This paper presents a model of quantitative easing (QE) at the zero lower bound (ZLB) on the short-term nominal interest rate. QE, which reduces the maturity of government debt, is effective at the ZLB because it generates expectations of future monetary expansion in a time-consistent equilibrium. Numerical experiments show that this effect can be substantial.

JEL Classification: E31, E52, E58, E61, E63

Keywords: government debt maturity; quantitative easing; time-consistent equilibrium; signaling; zero lower bound
The problem with Quantitative Easing is it works in practice, but it doesn’t work in theory.
Ben Bernanke, chairman of the Federal Reserve

1 Introduction

Starting at the onset of the economic crisis in 2008, the Federal Reserve expanded its balance sheet by large amounts, on the order of $3 trillion, mostly under the heading of quantitative easing (QE). At its peak in 2012, the cumulative amount of QE was equivalent to about 20% of annual GDP. The enormous scale of this policy has largely been explained by the fact that the Federal Reserve was unable to cut the federal funds rate further than it already had because of the zero lower bound (ZLB) on the short-term nominal interest rate. Meanwhile, high unemployment, slow growth, and low inflation called for further stimulus measures.

Many commentators argued that QE in the US prevented a much stronger contraction during the Great Recession and that QE was a key reason why the US recovered more rapidly than other countries. As pointed out by Federal Reserve chairman Ben Bernanke, however, while this might be true in practice, a coherent theoretical rationale has been hard to formulate. This paper formulates such a rationale for QE: it works because it allows the central bank to credibly commit, in a time-consistent equilibrium, to expansionary future policy at the ZLB. The paper accounts for QE in theory and shows numerical examples in which the effect is considerable.

What is QE? Our paper defines it as a policy in which a central bank buys long-term government debt with money. Since the nominal interest rate was zero when QE was implemented in the US, it makes no difference whether QE is implemented by printing money (or more precisely creating bank reserves) or by issuing short-term government debt: both types of assets are government-issued papers that yield a zero interest rate. From the perspective of the government as a whole, QE at the ZLB can then be thought of as reducing the maturity of outstanding government debt held by the public, as the government is replacing long-term debt with short-term debt. The main finding of this paper is that shortening the maturity of government debt is an effective way to commit to a future monetary expansion, an especially valuable tool when the central bank’s policy rate is constrained by the ZLB.

The main goal of QE in the US was to reduce long-term interest rates at the ZLB and thereby stimulate the economy. Indeed, several empirical studies find evidence of a reduction in long-term interest rates following these policy interventions by the Federal Reserve (see, for

---

As we emphasize later, our paper is thus best interpreted as providing a model for what has been called QE2 and QE3, as those policies were largely focused on purchases of long-term government bonds.
example, Gagnon et al. (2011), Krishnamurthy and Vissing-Jorgensen (2011), Hamilton and Wu (2012), Swanson and Williams (2013), and Bauer and Rudebusch (2013). An “irrelevance result” of this form was first stated by Wallace (1981). The Wallace irrelevance result was extended by Eggertsson and Woodford (2003) to a model with sticky prices and an explicit ZLB. Those results illustrate that, absent certain restrictions on asset trading that prevent arbitrage, a change in the relative supplies of various assets in the hands of the private sector has no effect on equilibrium quantities.

As pointed out by Eggertsson and Woodford (2003), and further illustrated in Woodford (2012) in the context of the financial crisis, QE may be effective not only because it reduces risk premia due to limits on arbitrage and/or differences in liquidity between private and public bonds. QE can also reduce long-term interest rates if it communicates to the private sector that the central bank will keep the short-term interest rates low once the ZLB is no longer a constraint; that is, it signals a change in the policy rule taken as given in Eggertsson and Woodford (2003). Krishnamurthy and Vissing-Jorgensen (2011) and Bauer and Rudebusch (2013) find strong evidence in support of such a “signaling channel” of various QE programs; that is, the programs generated expectations of low future federal funds rates.

The main contribution of this paper is to provide a formal theoretical model showing how

2 For example, Gagnon et al. (2011) estimate that the 2009 program, worth $1.75 trillion, reduced long-term interest rates by 58 basis points, while Krishnamurthy and Vissing-Jorgensen (2011) estimate that the 2010 program, worth $600 billion, reduced long-term interest rates by 33 basis points.

3 This may have motivated Ben Bernanke’s comment cited in the epigraph.
QE generates a credible commitment to future monetary expansion in a standard general equilibrium model. We formulate a Markov perfect equilibrium (MPE) as a game played by the government and the private sector. In the MPE, agents use the natural state variables of the game to predict the behavior of future governments. QE changes the dynamics of the endogenous state variables of the game, thereby affecting expectations about the future path of the policy instrument in a time-consistent equilibrium.

This paper proposes the rollover incentive as the main force that makes the government more keen to set lower rates the more short-term debt it issues. This is the principal effect of QE. The rollover incentive is best explained via a simple example that glosses over several subtleties discussed in detail in the body of the paper. Consider yourself as evaluating two mortgage contracts: a thirty-year loan with fixed interest rates and a loan contract with floating interest rates that are determined monthly. Now consider your incentives if you get the opportunity to set the federal funds rate. If you have a thirty-year loan, then your own interest rate is unaffected by an increase in the federal funds rate, leaving your interest costs unchanged. But if you have a flexible-rate loan, you have much to lose by increasing interest rates, for it would directly increase your interest payments. You are now rolling over your debt from one period to the next with higher interest rates. Accordingly, the more short-term debt you hold, the less willing you are to raise the federal funds rate.

A better-known effect is termed the balance sheet incentive: interest rate policy can change inflation and thus can change the real value of outstanding nominal debt. The balance sheet incentive, however, as we show below, is independent of the maturity structure of debt (in sharp contrast to the rollover incentive).

One objection to the main thesis of the paper is that many argue that fiscal considerations do not play a fundamental role in the decision making of the Federal Reserve. While the main analysis assumes a consolidated government budget constraint, we show that under certain conditions, the model can equivalently be interpreted as referring to an independent central bank that cares about it’s own balance sheet losses. We furthermore provide anecdotal evidence in support of this interpretation: Based upon recently declassified documents from the deliberations of the Federal Reserve Open Market Committee, we show that the Federal Reserve closely tracked possible balance sheet losses associated with its QE policy.

Our paper is organized as follows. Section 2 considers a three-period model that can be solved without resorting to any approximations. First, we show the deflation bias of discretionary monetary policy at the ZLB: in the presence of a deflationary shock, a central

---

4As we discuss later, the rollover incentive operates under fully flexible prices, fully fixed prices, or anywhere in between. In this example, a key subtlety that is glossed over is how to properly define the state variables of the game played by successive governments—namely, the maturity value of debt and its composition—and which asset-pricing conditions are constraints faced by the government in a given period in an MPE.
bank would like to commit to lower future policy rates but, because of a fundamental credibility problem, cannot do so. The paper then shows that the government can issue or buy a mix of long- and short-term debt to make the optimal monetary policy incentive compatible. This is the role QE plays in the paper. Section 3 considers an infinite-horizon model in a quantitative analysis and shows a substantial role for QE. Section 4 considers extensions, presents empirical evidence, and discusses the relationship of the results to the literature. Section 5 concludes.

2 A Simple Model

This section illustrates the role of the maturity of nominal government debt in affecting expectations about future policy in a time-consistent equilibrium, using minimal modeling ingredients.

