:

$$z_t^a = \rho_a z_{t-1}^a + \epsilon_t^a,$$

with $\epsilon_t^a \sim$ i.i.d. $N(0, \sigma_a^2)$. Because there is no price dispersion in steady state thanks to firms' price indexation, the aggregate output can be written up to the first-order approximation as:

$$Y_t = e^{z_t^a} h_t^{\theta}, \tag{25}$$

where Y_t and h_t denote the aggregate output and hours worked. The total cost of producing $Y_t(i)$ is equal to

$$W_t h_t(i) = W_t \left(\frac{Y_t(i)}{e^{z_t^a}}\right)^{\frac{1}{\theta}}.$$

In each period intermediate goods firms can change their price with probability ξ identically and independently across firms and over time. For each i the i-th intermediate good firm chooses the price, $\tilde{P}_t(i)$, to maximize its discounted sum of profits,

$$\max_{\tilde{P}_t(i)} E_t \sum_{s=0}^{\infty} (\xi)^s \bar{\Lambda}_{t+s|t} \left[P_{t+s}(i) Y_{t+s}(i) - W_{t+s}\left(\frac{Y_{t+s}(i)}{e^{z_{t+s}^a}}\right)^{\frac{1}{\theta}}\right],$$

subject to the demand curve,

$$Y_{t+s}(i) = \left(\frac{P_{t+s}(i)}{P_{t+s}}\right)^{\frac{\lambda_p}{1-\lambda_p}} Y_{t+s},$$

where

$$\begin{aligned}
&\bar{\Lambda}_{t+s|t} \equiv \omega_u \Lambda^u_{t+s|t} + (1-\omega_u)\,\Lambda^r_{t+s|t}\,, \\
&\Lambda^j_{t+s|t} = \beta^s_j d_{t+s|t}\left(c^j_{t+s}/c^j_t\right)^{-\sigma}\frac{1}{P_{t+s}}, \quad d_{t+s|t} = e^{z^b_{t+1}}e^{z^b_{t+2}}...e^{z^b_{t+s}} \\
&P_{t+s}(i) = \tilde{P}_t(i)\Pi^p_{t,t+s}, \\
&\Pi^p_{t+s|t} = \begin{cases} 1 & \text{if } s=0 \\ \prod_{k=1}^{s}(\pi_{t+k-1})^{\iota_p}(\pi)^{1-\iota_p} & \text{if } s=1,2,... \end{cases}.
\end{aligned}$$

Substituting the demand curve into the objective function yields

$$\max_{\tilde{P}_t(i)} E_t \sum_{s=0}^{\infty}(\xi)^s\bar{\Lambda}_{t+s|t}\left[\tilde{P}_t(i)\Pi^p_{t+s|t}\left(\frac{\tilde{P}_t(i)\Pi^p_{t+s|t}}{P_{t+s}}\right)^{\frac{\lambda_p}{1-\lambda_p}} Y_{t+s} - W_{t+s}\left(\frac{\tilde{P}_t(i)\Pi^p_{t+s|t}}{P_{t+s}}\right)^{\frac{\lambda_p}{(1-\lambda_p)\theta}}\left(\frac{Y_{t+s}}{e^{z^a_{t+s}}}\right)^{\frac{1}{\theta}}\right].$$

The first-order condition is

$$0 = E_t\sum_{s=0}^{\infty}(\xi)^s\bar{\Lambda}_{t+s|t}\left[\frac{1}{1-\lambda_p}\Pi^p_{t+s|t}Y_{t+s}(i) - W_{t+s}\frac{\lambda_p}{(1-\lambda_p)\,\theta}\left(\frac{Y_{t+s}(i)}{e^{z^a_{t+s}}}\right)^{\frac{1}{\theta}}\frac{1}{\tilde{P}_t(i)}\right].$$

Since $\tilde{P}_t(i)$ does not depend on i, index i is omitted hereafter. Define $\tilde{p}_t \equiv \tilde{P}_t/P_t$ and

$$\tilde{\Pi}^p_{t+s|t} = \begin{cases} 1 & \text{if } s=0 \\ \prod_{k=1}^{s}\frac{(\pi_{t+k-1})^{\iota_p}(\pi)^{1-\iota_p}}{\pi_{t+k}} & \text{if } s=1,2,... \end{cases}$$

The first-order condition can be transformed as

$$\begin{aligned}
0 = E_t\sum_{s=0}^{\infty}(\xi)^s\bar{\Lambda}_{t+s}P_{t+s}&\left[\frac{1}{1-\lambda_p}\frac{\Pi^p_{t+s|t}}{P_{t+s}}\left(\tilde{p}_t\tilde{\Pi}^p_{t+s|t}\right)^{\frac{\lambda_p}{1-\lambda_p}}Y_{t+s}\right. \\
&\left.-\frac{W_{t+s}}{P_{t+s}}\frac{\lambda_p}{(1-\lambda_p)\,\theta}\left(\tilde{p}_t\tilde{\Pi}^p_{t+s|t}\right)^{\frac{\lambda_p}{(1-\lambda_p)\theta}}\left(\frac{Y_{t+s}}{e^{z^a_{t+s}}}\right)^{\frac{1}{\theta}}\frac{1}{\tilde{P}_t}\right],
\end{aligned}$$

Denote $\bar{\lambda}_{t+s|t} \equiv \bar{\Lambda}_{t+s|t}P_{t+s}$. Also, denote $\lambda^j_{t+s|t} \equiv \beta^s_j d_{t+s|t}\left(c^j_{t+s}/c^j_t\right)^{-\sigma}$. Then, the above equation can be written as:

$$\tilde{p}_t = \left(\frac{\lambda_p}{\theta}\frac{\omega_u K^u_{p,t} + (1-\omega_u)\,K^r_{p,t}}{\omega_u F^u_{p,t} + (1-\omega_u)\,F^r_{p,t}}\right)^{\frac{(1-\lambda_p)\theta}{\theta-\lambda_p}} \tag{26}$$

where

$$F_{p,t}^{j} = \left(c_t^j\right)^{-\sigma} Y_t + \beta_j \xi_p E_t e^{z_{t+1}^b} (\tilde{\Pi}_{t+1|t}^p)^{\frac{1}{1-\lambda_p}} F_{p,t+1}^j, \tag{27}$$

$$K_{p,t}^{j} = \left(c_t^j\right)^{-\sigma} \left(\frac{Y_t}{e^{z_t^a}}\right)^{\frac{1}{\theta}} w_t + \beta_j \xi_p E_t e^{z_{t+1}^b} (\tilde{\Pi}_{t+1|t}^p)^{\frac{\lambda_p}{(1-\lambda_p)\theta}} K_{p,t+1}^j. \tag{28}$$

The aggregate price level evolves following

$$P_t = \left[\xi_p [(\pi_{t-1})^{\iota_p} (\pi)^{1-\iota_p} P_{t-1}]^{\frac{1}{1-\lambda_p}} + (1-\xi_p) \tilde{P}_t^{\frac{1}{1-\lambda_p}}\right]^{1-\lambda_p},$$

which can be written as

$$\tilde{p}_t = \left[\frac{1-\xi_p (\tilde{\Pi}_{t|t-1}^p)^{\frac{1}{1-\lambda_p}}}{1-\xi_p}\right]^{1-\lambda_p}. \tag{29}$$

The conditions, (26)-(29), summarize the price setting behavior of intermediate goods firms.

The aggregate nominal profits earned by intermediate goods firms are given by:

$$\Pi_t^m = \int_0^1 \left(P_t(i) Y_t(i) - W_t \left(\frac{Y_t(i)}{e^{z_t^a}}\right)^{\frac{1}{\theta}}\right) di \approx P_t Y_t - W_t \left(\frac{Y_t}{e^{z_t^a}}\right)^{\frac{1}{\theta}}.$$

Then, the aggregate real profits are given by $\pi_t^m = Y_t - w_t \left(Y_t / e^{z_t^a}\right)^{1/\theta}$.

A.1.4 Government

The govenment flow budget constraint is

$$(1+i_{t-1}) B_{t-1} + (1+\kappa P_{L,t}) B_{L,t-1} = B_t + P_{L,t} B_{L,t} + T_t,$$

where $T_t = \omega_u T_t^u + (1-\omega_u) T_t^r$. We assume that the lump-sum tax is imposed on households equaly so that $T_t^u = T_t^r = T_t$. Without loss of generality, we assume that the amount of short-term bonds issued is constant at $b_t = B_t / P_t = \bar{b}$. Then, the government flow budget constraint is reduced to:

$$(1+\kappa P_{L,t}) B_{L,t-1} = P_{L,t} B_{L,t} + T_t.$$

A.1.5 Central bank

As articulated in the main text, the central bank sets the nominal interest rate i_t, the shadow rate i_t^*, and the Taylor-rule-implied rate i_t^{Taylor} according to (1), (2), and (3). The central bank also launches QE following the simple rule (5).

A.1.6 Market clearing and equilibrium

As well as the good clearing market condition (19), there are market clearing conditions for labor, long-term government bonds, and short-term government bonds:

$$\omega_u h_t^u + (1-\omega_u)\, h_t^r = h_t, \tag{30}$$

$$\omega_u b_{L,t}^u + (1-\omega_u)\, b_{L,t}^r = b_{L,t}, \tag{31}$$

$$\omega_u b_t^u = b_t \tag{32}$$

Also, either the U-household's budget constraint or the R-household's budget constraint should be added as an equilibrium condition. Here the latter budget constraint is added:

$$c_t^r + P_{L,t} b_{L,t}^r = \left(\bar{R}_{L,t}/\pi_t\right) P_{L,t} b_{L,t-1}^r + w_t h_t^r - T_t^r/P_t + \Pi_t^r/P_t, \tag{33}$$

where

$$\frac{T_t^r}{P_t} = -\left(b_t + P_{L,t} b_{L,t}\right) + \frac{1+i_{t-1}}{\pi_t} b_{t-1} + \frac{1+\kappa P_{L,t}}{\pi_t} b_{L,t-1},$$
$$\frac{\Pi_t^r}{P_t} = y_t - w_t h_t.$$

The cost of trading long-term bonds, ζ_t, is specified as

$$\zeta_t = \bar{\zeta} \left(\frac{b_{L,t}}{b_L}\right)^{\rho_\zeta},$$

where $\bar{\zeta},\ \rho_\zeta > 0$. The cost is increasing in the amount of long-term bonds relative to its steady state value, $\zeta_t' > 0$.