2.1 Environment

2.1.1 Private Sector

There are three periods: 0, 1, and 2 (labeled short, medium, and long run respectively). A representative household solves

$$\max_{C_t, H_t, L_t, B_t} \mathbb{E}_0 \sum_{t=0}^2 \beta^t [\log C_t - \chi H_t] \xi_t$$

subject to the flow budget constraints

$$P_0 C_0 + B_0 + L_0 = B_{-1} + \int_0^1 Z_0(i) di + N_0 H_0 - P_0 T_0,$$

$$P_1 C_1 + B_1 = (1 + i_0) B_0 + \int_0^1 Z_1(i) di + N_1 H_1 - P_1 T_1,$$

$$P_2 C_2 = (1 + i_1) B_1 + (1 + R_0) L_0 + \int_0^1 Z_2(i) di + N_2 H_2 - P_2 T_2,$$

where $C_t$ is consumption, $H_t$ hours, and $\xi_t$ a preference shock. $B_{-1}$ is the initial wealth the household holds in period 0; $B_0$ and $B_1$ are one-period nominal risk-free bonds; and $L_0$ represents long-term nominal bonds issued in period 0 and repaid in period 2. The short-term interest rates in periods 0 and 1 are $i_0$ and $i_1$, while $R_0$ is the long-term interest rate in period 0. $N_t$ is the wage rate, $P_t$ the price index, $Z_t(i)$ firms’ profits (where $i$ indexes firms), and $T_t$ taxes. The model abstracts from money but imposes the ZLB directly.\(^5\)

---

\(^5\)Eggertsson and Woodford (2003) motivate this abstraction in more detail.
Consumption $C_t$ is a Dixit-Stiglitz aggregator, $C_t \equiv \left[ \int_0^1 c_t(i)^{\theta-1} \frac{\partial}{\partial i} \right]^{\theta} \pi di$, where $\theta > 1$ is the elasticity of substitution across goods with $i$ indexing varieties of goods. For each variety, there is a single firm, so that $y_t(i) = c_t(i)$. The firm has a linear production function, $y_t(i) = h_t(i)$, where $h_t(i)$ is labor. The firm maximizes profits

$$\max_{p_t(i),y_t(i)} \{ (1 + \tau)p_t(i)y_t(i) - N_t y_t(i) \}$$

subject to $y_t(i) = \left( \frac{p_t(i)}{P_t} \right)^{-\theta} Y_t$, where $Y_t = C_t$ is aggregate output and $\tau$ denotes a production subsidy. In periods 0 and 1, a fraction $\gamma$ of firms set prices flexibly while the remaining fraction index their prices to the past aggregate price index. In the terminal period, all prices are flexible. Further details are provided in Appendix A.1.

We abstract from government spending but assume that the government needs to repay the debt inherited in period 0, $B_{-1}$. As in Barro (1979), taxes are not one-to-one transfers of purchasing power from individuals to the government. Instead, they entail some collection costs or indirect misallocation costs. More concretely, the “production” of government revenues, $T_t$, requires labor input $f(T_t)$, where $f(.)$ is increasing and convex in $T_t$. This, together with the fact that production is linear in labor, implies that total hours can be expressed as

$$H_t = Y_t \Delta_t + f(T_t),$$


6Here, $T_t$ denotes tax revenues net of government workers’ salaries.

7In the quantitative model, utility is not assumed to be linear in labor, but the implications for taxes are the same. The reason is that taxes only affect allocations up to a second order in the quantitative model.

2.1.2 The Government’s Problem: A Maturity Value of Debt Characterization

The model can be summarized in three steps (details are provided in Appendix A). We obtain the government’s objective via a second-order approximation of household utility (1):

$$-\frac{1}{2} \mathbb{E}_0 \sum_{t=0}^{2} \beta^t \left[ (\Pi_t - 1)^2 + \lambda_y (Y_t - \bar{Y})^2 + \lambda_T (T_t)^2 + t.i.p. \right], \quad (2)$$

$$-\frac{1}{2} \mathbb{E}_0 \sum_{t=0}^{2} \beta^t \left[ (\Pi_t - 1)^2 + \lambda_y (Y_t - \bar{Y})^2 + \lambda_T (T_t)^2 + t.i.p. \right],$$

$$-\frac{1}{2} \mathbb{E}_0 \sum_{t=0}^{2} \beta^t \left[ (\Pi_t - 1)^2 + \lambda_y (Y_t - \bar{Y})^2 + \lambda_T (T_t)^2 + t.i.p. \right], \quad (2)$$
where we have normalized the weight on inflation to 1 and the weights $\lambda_y$ and $\lambda_T$ are functions of the structural parameters, whose derivation is relegated to Appendix A.2. The term including taxes appears because of the taxation costs discussed in Section 2.1.1. The cost of inflation is due to inefficient price dispersion.

The optimality conditions for the household’s and firm’s maximization problems are

$$\frac{1}{1 + i_t} = \beta \mathbb{E}_t \left[ \frac{Y_t \Pi_t^{-1} \xi_{t+1}}{Y_{t+1} \Pi_{t+1} \xi_t} \right] \quad \text{for } t = 0, 1, \quad (3)$$

$$\frac{1}{1 + R_0} = \beta^2 \mathbb{E}_0 \left[ \frac{Y_0 \Pi_0^{-1} \Pi_1^{-1} \xi_2}{Y \xi_0} \right], \quad (4)$$

$$i_t \geq 0 \text{ for } t = 0, 1, \quad (5)$$

$$(\Pi_t - 1) = \kappa(Y_t - \bar{Y}) \text{ for } t = 0, 1, \text{ and } Y_2 = \bar{Y}. \quad (6)$$

The household budget constraints, together with market clearing, imply the following flow government budget constraints:

$$x_0 \frac{w_0}{1 + R_0} + (1 - x_0) \frac{w_0}{1 + i_0} = \frac{w_{-1}}{\Pi_0} - T_0, \quad (7)$$

$$W_1 = \frac{w_0}{\Pi_1} + i_1(1 - x_0) \frac{w_0}{\Pi_1} - (1 + i_1)T_1, \quad (8)$$

$$0 = \frac{W_1}{\Pi_2} - T_2. \quad (9)$$

Here, $w_0 \equiv b_0(1 + i_0) + l_0(1 + R_0)$, $x_0 \equiv \frac{l_0(1 + R_0)}{w_0(1 + i_0) + l_0(1 + R_0)}$, and $W_1 \equiv w_1 + x_0 \frac{w_0}{\Pi_1}$, where $w_1 \equiv (1 + i_1)b_1$. The lower-case letters refer to the real value of the corresponding nominal debt; that is, $b_0 \equiv \frac{B_0}{P_0}$, $b_1 \equiv \frac{B_1}{P_1}$, $l_0 \equiv \frac{L_0}{P_0}$, and $b_{-1} \equiv w_{-1} \equiv \frac{B_{-1}}{P_{-1}}$.

The government’s problem is to maximize (2) by choosing $\{\Pi_t, T_t\}$ for $t = 0, 1, 2$, $\{i_t, Y_t\}$ for $t = 0, 1$, and $\{w_0, x_0, R_0, W_1\}$, subject to (3)-(9), taking as given the initial value of debt $w_{-1}$. We next discuss the government’s problem and interpret the policy variables.
2.1.3 The Government’s Problem: A Discussion

The most important feature of the characterization of the government’s problem in the last subsection is that its budget constraint in period 1, (8), is written in terms of the maturity value of debt issued in period 0, \( w_0 \), and a variable that denotes the fraction of that debt that is long term, \( x_0 \), while its budget constraint in period 2, (9), is written in terms of the maturity value of total debt issued in periods 0 and 1 that is due in period 2, \( W_1 \). We call this characterization the maturity-value notation. The reason we use this notation is that our paper focuses on an MPE (see Maskin and Tirole (2001) for a formal definition of this concept). The key restriction of an MPE is that the strategies of each player depend on the minimum set of state variables.

What are the state variables of the policy game in period 2? This depends on how the constraints are written. To see this, consider the government budget constraint in the terminal period. It might seem natural to write the constraint in terms of the real value of the short-term debt issued in period 0, \( b_0 = \frac{B_0}{P_0} \), and the long-term debt issued in period 0, \( l_0 = \frac{l_0}{P_0} \). Let us call this the debt-issuance notation. Written in this way, the period-2 budget constraint is

\[
0 = (1 + i_1) \frac{b_1}{\Pi_2} + (1 + R_0) \frac{l_0}{\Pi_1 \Pi_2} - T_2. \tag{10}
\]

Comparing (10) with (9), we see that using the debt-issuance notation implies there are five state variables in period 2—that is, \( i_1, b_1, R_0, l_0, \) and \( \Pi_1 \)—while according to the maturity-value notation, there is only a single state variable, \( W_1 \).

Deriving (7)-(9) and using those constraints to define a policy game can be summarized in four steps. First, all government debt issued is written in terms of its maturity value—that is, the total number of (real) future dollars that the debt issued in a given period promises to pay, inclusive of future interest rates. This yields \( w_0 \) and \( w_1 \). Second, define the fraction of the debt in period 0 issued in terms of long-term debt, yielding \( x_0 \). Third, identify the state variables of the policy game in each period. Fourth, write the budget constraints today in terms of both the state variable(s) for the next period and the state variables today.

What are the state variables in the three periods? The other constraints, (3)-(6), are perfectly forward looking and hence have no endogenous state variables, and thus we can focus on the budget constraints. As we discussed above, according to the debt-issuance notation, in period 2, the unique payoff-relevant state variable is the real value of all debt payments

\[\text{Equivalently, these are what Maskin and Tirole (2001) call the “coarsest history of player actions.”}\]

\[\text{Consider, for example, the case in which (6) is written in terms of the price level instead of inflation; that is } \frac{P_t - P_{t-1}}{P_{t-1}} = \kappa(Y_t - \bar{Y}). \text{ In that case, } R_t \text{ is a state variable of a policy game at time } t. \text{ Since this equation can obviously be written in terms of inflation } \Pi_t, \text{ which is determined at time } t, \text{ the MPE refinement suggests that the price level is not an appropriate state variable of the game.} \]
due at that time summarized by $W_1$. This variable is the sum of the maturity value of debt issued in period 1 (that is, $w_1$) and the portion of the maturity value of debt issued at time 0 that was issued in the form of long-term bonds (that is, $x_0 w_0$). Moving to period 1, the state variables are the maturity value of debt issued in period 0 (that is, $w_0$) and the portion of this debt that is short term (that is, $(1 - x_0) w_0$). As we will see, these two terms play a central role in the analysis and correspond to the balance sheet incentive and the rollover incentive respectively. Finally, in period 0, the state variable is $w_{-1}$.

The remainder of the model is relatively fairly standard. (3) is a standard asset-pricing condition for one-period risk-free nominal debt, (4) is the asset-pricing condition for two-period risk-free nominal debt, while (5) represents the ZLB, which we impose directly. (6) is the Phillips curve, which applies in periods 0 and 1 because of price indexation. For simplicity, there is no indexation in the terminal period, and thus there is long-run money neutrality with output pinned down by $\bar{\bar{Y}}$.

2.1.4 Thought Experiment and Parameterization

The following thought experiment is at the heart of our paper. Imagine that in the short run, there is an unexpected shock, $\xi_0 < \bar{\xi}$, which generates a recession. In the medium and long runs, there is no shock, so that $\xi_1 = \xi_2 = \bar{\xi}$. In the baseline parameterization, each period is a year and steady-state output is normalized to 1. The shock $\xi_0$ is chosen so that the ZLB is binding and output drops by 7.5% in the absence of any debt-maturity changes at the ZLB, while $\kappa$ is chosen so that the drop in inflation is 2.5%, implying that $\kappa = 0.33$. The initial public debt is $b_{-1} = 1$—that is, 100% of annual output. The discount factor is $\beta = 0.95$. The two welfare weights are $\lambda_y = \lambda_T = 0.01$. The values of these weights are motivated by the parameterization in the quantitative model, where they are discussed in detail using the underlying microfoundations of the quantitative model, along with sensitivity analysis.

2.2 The Deflation Bias

This section reviews optimal monetary policy at the ZLB; it abstracts from fiscal policy, which is equivalent to $\lambda_T = 0$ in (2). Then taxes are lump sum and do not affect welfare. It follows that both the level of debt and its composition are irrelevant. While this policy problem is well known, it is a helpful starting point since QE in our model works as a commitment device for

---

12Relative to the debt-issuance notation, the maturity-value notation reduces the number of state variables from five to one in the terminal period, while the number of state variables remains the same in periods 0 and 1. In the language of Maskin and Tirole (2001), the maturity-value notation thus represents a “coarser partition of the history of the players’ actions” than the debt-issued notation and thus is the correct specification of the state in an MPE.

13The flexible price limit is $\kappa^{-1} = 0$, where no firms index their prices to the past level of aggregate prices.
Figure 1: Impulse responses under optimal monetary policy at the zero lower bound

Note: The figure shows the responses of output, inflation, the nominal interest rate, and the real interest rate when a negative demand shock makes the zero lower bound binding in the short run.

To understand the deflation bias, consider first the solution in the absence of the ZLB, but taking the shock at time 0 into account. According to objective (2), the first-best outcome is $Y_t = \bar{Y}$ and $\Pi_t = 1$, and it can always be achieved if the ZLB is not binding, as all that is required is that the nominal interest rate satisfy $i_0 = \beta^{-1} \xi_0 - 1$ while remaining at steady-state in periods 1 and 2. If the shock in period 0 is sufficiently large, the ZLB is binding, making the first-best solution infeasible. The yellow line in Figure 1 shows the equilibrium solution under optimal monetary policy at the ZLB when the government can commit in period 0 to a future policy. The government commits to keeping the nominal interest rate low in period 1, resulting in output being above $\bar{Y}$ and inflation overshooting in period 1.

The commitment solution is not time consistent. This is because in periods 1 and 2, the first-best equilibrium is feasible; that is, $Y_1 = Y_2 = \bar{Y}$ and $\Pi_1 = \Pi_2 = 1$. If the government

---

14 All commitment problems are solved using a Lagrangian method described in Appendix A.3.
is not constrained by its promise in period 0, it will optimally choose this solution, as it corresponds to the best it can do from that time onward. Taking this solution as given, the time-consistent solution is not the commitment solution that promises an output boom and inflation, but instead the one that can be backed out from (3) and (6) by setting $i_0 = 0$, $Y_1 = \bar{Y}$, and $\Pi_1 = 1$, yielding $Y_0 = \beta^{-1} \xi_0 < \bar{Y}$ and $\Pi_0 = \kappa(\beta^{-1} \xi_0 - \bar{Y}) + 1 < \bar{\Pi}$. This solution is shown in Figure [I] with a blue line. Clearly, the commitment solution is preferable when considering the payoffs for all periods since it ameliorates a severe recession and deflation. The problem is how to make the optimal commitment credible. This is where QE comes in.

2.3 Markov Perfect Equilibrium and Government Debt Maturity

The reason the maturity structure of government debt matters under discretion is that it changes the incentives faced by future governments through the state variables of the game. The major challenge in solving for optimal policy under discretion is modeling expectations. Expectations depend on how the private sector believes future governments will set policy. The period-0 government, for example, can change these expectations because its choice of the duration of its debt, $x_0$, and its debt’s total maturity value, $w_0$, become state variables in the game played by the government in period 1. The key advantage of assuming a finite horizon, as in this section, is that the policy game can be solved backward to yield solutions for expectations as a function of state variables. We start by solving the government problem in the long run (period 2) and then move backward in time.

2.3.1 The Long Run: The Inflation-Tax Trade-off

Period 2’s government solves

$$V^2(W_1) = \max_{T_2, \Pi_2} \left\{-\frac{1}{2} (\Pi_2 - 1)^2 - \frac{1}{2} \lambda_T (T_2)^2\right\}$$

subject to (9), where $V^2(W_1)$ is the value function of the government and recall that in the long run we assume that $Y_t = \bar{Y}$.

Combine the optimality conditions (see Appendix A.4) to obtain

$$\left(\Pi_2 - 1\right)_{MC \ of \ \Pi_2} = \lambda_T T_2 \left\{ \frac{W_1}{(\Pi_2)^2} \right\}_{MB \ of \ \Pi_2}.$$  

15As we discuss in Section 4.1, under monetary and fiscal commitment, while the level of debt matters, the maturity structure of debt is irrelevant.
The left-hand side of (12) is the marginal cost of inflation, while the right-hand side is the marginal benefit. Inflation is beneficial because it reduces the need for taxes. The marginal benefit of inflation has two components. The first term on the right-hand side, \( \lambda T \), is the shadow value of \( W_1 \) in period 2; that is, it reflects the marginal benefit of lower debt payments, expressed in period-2 taxes. The second term, the expression in curly brackets, measures the marginal reduction of debt payments resulting from higher inflation. Since debt is nominal, higher inflation reduces the real value of debt payments (and thus the need for taxes).

(12) combined with (9) yields

\[
(\Pi_2)^4 - (\Pi_1)^3 - \lambda T (W_1)^2 = 0. 
\] (13)

(13) implicitly defines the policy function for inflation,

\[
\Pi_2 = \Pi^2(W_1), 
\] (14)

and, via (9), the policy function for taxes,

\[
T_2 = T^2(W_1). 
\] (15)

Using the implicit function theorem, it can be shown that \( \frac{\partial \Pi^2}{\partial W_1} > 0 \), \( \frac{\partial T^2}{\partial W_1} > 0 \), and \( \frac{\partial V^2}{\partial W_1} < 0 \) for \( W_1 > 0 \). Two important inputs into the solution in period 1 are \( \frac{\partial V^2}{\partial W_1} \) and \( \frac{\partial \Pi^2}{\partial W_1} \), denoted by \( V^2_{W_1} \) and \( \Pi^2_{W_1} \) respectively. The envelope theorem implies that

\[
V^2_{W_1} = -\lambda T \frac{W_1}{(\Pi_2)^2}, 
\] (16)

while the implicit function theorem implies that

\[
\Pi^2_{W_1} = 2\lambda T \frac{W_1}{4(\Pi_2)^3 - 3(\Pi_2)^2}. 
\] (17)

It is straightforward to compute the nonlinear policy functions (14) and (15) numerically, which are shown in Figure 2 with a solid line for the baseline parameterization.