The system of equations for the economy consists of 20 equations, (19)-(33), with the following endogenous variables:

$$c_t^u, c_t^r, h_t^u, h_t^r, h_t, b_{L,t}^u, b_{L,t}^r, b_{L,t}, b_t^u, y_t, w_t, i_t, i_t^*, i_t^T, Z_t, R_{L,t}, \pi_t, \tilde{p}_t, F_{p,t}^j, K_{p,t}^j.$$

A.2 Steady State and Calibration

The model period is quarterly. The inflation rate in steady state is equal to the targe rate of inflation, which is calibrated to be 2 percent, i.e. the gross inflation rate of $\pi = 1.005$. The marginal cost is unity in steady state $mc = 1$. We normalize $h = 1$ and set the value of ψ. The production function implies $y = 1$ in steady state. Then, in steady state, the amounts of short-term and long-term government bonds, $\bar{b}$ and $\bar{b}_L$, are interpreted as a debt-output ratio. Accordingly, we calibrate $\bar{b}$ and $\bar{b}_L$ to match debt-GDP ratios for short-term and long-term government bonds, respectively. The cost of trading long-term bonds is given by $\zeta = \zeta_1 \left(\bar{b}_L\right)^{\zeta_2}$. From equations (21) and (22), it is straight-forward to get $R = \pi/\beta_u$ and $R_L = (1+\zeta)\,\pi/\beta_u$. We set a target value of the short-term rate in steady state, say, $R = 3$ percent annually and set β_u as $\beta_u = \pi/(1+i)$. That for the R-household is set as $\beta_r = \pi/R_L$ to satisfy equation (24). The term premium $\zeta = \bar{\zeta}$ is derived as:

$$\bar{\zeta} = \frac{R_L}{1+i} - 1,$$

where R_L is the target long-term rate. In steady state, $\bar{R}_L = R_L$ and $P_L = \left(\bar{R}_L - \kappa\right)^{-1}$.The first-order approximation of the term premium around the steady state is given by:

$$\zeta_t - \zeta = \rho_\zeta \frac{\zeta}{b_L} \left(b_{L,t} - b_L\right).$$

The empirical literature on long-term asset purchases can be used to calibrate the parameter ρ_ζ. Suppose that the long-term government bonds to GDP ratio is 100 percent and the term premium is 100 bps in steady state, which are not far from those in 2012, a year preceding to the introduction of the Quantitative and Qualitative Easing, in Japan. Then ρ_ζ is calibrated as:

$$\rho_\zeta = \frac{\zeta_t - \zeta}{P_L b_{L,t} - P_L b_L} \frac{1}{100}.$$

Accoding to the empirical studies on long-term bond purchases by the Bank of Japan, such a purchase of 10 percent of GDP lowers the long-term rate by 3 $\sim$ 35bps. This implies $\rho_\zeta = 0.3 \sim 3.5$.

In the following, given the ratio, c^u/c^r, we compute the steady state. (Why is c^u/c^r indeterminate? Answer: in equilibrium, there is an initial condition for government debt holdings. The relative consumption depends on the initial condition, but here the condition is treated as an endogenous object. Here, instead of specifying the initial condition, we specify c^u/c^r and determine the households' government bonds holdings.) Denote ξ a fraction of labor supplied by U-households. We express labor supply by each household as follows:

$$\omega_u h^u = \xi h, \quad (1-\omega_u)\, h^r = (1-\xi)\, h$$

The relevant equations for computing the other variables in steady state are:

$$
\begin{aligned}
w &= \psi \left(c^u\right)^\sigma \left(\frac{\xi}{\omega_u} h\right)^{1/\nu}, \\
w &= \psi \left(c^r\right)^\sigma \left(\frac{1-\xi}{1-\omega_u} h\right)^{1/\nu}, \\
w &= \theta h^{\theta-1}, \\
y &= h^\theta, \\
y &= \omega_u c^u + \left(1-\omega_u\right) c^r, \\
\bar{b}_L &= \omega_u b_L^u + \left(1-\omega_u\right) b_L^r, \\
\bar{b} &= \omega_u b^u, \\
c^r + P_L b_L^r &= \frac{\bar{R}_L}{\pi} P_L b_L^r + w \frac{1-\xi}{1-\omega_u} h - \left[\left(\frac{1+i}{\pi} - 1\right) b + \left(\frac{\bar{R}_L}{\pi} - 1\right) P_L b_L\right] + \left(y - wh\right).
\end{aligned}
$$