Moreover, analytical solutions can be derived by approximating the policy functions local to a point at which there are no tax distortions while taking \( W_1 \) as given. This approximation leads to the following proposition:

\footnote{The shadow value is equal to the Lagrange multiplier on (9). See Appendix A.4.}

\footnote{As the expansion is made around the point \( \lambda T = 0 \), the approximation is accurate to an order \( O(||\lambda T||^2) \), where \( W_1 \) is treated as a parameter.}
**Figure 2:** Long-run policy functions

Note: The figure shows policy functions for taxes and inflation in the long run for the model with optimal monetary and fiscal policy under discretion. The only state variable for the long run is the total maturity value of debt ($W_1$) inherited from the short and medium runs. The dashed lines show the approximate, analytical policy functions discussed in the text.

**Proposition 1** *Local to no tax distortions ($\lambda_T = 0$), given $W_1$, the MPE in period 2 is*

\[
\Pi_2 = 1 + (W_1)^2 \lambda_T \quad \text{and} \quad T_2 = W_1 - (W_1)^3 \lambda_T. \tag{18}
\]

*Proof.* First, at $\lambda_T = 0$, $\Pi_1 = 1$; and by (9), $T_2 = W_1$. (18) then follows directly from a first-order Taylor expansion of (13), and (19) follows from a first-order Taylor expansion of (15) (details are provided in Appendix A.6).

Proposition 1 suggests that inflation increases quadratically with $W_1$ while taxes increase linearly. Hence, in trading off higher taxes vs. higher inflation, as the long-run debt burden increases, the government optimally chooses to rely more and more on inflation. Figure 2 shows that the approximated policy functions shown in Proposition 1, shown by a dashed line, match the nonlinear policy functions almost exactly.

### 2.3.2 The Medium Run: The Rollover and Balance Sheet Incentives

Period 1’s government solves

\[
V^1(x_0, w_0) = \max_{n_1, r_1, y_1, i_1, w_1} \left\{ -\frac{1}{2}(\Pi_1 - 1)^2 - \frac{1}{2} \lambda_y (Y_1 - \bar{Y})^2 - \frac{1}{2} \lambda_T (T_1)^2 + \beta V^2(W_1) \right\} \tag{20}
\]

subject to (5), (6), (8), and

\[
1 + i_1 = \beta^{-1} \left( \frac{\bar{Y}}{Y_1} \right) \bar{\Pi}^2(W_1), \tag{21}
\]
Figure 3: Medium-run policy functions

Note: The figure shows policy functions for inflation and the nominal interest rate in the medium run with optimal monetary and fiscal policy under discretion. The vertical yellow line shows the optimal value of the maturity of debt ($x_0 = 1.31$). The maturity value of debt ($w_0$) is fixed at the optimal value (72%). The dashed lines show the approximate, analytical policy functions discussed in the text.

where $V^1(x_0, w_0)$ is the value function of the government, $\Pi^2(W_1)$ is given by (14), and $V^2(W_1)$ is given by (11). The government treats $\Pi^2(W_1)$ and $V^2(W_1)$ as given but understands how they change with the state variable $W_1$.

The optimality conditions are shown in Appendix A.5. They contain the derivatives $V^2_W$ and $\Pi^2_W$, which are given by (16) and (17). The fact that we have an explicit solution for these derivatives is the main advantage of the assumption of a finite horizon. The government’s optimality conditions and the constraints (5), (6), (8), and (21) define the equilibrium. These equilibrium conditions then implicitly define the policy function for each of the endogenous variables in period 1 as a function of the two state variables $x_0$ and $w_0$. That is,

$$Y_1 = Y^1(x_0, w_0), \Pi_1 = \Pi^1(x_0, w_0), i_1 = i^1(x_0, w_0), W_1 = W^1(x_0, w_0), \text{ and } T_1 = T^1(x_0, w_0).$$

(22)

The two most economically interesting policy functions are shown in Figure 3 (Appendix A.7.2 contains the full set of policy functions). A lower $x_0$—that is, a shorter maturity of debt in period 0—triggers a lower nominal interest rate and higher inflation in period 1. This suggests that the government in period 0 can credibly commit the government in period 1 to a lower interest rate and higher inflation by shortening the duration of debt in period 0. This is the key mechanism that gives QE its power. It thus deserves a careful description.

Below we derive a closed-form approximation of the policy functions for inflation and the nominal interest rate shown in Figure 3. The steps to get there clarify the key mechanisms of
the model, and we discuss them first. Consider the budget constraint in period 1:

\[ W_1 = w_0 \frac{\Pi_1}{1 + i_1} + i_1 (1 - x_0) w_0 \frac{\Pi_1}{1 + i_1} - (1 + i_1) T_1. \]  

(23)

Recall that \( W_1 \) is the state variable of the game in the next period and thus determines the continuation value \( \beta V_2(W_1) \) in (20). This variable is determined by three terms, all of which play a key role in determining equilibrium inflation: the balance sheet incentive, the rollover incentive, and the tax-smoothing incentive. We discuss each in detail below.

Combine the optimality conditions of the government’s problem in period 1 (see Appendix A.5) to yield a condition analogous to (12) in period 2:

\[ \lambda_T T_1 \left( \Pi_1 - 1 \right) + \frac{w_0}{\Pi_1^2} + \kappa^{-1} w_0 (1 + i_1) \left( 1 - x_0 \right) \frac{\Pi_1}{Y_1 \Pi_1^2} + \frac{i_1 (1 - x_0) w_0}{\Pi_1^2} - \kappa^{-1} (1 + i_1) T_1 \right\}. \]

\[ \frac{\partial W_1}{\partial \Pi_1} \]

(24)

The left-hand side of (24) is the marginal cost of inflation, and the right-hand side is the marginal benefit. Again, inflation is beneficial because it reduces the need for taxes. The first term on the right-hand side, \( \lambda_T T_1 \), reflects the marginal benefit of reducing future debt payments (that is, \( W_1 \)), expressed in terms of period-1 taxes. The second term, the expression in curly brackets, is equal to \( \frac{\partial W_1}{\partial \Pi_1} \) and reflects how much period-1 inflation reduces the debt that is rolled over to the terminal period.

While the consolidated optimality condition (24) can be used to approximate the policy function for inflation (as we will see below), the key to approximating the policy function for the interest rate is to combine (6) with (21) to yield

\[ 1 + i_1 = \beta^{-1} \left( \frac{\bar{Y}}{Y_1} \right) \bar{\Pi}^2(W_1) = \beta^{-1} \left( \frac{\bar{Y}}{\kappa^{-1} (\Pi_1 - 1) + Y} \right) \bar{\Pi}^2(W_1). \]

(25)

According to the first term in (25), the nominal interest rate in period 1 declines as inflation rises. The reason, which we term the sticky-price effect, is that higher inflation increases output in period 1 relative to period 2, putting downward pressure on the interest rate (given expectations about future inflation). The second term is the Fisher effect. It captures the fact that higher inflation expectations increase the nominal interest rate.

An approximation of (24) local to the point \( \lambda_T = 0, \Pi_1 = Y_1 = 1, 1 + i_1 = \beta^{-1} \), and
\( T_1 = \frac{w_0}{2} \), taking the state variables \( x_0 \) and \( w_0 \) as given (see Appendix A.6), yields

\[
MC_1 = (1 + \frac{\lambda y}{K^2})(\Pi_1 - 1) \quad \text{and} \quad MB_1 = \frac{\beta\lambda T w_0}{2} \left\{ \begin{array}{l}
\text{BSI} + \kappa^{-1} \beta^{-1}(1 - x_0)w_0 + \beta^{-1}(1 - \beta)(1 - x_0)w_0 - \kappa^{-1}\beta^{-1}w_0 \\
\text{ROI 1} \\
\text{ROI 2} \\
\text{TSI}
\end{array} \right\},
\]

where \( MC_1 \) is the marginal cost of \( \Pi_1 \) (the left-hand side of (24)) and \( MB_1 \) is the marginal benefit of \( \Pi_1 \) (the right-hand side of (24)). Similarly, an approximation of (25) then results in the following proposition:

**Proposition 2** *Local to no tax distortion* (\( \lambda_T = 0 \)),

(i) the policy functions for inflation and the interest rate in the MPE are given by

\[
\Pi_1 = 1 + \left( \frac{1}{1 + \frac{\lambda y}{K^2}} \right) \left\{ 1 - \frac{\kappa^{-1}\beta^{-1}}{2} + \beta^{-1}(1 - \beta + \kappa^{-1})(1 - x_0) \right\} \frac{\beta\lambda T w_0^2}{2} \quad \text{and}
\]

\[
i_1 = \beta^{-1}(1 - \beta) - \beta^{-1}\kappa^{-1}(\Pi_1 - 1),
\]

(ii) \( \frac{\partial \Pi_1}{\partial x_0} < 0, \frac{\partial i_1}{\partial x_0} > 0, \)

(iii) \( \frac{\partial \Pi_1}{\partial w_0} > 0, \frac{\partial i_1}{\partial w_0} < 0 \) iff \( x_0 < \frac{1+\kappa^{-1}}{1-\beta+\kappa^{-1}} \).