Rearranging the above equation, we obtain

$$
c^u = \left[\frac{\theta}{\psi} \left(\frac{\omega_u}{\xi}\right)^{1/\nu} h^{-1/\nu-1+\theta}\right]^{1/\sigma}, \tag{34}
$$

$$
c^r = \left[\frac{\theta}{\psi} \left(\frac{1-\omega_u}{1-\xi}\right)^{1/\nu} h^{-1/\nu-1+\theta}\right]^{1/\sigma}, \tag{35}
$$

$$
h^\theta = \omega_u c^u + \left(1-\omega_u\right) c^r, \tag{36}
$$

$$
c^r = \left(\frac{\bar{R}_L}{\pi} - 1\right) P_L b_L^r + \left(\theta \frac{1-\xi}{1-\omega_u} + 1 - \theta\right) h^\theta - \left[\left(\frac{1+i}{\pi} - 1\right) b + \left(\frac{\bar{R}_L}{\pi} - 1\right) P_L b_L\right]. \tag{37}
$$

This is four equations with five unknowns, c^u, c^r, ξ, b_L^r and h. But we have a target value for c^u/c^r. Combining equations (34) and (35) yields:

$$
\xi = \frac{\omega_u}{1-\omega_u} \left[\left(\frac{c^u}{c^r}\right)^{\nu\sigma} + \frac{\omega_u}{1-\omega_u}\right]^{-1}.
$$

Substituting out c^u and c^r in equations (36) yields

$$
h^\theta = \omega_u \left[\frac{\theta}{\psi} \left(\frac{\omega_u}{\xi}\right)^{1/\nu} h^{-1/\nu-1+\theta}\right]^{1/\sigma} + \left(1-\omega_u\right) \left[\frac{\theta}{\psi} \left(\frac{1-\omega_u}{1-\xi}\right)^{1/\nu} h^{-1/\nu-1+\theta}\right]^{1/\sigma},
$$

or

$$
1 = \left[\omega_u \left(\frac{\omega_u}{\xi}\right)^{\frac{1}{\nu\sigma}} + \left(1-\omega_u\right) \left(\frac{1-\omega_u}{1-\xi}\right)^{\frac{1}{\nu\sigma}}\right] \left(\frac{\theta}{\psi}\right)^{\frac{1}{\sigma}}.
$$

Hence, ψ is given by:

$$\psi = \left[\omega_u \left(\frac{\omega_u}{\xi}\right)^{\frac{1}{\nu\sigma}} + (1-\omega_u)\left(\frac{1-\omega_u}{1-\xi}\right)^{\frac{1}{\nu\sigma}}\right]^{\sigma} \theta.$$

The value of long-term bonds held by the R-household, $P_L b_L^r$, is given by equation (37) as:

$$P_L b_L^r = \frac{\left(\frac{1-\omega_u}{1-\xi}\right)^{\frac{1}{\nu\sigma}} \left(\frac{\theta}{\psi}\right)^{\frac{1}{\sigma}} - \left(\theta\frac{1-\xi}{1-\omega_u} + 1 - \theta\right) + \left[\left(\frac{1+i}{\pi} - 1\right) b + \left(\frac{\bar{R}_L}{\pi} - 1\right) P_L b_L\right]}{\frac{\bar{R}_L}{\pi} - 1}.$$

A.3 Log-Linearized equations

A.3.1 Euler equation

In this section we derive equation (6). Log-linearlizing equations (21), (22), (24) and (19), we obtain

$$0 = E_t\left[-\sigma\left(\hat{c}^u_{t+1} - \hat{c}^u_t\right) + \hat{\imath}_t - \hat{\pi}_{t+1} + z^b_{t+1}\right], \tag{38}$$

$$\frac{\zeta}{1+\zeta}\hat{\zeta}_t = E_t\left[-\sigma\left(\hat{c}^u_{t+1} - \hat{c}^u_t\right) + \hat{R}_{L,t+1} - \hat{\pi}_{t+1} + z^b_{t+1}\right], \tag{39}$$

$$0 = E_t\left[-\sigma\left(\hat{c}^r_{t+1} - \hat{c}^r_t\right) + \hat{R}_{L,t+1} - \hat{\pi}_{t+1} + z^b_{t+1}\right], \tag{40}$$

$$\hat{y}_t = \frac{\omega_u c^u}{y}\hat{c}^u_t + \frac{(1-\omega_u)\, c^r}{y}\hat{c}^r_t. \tag{41}$$

From equation (41), we obtain:

$$\hat{c}^u_t = \frac{y}{\omega_u c^u}\left\{\hat{y}_t - \frac{(1-\omega_u)\, c^r}{y}\hat{c}^r_t\right\}.$$

Subtracting $\hat{c}^u_{t+1}$ from $\hat{c}^u_t$ yields:

$$\begin{aligned}
\hat{c}^u_{t+1} - \hat{c}^u_t &= \frac{y}{\omega_u c^u}\left\{\hat{y}_{t+1} - \hat{y}_t - \frac{(1-\omega_u)\, c^r}{y}\left(\hat{c}^r_{t+1} - \hat{c}^r_t\right)\right\}, \\
&= \frac{y}{\omega_u c^u}\left\{\hat{y}_{t+1} - \hat{y}_t - \frac{(1-\omega_u)\, c^r}{y}\frac{\left(\hat{R}_{L,t+1} - \hat{\pi}_{t+1} + z^b_{t+1}\right)}{\sigma}\right\}, \qquad (42)
\end{aligned}$$

where equation (40) was used in the second equality. Substituting equation (42) into equation (38) yields:

$$
\begin{aligned}
0 &= E_t\left[-\sigma\left(\hat{c}^u_{t+1}-\hat{c}^u_t\right)+\hat{\imath}_t-\hat{\pi}_{t+1}+z^b_{t+1}\right], \\
&= E_t\left[-\frac{\sigma y}{\omega_u c^u}\left(\hat{y}_{t+1}-\hat{y}_t\right)+\frac{\sigma y}{\omega_u c^u}\frac{\left(1-\omega_u\right)c^r}{y}\frac{\left(\hat{R}_{L,t+1}-\hat{\pi}_{t+1}+z^b_{t+1}\right)}{\sigma}\right. \\
&\left.+\hat{\imath}_t-\hat{\pi}_{t+1}+z^b_{t+1}\right], \\
0 &= E_t\left[-\sigma\left(\hat{y}_{t+1}-\hat{y}_t\right)+\frac{\left(1-\omega_u\right)c^r}{y}\hat{R}_{L,t+1}+\frac{\omega_u c^u}{y}\hat{\imath}_t-\hat{\pi}_{t+1}+z^b_{t+1}\right],
\end{aligned}
\tag{43}
$$

where equation (19) in steady state was used in the third equality. Also, substituing equation (42) into equation (39) yields:

$$
\begin{aligned}
\frac{\zeta}{1+\zeta}\hat{\zeta}_t &= E_t\left[-\sigma\left(\hat{c}^u_{t+1}-\hat{c}^u_t\right)+\hat{R}_{L,t+1}-\hat{\pi}_{t+1}\right] \\
&= E_t\left[-\sigma\frac{y}{\omega_u c^u}\left\{\hat{y}_{t+1}-\hat{y}_t-\frac{\left(1-\omega_u\right)c^r}{y}\frac{\left(\hat{R}_{L,t+1}-\hat{\pi}_{t+1}+z^b_{t+1}\right)}{\sigma}\right\}\right. \\
&\left.+\hat{R}_{L,t+1}-\hat{\pi}_{t+1}+z^b_{t+1}\right], \\
&= E_t\left[-\frac{\sigma y}{\omega_u c^u}\left(\hat{y}_{t+1}-\hat{y}_t\right)+\frac{y}{\omega_u c^u}\left(\hat{R}_{L,t+1}-\hat{\pi}_{t+1}+z^b_{t+1}\right)\right],
\end{aligned}
$$

or

$$
\begin{aligned}
E_t\left(\hat{R}_{L,t+1}-\hat{\pi}_{t+1}\right) &= \sigma E_t\left(\hat{y}_{t+1}-\hat{y}_t\right)-E_t\left(z^b_{t+1}\right)+\frac{\omega_u c^u}{y}\frac{\zeta}{1+\zeta}\hat{\zeta}_t \\
&= \sigma E_t\left(\hat{y}_{t+1}-\hat{y}_t\right)-E_t\left(z^b_{t+1}\right)+\frac{\omega_u c^u}{y}\frac{\zeta}{1+\zeta}\rho_\zeta\widehat{b}_{L,t}.
\end{aligned}
\tag{44}
$$

Combining equations (43) and (44) yields:

$$
\begin{aligned}
0 &= E_t\left[-\sigma\left(\hat{y}_{t+1}-\hat{y}_t\right)+\frac{\left(1-\omega_u\right)c^r}{y}\hat{R}_{L,t+1}+\frac{\omega_u c^u}{y}\hat{\imath}_t-\hat{\pi}_{t+1}+z^b_{t+1}\right] \\
&= E_t\left[-\sigma\left(\hat{y}_{t+1}-\hat{y}_t\right)+\frac{\left(1-\omega_u\right)c^r}{y}\left(\sigma\left(\hat{y}_{t+1}-\hat{y}_t\right)-z^b_{t+1}+\frac{\omega_u c^u}{y}\frac{\zeta}{1+\zeta}\rho_\zeta\widehat{b}_{L,t}+\hat{\pi}_{t+1}\right)\right. \\
&\left.+\frac{\omega_u c^u}{y}\hat{\imath}_t-\hat{\pi}_{t+1}+z^b_{t+1}\right] \\
&= E_t\left[-\frac{\omega_u c^u\sigma}{y}\left(\hat{y}_{t+1}-\hat{y}_t\right)+\frac{\omega_u c^u}{y}\hat{\imath}_t-\frac{\omega_u c^u}{y}\left(\hat{\pi}_{t+1}-z^b_{t+1}\right)+\frac{\left(1-\omega_u\right)c^r}{y}\frac{\omega_u c^u}{y}\frac{\zeta}{1+\zeta}\rho_\zeta\widehat{b}_{L,t}\right], \\
0 &= E_t\left[-\sigma\left(\hat{y}_{t+1}-\hat{y}_t\right)+\hat{\imath}_t-\hat{\pi}_{t+1}+z^b_{t+1}+\frac{\left(1-\omega_u\right)c^r}{y}\frac{\zeta}{1+\zeta}\rho_\zeta\widehat{b}_{L,t}\right],
\end{aligned}
$$