**Proof.** The first part of the proposition follows from equating (26) and (27) and from approximating (25) (further details are provided in Appendix A.6). The second and third parts of the proposition follow directly by differentiating (28) and (29).

Proposition 2 shows in closed form what we showed numerically in Figure 3: a lower \( x_0 \) (that is, a shorter maturity of debt in period 0) triggers higher inflation and a lower nominal interest rate in period 1. Figure 3 plots the approximated policy functions (28) and (29) with a dashed line in comparison with the exact policy functions shown by a solid line. Evidently, the approximation is relatively accurate, suggesting that the approximated marginal cost and benefit of inflation capture the key mechanism driving the shape of the policy functions.

Figure 4 plots the approximated marginal cost and benefit of inflation. The marginal cost of inflation increases linearly with inflation. The marginal benefit, however, is independent of inflation, as shown by the horizontal line. The intersection of the two curves gives equilibrium inflation, as given by (28), at point A. Consider comparative statics for \( x_0 \) and \( w_0 \). A reduction in \( x_0 \) increases the marginal benefit of inflation at any level of inflation—thus shifting the curve

\(18\) Local to no tax distortions, the time path for taxes is indeterminate. Assuming \( T_1 = \frac{w_0}{2} \), however, provides a relatively good approximation, as seen in Figure 3.
Note: The figure shows the approximate marginal cost and benefit of inflation in the medium run with optimal monetary and fiscal policy under discretion. The marginal cost of inflation is increasing in inflation. The marginal benefit does not depend on inflation, but instead depends on the maturity of debt ($x_0$). The intersection of the two lines gives equilibrium inflation.

Figure 4: Medium-run marginal cost and benefit of inflation

The forces at play are what we term the rollover incentives 1 and 2 (ROI 1 and ROI 2) in (27). ROI 1 arises because higher inflation reduces the real interest rate of the short-term debt rolled over to the next period. This reduces the government’s financing cost. ROI 2 arises because the interest rate cost of short-term debt is nominal. Hence, for a given $w_0$, increasing the fraction of short-term debt increases the benefit of inflation. ROI 1 and ROI 2 only apply to the fraction of the debt that is short term—that is, $(1 - x_0)w_0$. Intuitively, this is because the interest rate on the long-term debt has already been determined, while the interest rate on the short-term debt that is rolled over is determined in period 1. A reduction in $x_0$ therefore amplifies these incentives by increasing the share of debt being rolled over.

A reduction in $x_0$ always increases the marginal benefit of inflation. The same is not true for an increase in $w_0$. The condition $x_0 < \frac{1 + \frac{1}{\beta \kappa}}{1 - \beta + \kappa}$ must be met for a higher $w_0$ to be inflationary. To interpret this condition, it is again useful to consider the marginal benefit of inflation (described in (27)). Consider first the balance sheet incentive (BSI). This incentive refers to the fact that since debt is nominal, inflation in period 1 reduces the real value of debt. One can interpret the numerator of the fraction as the depreciation rate of the nominal debt. The condition for a higher $w_0$ to be inflationary is that this depreciation rate be lower than the rate of inflation in period 1. The condition is thus equivalent to the condition that the inflation rate in period 1 be lower than the real interest rate on the short-term debt in period 1.

An increase in $x_0$ would shift it down and hence decrease equilibrium inflation at point C.

Note that with flexible prices, as $\kappa^{-1} = 0$, ROI 1 goes away but ROI 2 does not. Thus, because of ROI 2, a lower $x_0$ triggers higher inflation in period 1, even under flexible prices. The period-1 policy functions under flexible prices are given in Figures A.5 and A.6 in Appendix A.7.4.

In our numerical example, the condition is $x_0 < 0.82$. Appendix A.7.2 contains Figure A.3, which shows the dependence of period-1 variables on $w_0$ and illustrates how in our baseline parameterization with high $x_0$, equilibrium inflation is indeed declining in $w_0$. 

19 An increase in $x_0$ would shift it down and hence decrease equilibrium inflation at point C.

20 Note that with flexible prices, as $\kappa^{-1} = 0$, ROI 1 goes away but ROI 2 does not. Thus, because of ROI 2, a lower $x_0$ triggers higher inflation in period 1, even under flexible prices. The period-1 policy functions under flexible prices are given in Figures A.5 and A.6 in Appendix A.7.4.

21 In our numerical example, the condition is $x_0 < 0.82$. Appendix A.7.2 contains Figure A.3, which shows the dependence of period-1 variables on $w_0$ and illustrates how in our baseline parameterization with high $x_0$, equilibrium inflation is indeed declining in $w_0$. 

16
the debt. Since both long- and short-term debt are nominal, however, this incentive does not depend on \( x_0 \). The marginal benefit of inflation via the BSI is always higher the higher is \( w_0 \). Conversely, because of the tax-smoothing incentive (TSI), the marginal benefit of inflation is always lower the higher is \( w_0 \). The ROI can either increase or decrease the marginal benefit of inflation associated with a higher \( w_0 \), depending on whether \( x_0 \) is greater than or smaller than 1. Which force prevails, and what is the mechanism?

To understand how the forces interact, it is useful to consider the special case \( x_0 = 1 \). In that case, the government only issues long-term debt in period 0 and there is no short-term debt to be rolled over in period 1. Yet the government still finds it optimal to levy taxes in period 1 (to smooth tax collection across periods). The tax collection in period 1 is then invested in short-term bonds that are rolled over at interest rate \( i_1 \). In this case, the government is a net debtor in long-term bonds and a saver in short-term bonds. Thus, in evaluating whether \( w_0 \) increases the marginal benefit of inflation, one needs to take into account that while inflation reduces the real value of the long-term debt (because of the BSI), it reduces the return of the assets (short-term nominal bonds) the government is holding on its balance sheet. The condition \( x_0 < \frac{\frac{1}{1-\beta+\kappa}+1}{\kappa-1} \) guarantees that the former force is always stronger so that the marginal benefit of inflation is increasing in \( w_0 \). Now consider the case in which \( x_0 \neq 1 \). When \( x_0 > 1 \), the ROI reinforces the TSI (because the government enters the period with short-term assets), while when \( x_0 < 1 \), the ROI reinforces the BSI.

To sum up, a lower \( x_0 \) always triggers expectations of higher inflation and a lower interest rate in period 1—and this is the central insight of our paper—while the effect of a higher \( w_0 \) depends on the overall balance sheet position of the government. The MPE for periods 1 and 2 can be used to formalize a model of QE in period 0, the subject of Sections 2.3.3 and 2.3.4.

### 2.3.3 The Short Run: Quantitative Easing as Comparative Statics in Policy

Consider a situation in which a shock in period 0 triggers the ZLB. Suppose that monetary and fiscal policies are set in periods 1 and 2 according to the MPE described in Sections 2.3.1 and 2.3.2. Consider now a policy regime in period 0 in which \( i_0 = 0 \) and there is some fixed (real) amount of government nominal debt; that is, \( D_0 = \frac{P_0}{P_0} + \frac{L_0}{P_0} = \bar{D} \). Now consider a policy in which in period 0 the government shortens the duration of its outstanding debt via open market operations. QE in period 0 can now be thought of as comparative static: what happens if the government purchases, say, $100 billion worth of long-term government bonds with short-term bonds (or reserves) via QE, holding total government debt, \( \bar{D} \), constant? The magnitude of QE directly maps onto the state variables \( x_0 \) and, to a smaller degree, \( w_0 \). We find this thought experiment to be interesting for it corresponds to the type of question policy maker ask themselves when deciding on the size of QE at a given point in time.
Figure 5: Impulse responses at the zero lower bound under various maturities of debt issued in the short run with otherwise optimal monetary and fiscal policy

Note: The figure shows the impulse response of output, inflation, the nominal interest rate, and taxes when a negative demand shock makes the zero lower bound binding in the short run. The model is solved for two different maturities of debt ($x_0$) issued in the short run under discretion (indicated by different colors), holding fixed the stock of debt issued ($\bar{D}$). The figure also shows the results for the model with optimal monetary and fiscal policy under commitment and the model in which the government deviates in periods 1 and 2 from the path of the nominal rate and inflation under $x_0 = 0$.

While it is clear enough that QE reduces the duration of the government debt—that is, it lowers $x_0$—the effect on $w_0$ is less obvious. Since $i_0 = 0$ and there is perfect foresight after the realization of the shock in period 0, simple manipulations yield $w_0 = \frac{(1+i_1)\bar{D}}{1+i_1(1-x_0)}$. This expression says that a lower $x_0$ is associated with a smaller $w_0$ for a given $i_1$. The reason is that $w_0$ measures the maturity value of debt at time 0, which includes future interest payments. Less long-term debt implies lower interest rate payments and thus lower maturity value in period 0, unless there is a simultaneous increase in $i_1$ to compensate. In numerical experiments, however, this effect is trivial, as we will see.

Figure 5 shows the solution of the model for an illustrative QE policy and compares it with some alternative policies. First, the yellow line shows the equilibrium under the assumptions that $x_0$ and $w_0$ are fixed at time 0 at the levels that would be optimal in the absence of
the ZLB shock—that is \((x_0, w_0) = (1.31, 0.723)\)—and that there is an MPE in periods 1 and 2. Next, the blue line shows the equilibrium for optimal fiscal and monetary policies under commitment. This commitment, however, is not credible in an MPE when the state variables are \((x_0, w_0) = (1.31, 0.723)\), as depicted by the yellow line.

What can the government do? The green line shows the effect of QE whereby the government shortens the maturity structure of debt so that \(x_0 = 0\) (thus issuing only one-period debt) while keeping \(D_0\) fixed. This implies a trivial adjustment of \(w_0\) from 0.723 to 0.721. Comparing the yellow with the green line, the solution illustrates that QE brings about a lower interest rate and a higher inflation rate in period 1; this force arises exclusively because of the shortening of the duration of debt. The expectation of this outcome then reduces the drops in output and inflation in period 0. QE, in other words, enables the government to credibly commit to a monetary expansion.

To understand why QE makes a commitment to low future interest rates credible, it is helpful to understand the cost of deviating from the loose monetary policy implied by this policy in period 1. The dashed-red line, labeled “Reneging” in Figure 5 shows the evolution of each variable if, in period 0, the government announces a plan to keep future interest rates low in line with the solution given by the green line (assuming the private sector believes this announcement), implements QE so that \((x_0, w_0) = (0, 0.721)\), but then deviates from the plan by choosing the equilibrium that had been optimal in the absence of QE (that is, the MPE associated with \((x_0, w_0) = (1.31, 0.723)\)). As in the analysis of the deflation bias, the government obtains better outcomes for output and inflation via this deviation strategy.

What is new is the effect the policy has on taxation: the government needs to raise taxes by more than it otherwise would have (compare the dashed-red line tax path to the green line’s path). The reason for the tax cost of reneging on the the optimal discretionary policy is twofold. First, the lower rate of inflation increases the real value of debt. This is the balance sheet incentive. Second, the government needs to pay a higher interest rate on the debt it is rolling over from period 1 to period 2. This is the rollover incentive.

The effect of QE can be characterized analytically if we assume that QE operates only via \(x_0\) (thus abstracting from its effect on \(w_0\), which is quantitatively trivial, as we have just seen), as given in the following proposition:

**Proposition 3** If (i) there is a policy regime with fixed \(w_0\) in period 0, (ii) there is an MPE in periods 1 and 2, and (iii) the policy function \(\Pi_1(w_0, x_0)\) is approximated using Proposition 2.

---

22 Section 4.1 further discusses the solution under commitment, where, in particular, the maturity structure of government debt in period 0 is irrelevant.

23 In computing the path for taxes when the government reneges, one can only pin down the sum of taxes in periods 1 and 2. We assume taxes are the same in both periods.
then
\[
\frac{\partial Y_0}{\partial x_0} = - \left( \frac{\lambda T}{1 + \frac{\lambda y}{\kappa}} \right) \beta^{-1}(1 - \beta + \kappa^{-1})(2\Pi_1 - 1) \frac{w_0^2}{2} \xi_0 < 0.
\]

Proof. From (3) we obtain \(Y_0 = \beta^{-1}Y_1\Pi_1\xi_0 = \beta^{-1}\kappa^{-1}(\Pi_1 - 1)\Pi_1\xi_0\). Using (28) to approximate \(\Pi_1\) and taking the partial derivative with \(x_0\) yields the result. ■

Proposition (3) says that when the ZLB is binding in period 0, a reduction in the maturity of government debt in period 0 leads to an increase in output for a given \(w_0\). This effect is higher the larger is the value of outstanding debt; it is increasing in the degree of tax distortions (\(\lambda_T\)); and it is stronger the higher is the degree of nominal rigidity. This statistic, \(\frac{\partial Y_0}{\partial x_0}\), can be used to answer an important policy question: what is the output effect at the ZLB of a certain amount of QE? The main objective of Section 3 is to obtain a quantitative answer to this question.

2.3.4 The Short Run: Quantitative Easing as an Optimal Time-Consistent Policy

The MPE in the short run provides an interpretation of QE as an optimal time-consistent policy for government debt and its maturity. Period 0’s government solves

\[
V^0(w_{-1}, \xi_0) = \max_{n_0, T_0, y_0, i_0, R_0, x_0, w_0} \left\{ -\frac{1}{2} (\Pi_0 - 1)^2 - \frac{1}{2} \lambda T T_0^2 - \frac{1}{2} \lambda y (Y_0 - \bar{Y})^2 + \beta V^1(x_0, w_0) \right\}
\]

subject to

\[
\frac{x_0 w_0}{1 + R_0} + \frac{(1 - x_0) w_0}{1 + i_0} = \frac{w_{-1}}{\Pi_0} - T_0,
\]

\[
\frac{1}{1 + i_0} = \beta \frac{Y_0}{Y^1(x_0, w_0)} \Pi^1(x_0, w_0) \frac{1}{\xi_0}, i_0 \geq 0,
\]

\[
\frac{1}{1 + R_0} = \beta^2 \frac{Y_0}{Y \Pi^1(x_0, w_0) \Pi^2(W^1(x_0, w_0))} \xi_0,
\]

\[
(\Pi_0 - 1) = \kappa (Y_0 - \bar{Y}),
\]

where \(V^0(w_{-1}, \xi_0)\) is the value function of the government and the state variables are \(w_{-1}\) and \(\xi_0\). Moreover, \(V^1(x_0, w_0)\) is given by (20), and the expectation functions—\(\Pi^1(\cdot), \bar{Y}^1(\cdot), \bar{W}^1(\cdot)\), and \(\Pi^2(\cdot)\)—are given by (22) and (14)24.

The green line in Figure 6 shows the solution for the MPE if policy is conducted optimally and the ZLB is binding in period 0. The yellow line shows the solution for optimal policy under commitment. The purple line shows the solution under discretion if the government

24We take optimality conditions for all variables, other than \(x_0\) and \(w_0\), and then evaluate the objective numerically on a grid of \(x_0\) and \(w_0\) in Mathematica, taking as given the solution derived for periods 1 and 2. We check for a global optimum.
Figure 6: Impulse responses under optimal quantitative easing and other policies at the zero lower bound

Note: The figure shows the responses of output, inflation, the nominal interest rate, and taxes when a negative demand shock makes the zero lower bound binding in the short run. The yellow line shows the solution for optimal policy under commitment. The green line shows the solution for optimal policy under discretion—that is, optimal quantitative easing. The purple line shows the solution for optimal policy under discretion with the maturity of debt in period 0 set to the value that would be optimal with no shock.

keeps the maturity of the debt fixed at time 0 at what would have been optimal in the absence of the shock; $x_0 = 1.31$. In this case, the government is allowed to optimize over $w_0$; we label this constrained discretion.
Figure 7: Government balance sheet dynamics under discretion at the zero lower bound

Note: The figure shows the evolution of the government balance sheet when a negative demand shock makes the zero lower bound binding in the short run. Panel (a) shows the solution for optimal policy under discretion—that is, optimal quantitative easing. Panel (b) shows the solution for optimal policy under discretion with the maturity of debt in period 0 set to the value that would be optimal with no shock—that is, constrained discretion.