or

$$\begin{aligned}
\hat{y}_t &= E_t\hat{y}_{t+1} - \frac{1}{\sigma}\left(\hat{\imath}_t - E_t\hat{\pi}_{t+1} + E_t z_{t+1}^b\right) - \frac{1}{\sigma}\frac{(1-\omega_u)\,c^r}{y}\frac{\zeta}{1+\zeta}\rho_\zeta \widehat{b}_{L,t} \\
&= E_t\hat{y}_{t+1} - \frac{1}{\sigma}\left(\hat{\imath}_t - E_t\hat{\pi}_{t+1}\right) - \frac{1}{\sigma}\frac{(1-\omega_u)\,c^r}{y}\frac{\zeta}{1+\zeta}\rho_\zeta \widehat{b}_{L,t} - \frac{\rho_b}{\sigma} z_t^b.
\end{aligned}$$

This equation shows that a central bank's government bond purchase – a decrease in $\hat{b}_{L,t}$ – stimulates output, given $E_t\hat{y}_{t+1}$ and the real rate $\hat{R}_t - E_t\hat{\pi}_{t+1}$. Similarly, the negative preference shock, $z_t^b < 0$, discounts future utility more and stimulates today's consumption and output. This completes the derivation of equation (6) where

$$\begin{aligned}
\chi_b &= \frac{1}{\sigma}\frac{(1-\omega_u)\,c^r}{y}\frac{\zeta}{1+\zeta}\rho_\zeta > 0, \\
\chi_z &= \frac{\rho_b}{\sigma} > 0.
\end{aligned}$$

As shown in the introduction, λ in equation (7) summarizes the efficacy of the unconventional monetary policy (i.e. long-term government bond purchases), and it is given by $\lambda^* = \sigma\chi_b\gamma/(1+i)$. Specifically:

$$\lambda^* = \frac{(1-\omega_u)\,c^r}{y}\frac{\zeta}{1+\zeta}\frac{1}{1+i}\rho_\zeta\gamma.$$

The case of $\lambda^* = 1$ corresponds to the "fully effective" unconventional monetary policy, which makes the zero lower bound irrelevant. Such a case can be achieved, e.g. when the central bank responds to the shadow rate aggressively enough to satisfty:

$$\gamma = \left[\frac{(1-\omega_u)\,c^r}{y}\frac{\zeta}{1+\zeta}\frac{1}{1+i}\rho_\zeta\right]^{-1}.$$

Or it can be achieved, e.g. when the elasticity of the premium ξ with respect to the amount of long-term bonds is so high that it is given by:

$$\rho_\xi = \left[\frac{(1-\omega_u)\,c^r}{y}\frac{\zeta}{1+\zeta}\frac{1}{1+i}\gamma\right]^{-1}.$$

A.3.2 Phillips curve

The Phillips curve can be derived from equations (26)-(29). Log-linearizing equation (29) yields:

$$\widehat{\tilde{p}}_t = -\frac{\xi_p}{1-\xi_p}\widehat{\tilde{\Pi}}^p_{t|t-1}, \tag{45}$$

where

$$\widehat{\widetilde{\Pi}}^{p}_{t|t-1} = (1-\nu_p)\,\hat{\pi}_{t-1} - \hat{\pi}_t.$$

Log-linearzing equation (26) yields:

$$\begin{aligned}\frac{\theta-\lambda_p}{(1-\lambda_p)\,\theta}\widehat{\tilde{p}}_t = &\frac{\omega_u K_p^u}{\omega_u K_p^u + (1-\omega_u)\,K_p^r}\hat{K}^u_{p,t} + \frac{(1-\omega_u)\,K_p^r}{\omega_u K_p^u + (1-\omega_u)\,K_p^r}\hat{K}^r_{p,t} \\ &- \frac{\omega_u F_p^u}{\omega_u F_p^u + (1-\omega_u)\,F_p^r}\hat{F}^u_{p,t} - \frac{(1-\omega_u)\,F_p^r}{\omega_u F_p^u + (1-\omega_u)\,F_p^r}\hat{F}^r_{p,t}. \end{aligned} \tag{46}$$

Combining equations (45) and (46) leads to:

$$\begin{aligned}-\frac{\xi_p}{1-\xi_p}\frac{\theta-\lambda_p}{(1-\lambda_p)\,\theta}\left[(1-\nu_p)\,\hat{\pi}_{t-1} - \hat{\pi}_t\right] = &\frac{\omega_u K_p^u}{\omega_u K_p^u + (1-\omega_u)\,K_p^r}\hat{K}^u_{p,t} + \frac{(1-\omega_u)\,K_p^r}{\omega_u K_p^u + (1-\omega_u)\,K_p^r}\hat{K}^r_{p,t} \\ &- \frac{\omega_u F_p^u}{\omega_u F_p^u + (1-\omega_u)\,F_p^r}\hat{F}^u_{p,t} - \frac{(1-\omega_u)\,F_p^r}{\omega_u F_p^u + (1-\omega_u)\,F_p^r}\hat{F}^r_{p,t}. \end{aligned} \tag{47}$$

Log-linearizing equation (27) and (28) yields:

$$\begin{aligned}\hat{F}^j_{p,t} &= (1-\beta\xi_p)\left(-\sigma\hat{c}^j_t + \hat{Y}_t\right) + \beta\xi_p E_t\left(z^b_{t+1} + \frac{1}{1-\lambda_p}\widehat{\widetilde{\Pi}}^{p}_{t+1|t} + \hat{F}^j_{p,t+1}\right), \\ \hat{K}^j_{p,t} &= (1-\beta\xi_p)\left(-\sigma\hat{c}^j_t + \frac{1}{\theta}\hat{Y}_t - \frac{1}{\theta}z^a_t + \hat{w}_t\right) + \beta\xi_p E_t\left(z^b_{t+1} + \frac{\lambda_p}{(1-\lambda_p)\,\theta}\widehat{\widetilde{\Pi}}^{p}_{t+1|t} + \hat{K}^j_{p,t+1}\right),\end{aligned}$$

for $j = r$ and u. The terms involving $\hat{F}^u_{p,t}$ and $\hat{F}^r_{p,t}$ in equation (47) is calculated as follows.

$$\begin{aligned}&\frac{\omega_u F_p^u}{\omega_u F_p^u + (1-\omega_u)\,F_p^r}\hat{F}^u_{p,t} + \frac{(1-\omega_u)\,F_p^r}{\omega_u F_p^u + (1-\omega_u)\,F_p^r}\hat{F}^r_{p,t} \\ &= (1-\beta\xi_p)\left(-\sigma\frac{\omega_u F_p^u\hat{c}^u_t + (1-\omega_u)\,F_p^r\hat{c}^r_t}{\omega_u F_p^u + (1-\omega_u)\,F_p^r} + \hat{Y}_t\right) \\ &+ \beta\xi_p E_t\left(z^b_{t+1} + \frac{1}{1-\lambda_p}\widehat{\widetilde{\Pi}}^{p}_{t+1|t} + \frac{\omega_u F_p^u}{\omega_u F_p^u + (1-\omega_u)\,F_p^r}\hat{F}^u_{p,t+1} + \frac{(1-\omega_u)\,F_p^r}{\omega_u F_p^u + (1-\omega_u)\,F_p^r}\hat{F}^r_{p,t+1}\right).\end{aligned}$$

Similarly, the terms involving $\hat{K}^u_{p,t}$ and $\hat{K}^r_{p,t}$ in equation (47) is calculated as:

$$\begin{aligned}&\frac{\omega_u K_p^u}{\omega_u K_p^u + (1-\omega_u)\,K_p^r}\hat{K}^u_{p,t} + \frac{(1-\omega_u)\,K_p^r}{\omega_u K_p^u + (1-\omega_u)\,K_p^r}\hat{K}^r_{p,t} \\ &= (1-\beta\xi_p)\left(-\sigma\frac{\omega_u K_p^u\hat{c}^u_t + (1-\omega_u)\,K_p^r\hat{c}^r_t}{\omega_u K_p^u + (1-\omega_u)\,K_p^r} + \frac{1}{\theta}\hat{Y}_t - \frac{1}{\theta}z^a_t + \hat{w}_t\right) \\ &+ \beta\xi_p E_t\left(z^b_{t+1} + \frac{\lambda_p}{(1-\lambda_p)\,\theta}\widehat{\widetilde{\Pi}}^{p}_{t+1|t} + \frac{\omega_u K_p^u}{\omega_u K_p^u + (1-\omega_u)\,K_p^r}\hat{K}^u_{p,t+1} + \frac{(1-\omega_u)\,K_p^r}{\omega_u K_p^u + (1-\omega_u)\,K_p^r}\hat{K}^r_{p,t+1}\right).\end{aligned}$$

Let the right-hand-side of equation (47) denote as $\hat{X}_t$. Then, using the above relationships just derived, $\hat{X}_t$ can be written as:

$$\hat{X}_t = (1-\beta\xi_p)\left[\left(\frac{1}{\theta}-1\right)\hat{Y}_t - \frac{1}{\theta}z_t^a + \hat{w}_t\right] \\ + \beta\xi_p E_t\left(-\frac{\lambda_p-\theta}{(\lambda_p-1)\theta}\widehat{\tilde{\Pi}}^p_{t+1|t} + \hat{X}_{t+1}\right).$$

Because $\hat{X}_t$ is the right-hand-side of equation (47), equation (47) can be written as:

$$-\frac{\xi_p}{1-\xi_p}\frac{\lambda_p-\theta}{(\lambda_p-1)\theta}\left[(1-\nu_p)\hat{\pi}_{t-1}-\hat{\pi}_t\right] = (1-\beta\xi_p)\left[\left(\frac{1}{\theta}-1\right)\hat{Y}_t - \frac{1}{\theta}z_t^a + \hat{w}_t\right] \\ + \beta\xi_p E_t\left(-\frac{\lambda_p-\theta}{(\lambda_p-1)\theta}\widehat{\tilde{\Pi}}^p_{t+1|t} - \frac{\xi_p}{1-\xi_p}\frac{\lambda_p-\theta}{(\lambda_p-1)\theta}\left[(1-\nu_p)\hat{\pi}_t-\hat{\pi}_{t+1}\right]\right),$$

or

$$\hat{\pi}_t = \frac{\xi_p(1-\nu_p)}{(\xi_p+1-\upsilon_p)}\hat{\pi}_{t-1} + \frac{(1-\beta\xi_p)(1-\xi_p)(\lambda_p-1)\theta}{(\lambda_p-\theta)(\xi_p+1-\upsilon_p)}\left[\left(\frac{1}{\theta}-1\right)\hat{Y}_t - \frac{1}{\theta}z_t^a + \hat{w}_t\right] + \frac{\beta\xi_p}{(\xi_p+1-\upsilon_p)}E_t\hat{\pi}_{t+1}.$$