Constrained discretion is qualitatively similar to the optimal monetary policy under discretion when \( \lambda_T = 0 \). In contrast, the fully optimal time-consistent solution, in which \( x_0 \) is chosen freely, almost completely replicates the optimal-commitment solution (yellow versus green line) in terms of output, inflation, and the interest rate.\(^{25}\) In the fully optimal solution, the government reduces maturity substantially to \( x_0 = -1.24 \) instead of keeping it at 1.31, which was the optimal duration in the absence of the ZLB.\(^{26}\)

Figure 7 shows the balance sheet position of the government—that is, the maturity value of debt \((w_{-1}, w_0, W_1)\) and the long- and short-term debt issued \((b_0, l_0, b_1)\). Panel (a) shows the balance sheet positions of the government when it structures both the level and composition of its debt optimally—that is, under optimal discretion. In period 0, the government issues only short-term debt and buys long-term debt from the private sector, thus closely replicating the commitment solution. Panel (b) shows the asset position of the government when the maturity composition is kept fixed in period 0 as if there were no shock, while the level of

\(^{25}\)The solutions will not coincide fully, however. Under commitment, there will never be inflation in period 2, while in the MPE, there will always be some inflation in period 2 because of inherited nominal debt. There is also a more pronounced difference in taxes, which further shows that the solutions are not exactly the same. As is natural, for taxes, they are higher in period 1 under optimal discretion compared with commitment. While our overall conclusions are independent of parameterization, in Figure A.4 in Appendix A.7.3 we present an example in which the differences are bigger for interest rates and inflation in all periods.

\(^{26}\)The negative value of \( x_0 \) provides an interpretation of some of the unconventional credit policies by the Federal Reserve, often referred to as QE1. A negative \( x_0 \) implies that the central bank is printing short-term reserves and buying long-term privately issued bonds. To the extent that the duration of the asset the Federal Reserve took on its balance sheet during the credit-easing operation is longer than that of reserves (which are fully short term), these policies change the inflation incentive of the government via the term structure.
debt is chosen optimally—that is, under constrained discretion. The government issues long-
term debt to purchase short-term debt, which is optimal only in the absence of the shock, as
it eliminates the inflation bias created by nominal government debt in period 1.

3 A Quantitative Model

This section considers a standard New Keynesian model that has been subjected to quanti-
tative evaluation and can be parameterized based on the existing literature (see, for example,
Woodford (2003) and Gali (2007) for textbook treatments). The standard model is extended
to include the costs of taxation and long-term government debt. Since the cost of taxation is
the key new parameter, it is the main focus of the calibration and sensitivity analyses.

3.1 Environment

A representative household maximizes

$$E_t \sum_{t=0}^{\infty} \beta^t [u(C_t) + g(G_t) - v(H_t)] \xi_t,$$

subject to a sequence of budget constraints

$$P_t C_t + B_t + S_t L_t \leq N_t H_t + (1 + i_{t-1}) B_{t-1} + (1 + \rho_{t-1} Q_t) L_{t-1} - P_t T_t + \int_0^1 Z_t(i) di,$$

where we use the same notation as in the simple model. Here, $u(.)$ and $g(.)$ are both increasing
and concave in their arguments while $v(.)$ is increasing and convex. $C_t$ is private consumption,
and $G_t$ is government consumption, where both are Dixit-Stiglitz aggregates of a continuum
of varieties of goods.

$B_t$ is a one-period riskless nominal government bond while $L_t$ is a perpetuity nominal
government bond with price $S_t$ and decay factor $\rho_t$. A perpetuity issued in period $t$ pays $\rho_t^j$
dollars $j + 1$ periods later for some decay factor $0 \leq \rho_t < \beta^{-1}$. At time $t$, the household can
sell its existing perpetuities, $L_{t-1}$, which have a decay factor $\rho_{t-1}$, at price $Q_t$. The model
thus incorporates debt of arbitrary duration. A value of $\rho_t = 0$, for example, implies that $L_t$
reduces to a one-period short-term bond, while $\rho_t = 1$ corresponds to a classic consol bond of
infinite duration. More generally, with stable prices, the duration of a perpetuity with decay
factor $\rho_t$ is $(1 - \beta \rho_t)^{-1}$.

The maximization problem and optimality conditions are standard, but with an additional
pricing condition for existing and new perpetuities. The standard Euler equation, which prices
one-period short-term bonds, is reported below in (32) while the asset-pricing equations for
current and existing long-term perpetuities are shown in (33).

A continuum of monopolistically competitive firms, indexed by \(i\), produce varieties of goods using a production function that is linear in labor, \(y_t(i) = H_t(i)\). Following Rotemberg (1982), we model firms as facing a cost of changing prices given by \(d \left( \frac{p(i)}{p_{t-1}(i)} \right)\), where \(d' > 0\) and \(d''(\cdot) < 0\). The firm maximizes expected discounted profits facing a subsidy \(\tau\) that is set so that steady-state output is efficient. The firms’ problem, after we impose a symmetric equilibrium, implies the nonlinear Phillips curve reported below in (34).

At the beginning of each period \(t\), the government buys all existing debt (perpetuities that have duration \(\rho_{t-1}\)) at price \(Q_t\) and issues new debt of duration \(\rho_t\) at price \(S_t\). The government issues no short-term bonds in equilibrium so that \(B_t = 0\). If we define \(l_t = \frac{L_t}{F_t}\) and \(\Pi_t\) as gross inflation, the government budget constraint is (35). Government policy—which is the choice of the tax \(T_t\), interest rate \(i_t\), debt \(l_t\), and maturity of government debt \(\rho_t\)—is also constrained by the ZLB, as reported in (32), and the resource constraint, written in (36).

The equilibrium conditions and constraints based on the problems described above are

\[
\frac{1}{1 + i_t} = \mathbb{E}_t \left[ \beta \frac{u_c(C_{t+1})\xi_t}{u_c(C_t)\xi_t\Pi_{t+1}} \right], \quad i_t \geq 0, \tag{32}
\]

\[
S_t = \mathbb{E}_t \left[ \beta \frac{u_c(C_{t+1})\xi_{t+1}}{u_c(C_t)\xi_t\Pi_{t+1}} (1 + \rho_t S_{t+1}) \right], \quad Q_t = \mathbb{E}_t \left[ \beta \frac{u_c(C_{t+1})\xi_{t+1}}{u_c(C_t)\xi_t\Pi_{t+1}} (1 + \rho_{t-1} Q_{t+1}) \right], \tag{33}
\]

\[
\theta Y_t \left[ u_c(C_t) - v_y(Y_t) \right] \xi_t + u_c(C_t)\xi_t d' (\Pi_t) \Pi_t = \mathbb{E}_t \left[ \beta u_c(C_{t+1})\xi_{t+1} d' (\Pi_{t+1}) \Pi_{t+1} \right], \tag{34}
\]

\[
S_t l_t = (1 + \rho_{t-1} Q_t) l_{t-1} \Pi_t^{-1} + (\bar{F} - T_t), \quad \text{and} \tag{35}
\]

\[
Y_t = C_t + F. \tag{36}
\]

\footnote{Unlike in the simple model, we here assume that the tax-collection costs are in terms of the final good rather than labor.}

\footnote{If government is also allowed to choose \(F_t\), this introduces one additional policy instrument. This does not, however, change the fact that taxes have a welfare effect via diverting resources away from welfare-contributing government consumption, \(G_t\), to tax collection, which contributes nothing to utility. Eggertsson (2006, 2008) considers the case in which both \(F_t\) and \(T_t\) are optimally chosen in a model with one-period debt. Rather than assuming \(F_t = \bar{F}\), the model could alternatively be written assuming that \(G_t = G\) and would yield identical results. In either case, the interpretation is that taxes have a negative effect on welfare, as they require resources to “produce” the tax revenues. In the former case, higher taxes result in lower government spending, \(G_t\), and lower welfare. In the latter, higher taxes result in lower private consumption, \(C_t\), and lower welfare. We thank an anonymous referee for stressing to us the importance of this alternative interpretation.}
3.2 Markov Perfect Equilibrium and Solution Method

Define the expectation variables \( f^E_t \equiv \mathbb{E}_t \left[ \frac{u_C(C_{t+1})\xi_{t+1}}{\Pi_{t+1}} \right] \), \( g^E_t \equiv \mathbb{E}_t \left[ \frac{u_C(C_{t+1})\xi_{t+1}}{\Pi_{t+1}} \right] (1 + \rho_{t}S_{t+1}) \), \( j^E_t \equiv \mathbb{E}_t \left[ \frac{u_C(C_{t+1})\xi_{t+1}}{\Pi_{t+1}} \right] (1 + \rho_{t-1}Q_{t+1}) \), and \( h^E_t \equiv \mathbb{E}_t [u_C(C_{t+1})\xi_{t+1}d' (\Pi_{t+1}) \Pi_{t+1}] \). An examination of conditions (32)-(36) reveals that the physical state variables of the model at time \( t \) are \( \rho_{t-1}, l_{t-1}, \) and the exogenous shock \( \xi_t \). The assumption of MPE implies that expectations are only a function of the state variables

\[
    f^E_t = \tilde{f}^E(l_t, \rho_t, \xi_t), \quad g^E_t = \tilde{g}^E(l_t, \rho_t, \xi_t), \quad j^E_t = \tilde{j}^E(l_t, \rho_t, \xi_t), \quad \text{and} \quad h^E_t = \tilde{h}^E(l_t, \rho_t, \xi_t),
\]

where \( \tilde{f}^E(\cdot), \tilde{g}^E(\cdot), \tilde{j}^E(\cdot), \) and \( \tilde{h}^E(\cdot) \) are unknown functions.