From equations (20) and (23), the wage $\hat{w}_t$ can be written as:

$$\begin{aligned}\hat{w}_t &= \omega_u\left(\sigma\hat{c}_t^u + \frac{1}{\nu}\hat{h}_t^u\right) + (1-\omega_u)\left(\sigma\hat{c}_t^r + \frac{1}{\nu}\hat{h}_t^r\right),\\ &= \sigma\hat{Y}_t + \frac{1}{\nu}\hat{h}_t,\\ &= \left(\sigma + \frac{1}{\nu\theta}\right)\hat{Y}_t - \frac{1}{\nu\theta}z_t^a,\end{aligned}$$

where the market clearing conditions (19) and (30) were used in the second equality and the production function (25) was used in the third equality. In our calibration, $c^u = c^r$ so that the second equality holds. Using the expression for $\hat{w}_t$ the Phillips curve can be written as

$$\hat{\pi}_t = \frac{\xi_p(1-\nu_p)}{(\xi_p+1-\upsilon_p)}\hat{\pi}_{t-1} \\ + \frac{(1-\beta\xi_p)(1-\xi_p)(\lambda_p-1)\theta}{(\lambda_p-\theta)(\xi_p+1-\upsilon_p)}\left[\frac{\nu+\nu\theta(\sigma-1)+1}{\nu\theta}\hat{Y}_t - \frac{1+\nu}{\nu\theta}z_t^a\right] + \frac{\beta\xi_p}{(\xi_p+1-\upsilon_p)}E_t\hat{\pi}_{t+1}. \tag{48}$$

In the case of no price indexation to the past inflation rate and a linear production function, that is, in the case of $\nu_p = 1$ and $\theta = 1$, the Phillips curve is collapsed to the standard form:

$$\hat{\pi}_t = \frac{(1-\beta\xi_p)(1-\xi_p)}{\xi_p}\left(\left(\sigma+\frac{1}{\nu}\right)\hat{Y}_t - \frac{1+\nu}{\nu}z_t^a\right) + \beta E_t\hat{\pi}_{t+1}.$$

A.3.3 System of log-linearized equations

This model has a closed system of four equations with six variables $\left\{y_t, \pi_t, i_t, i_t^*, i_t^T, Z_t\right\}$. The equations consist of (7) and (48) as well as a set of interest rate rules (1)-(3).

B Data description

We construct our quarterly data by taking averages of their monthly counterparts. The inflation rate is computed from the implicit price deflator (GDPDEF) as $\pi_t = 400 * ln(\frac{P_t}{P_{t-1}})$. The output gap is calculated as $100\% * \frac{GDPC1-GDPPOT}{GDPPOT}$, where GDPC1 is the series for the US real GDP and GDPPOT is the US real potential GDP. The long-term interest rate comes from the 10-year Treasury constant maturity rate (GS10). All these series are from the FRED database and are available from 1949Q1 to 2019Q1. The effective federal funds rate (FEDFUNDS) is available from 1954Q3 to 2019Q1.

Our money growth data are computed from 12 alternative indicators as listed in Table 13 as $m_t = 400 * ln(\frac{M_t}{M_{t-1}})$, where M_t is the particular money series considered. All M_t values are quarterly and computed by taking averages of their corresponding monthly values. The traditional monetary aggregates (MB, M1, M2, M2M, MZM), excess reserves, and securities held outright are from the FRED database. The Divisia monetary aggragates (DIVM1, DIVM2, DIVM2M, DIVMZM, DIVM4) are from the Center for Financial Stability Divisia database.

Table 13: Monetary Aggregates Data used in the Model

Monetary Aggregate (M_t)	Series Code in the Corresponding Database	Available Sample Periods
Monetary Base (MB)	MBSL	1948Q1-2019Q1
M1	M1SL	1959Q2-2019Q1
M2	M2SL	1959Q2-2019Q1
M2M	M2MSL	1959Q2-2019Q1
MZM	MZMSL	1959Q2-2019Q1
Excess Reserves	EXCSRESNS	1984Q3-2019Q1
Securities Held Outright	WSECOUT	1989Q3-2019Q1
Divisia M1 (DIVM1)	Divisia M1	1967Q2-2019Q1
Divisia M2 (DIVM2)	Divisia M2	1967Q2-2019Q1
Divisia M2M (DIVM2M)	Divisia M2M	1967Q2-2019Q1
Divisia MZM (DIVMZM)	Divisia MZM	1967Q2-2019Q1
Divisia M4 (DIVM4)	DM4	1967Q2-2019Q1