The government’s problem is

\[
    V(l_{t-1}, \rho_{t-1}, \xi_t) = \max_{i_t, T_t, l_t, \rho_t} \left[ U(\cdot) + \beta \mathbb{E}_t V(l_t, \rho_t, \xi_{t+1}) \right]
\]

subject to

\[
    1 + i_t = \frac{u_c(C_t)\xi_t}{\beta f^E_t}, \quad i_t \geq 0, \quad (39)
\]

\[
    S_t = \frac{1}{u_c(C_t)\xi_t} \beta g^E_t, \quad Q_t = \frac{1}{u_c(C_t)\xi_t} \beta j^E_t, \quad (40)
\]

\[
    h^E_t = \theta Y_t [u_c(C_t)\xi_t - v_y(Y_t)] \xi_t + u_c(C_t)\xi_t d'(\Pi_t) \Pi_t, \quad (41)
\]

\[
    S_l_t = (1 + \rho_{t-1}Q_t) l_{t-1} \Pi_t^{-1} + (\tilde{F} - T_t), \quad \text{and} \quad (42)
\]

\[
    Y_t = C_t + \tilde{F} + d'(\Pi_t), \quad (43)
\]

where \( f^E_t, g^E_t, h^E_t, \) and \( j^E_t \) are given by (37) and \( V(l_{t-1}, \rho_{t-1}, \xi_t) \) is the value function.

An equilibrium is defined as a collection of stochastic processes for the endogenous variables, where the processes solve both constraints (32)-(36) and the optimality conditions of the government’s problem (38). The optimality conditions, reported in Appendix C.4, contain derivatives of the expectation functions, such as \( \tilde{f}^E_t \) and \( \tilde{f}^E_{\rho} \). The expectation functions are unknown and hence so are their derivatives. The simple model was solved backward from a terminal period in which there were no expectations of the future. The solution of the terminal period could then be used to construct expectations for the second-to-last period, which in turn could be used for the period before. In an infinite-horizon model, this is not feasible.

The solution method follows the literature in making the assumption that the expectation functions are smooth and differentiable (see, for example, Soderlind (1999) and Klein et al. (2008)). This allows for an approximation of the expectation functions by using perturbation methods, which rely on the implicit function theorem, and the method of undetermined coef-
The solution method is discussed in Appendices C.7 and C.8. A key simplification is that the model is assumed to be efficient in steady state for positive debt at any given maturity (the value of debt can be taken from the data). This allows for a full characterization of the steady state in the MPE in closed form, which is used as an approximation point.

3.3 Thought Experiment

The thought experiment here is the infinite-horizon analog to the thought experiment in Section 2.3.3. The shock follows a two-state Markov process with an absorbing state, as in Eggertsson and Woodford (2003). In period 0 there is an unexpected shock so that $\xi_0 < \bar{\xi}$. In the following periods, there is a probability $1 - \mu$ that the shock will revert back to steady state so that $\xi_1 = \bar{\xi}$ and a probability $\mu$ that the shock will remain in its low state so that $\xi_1 = \xi_0 = \xi_S < \bar{\xi}$. If the shock reverts to steady state, it stays there forever. If the shock remains, then it behaves exactly the same in the following period; that is, it has a probability $1 - \mu$ of reverting to steady state and probability $\mu$ of remaining so that $\xi_2 = \xi_1 = \xi_0 < \bar{\xi}$. The shock takes the same structure in all future periods. This means that the expected duration of the shock is $\frac{1}{\mu}$ quarters. The stochastic period in which the shock returns to steady state is denoted by $K$. The period prior to that, $t < K$, is the short run (denoted by $S$), while $t \geq K$ is the long run. The shock is chosen to be large enough so that the ZLB is always binding in the short run. In an approximate equilibrium, the shock is defined as $r^e_s = \bar{r} + \sigma^{-1}(\hat{\xi}_t - E_t\hat{\xi}_{t+1})$, interpreted as the real interest rate needed for output to reach its flexible price level.

In the short run—that is, $t < K$—policy is (possibly) suboptimal as in Section 2.3.3. This allows for a straightforward match to particular QE episodes observed in the US in the Great Recession. In the long run, in which $t \geq K$, policy is set according to the MPE. QE shortens the duration of government debt in the short run—that is, it leads to a smaller $\rho_t$—holding fixed the total value of the debt issued ($S_t \ell_t = \bar{D}$). This means outcomes can be compared with empirical estimates from asset markets during QE episodes as a test of the model.

Consider a variable $x$, and denote by $\Delta x$ the change in this variable in response to QE. A QE episode is mapped onto a reduction of $\Delta \rho$ and a resulting change in variables of interest, $\Delta x$. We can compare the model prediction to the data, which we do in Section 3.5.3.

---

This allows us to characterize the steady state in closed form in the MPE. This is a major simplification relative to Klein et al. (2008), who rely on nontrivial approximations to characterize the steady state. Leeper et al. (2019) similarly consider a time-consistent equilibrium in the New Keynesian model with long-term debt in which the steady state is not efficient, approximating it using collocation methods.
3.4 Parameterization

The model is fully parameterized, with choices of parameters (discussed in detail below), related to (i) fiscal policy and QE ($\rho, \Delta \rho, \psi, \hat{b}_S$); (ii) structural parameters of the model from the private sector ($\beta, \kappa, \sigma$); (iii) the size and persistence of the shock ($r^e_S, \mu$); and (iv) the welfare-based weights the government puts on variability in output and taxation ($\lambda_y, \lambda_T$). We discuss calibration of each in turn, with special attention to $\lambda_T$, which is new.

3.4.1 Calibration of Baseline Parameters

The fiscal policy/QE parameters are directly backed out of the data. The model assumes a consolidated government budget constraint so that, for government debt, the relevant measure is all government debt held by the public (thus, government debt held by the Federal Reserve is netted out). Reserves issued by the Federal Reserve are treated as short-term government debt. The duration of the consolidated government’s debt is shown in Figure C.8 in Appendix C.9.1, where different phases of QE, termed QE1, QE2, and QE3, are discussed. The baseline QE experiment is QE2, which started on November 2010. The baseline maturity is 16.87 quarters—the level at the beginning of QE2—resulting in $\rho = 0.9502$. The reductions in maturity due to QE2 and QE3 are shown in Table 1. See Appendix C.9.1 for details on the construction of these statistics.

In addition to debt duration and its change due to QE2 (that is, $\rho(\text{pre-QE2})$ and $\Delta \rho(\text{QE2})$), two other parameters are directly calibrated in Table 1: $\psi$ and $\hat{b}_S$. The parameter $\psi \equiv \frac{L^S}{T} = \frac{D}{T}$ is the steady-state level of the debt-to-taxes ratio. The long-run average of the ratio of the market value of debt to output is computed using data from the Federal Reserve Bank of Dallas (1942-2010), while NIPA data from Bureau of Economic Analysis (1947-2010) is used to measure the ratio of taxes to output, yielding $\psi = 7.2$. Debt in the short run, $\hat{b}_S$, is 30% above its steady-state value based on data from the Federal Reserve Bank of Dallas (1942-2010).
Table 1: Calibrated/estimated parameter values in the quantitative model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.99</td>
<td>Discount factor</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.5</td>
<td>Scaled IES</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.02</td>
<td>Slope of Phillips curve</td>
</tr>
<tr>
<td>$\lambda_y$</td>
<td>0.029</td>
<td>Weight on output</td>
</tr>
<tr>
<td>$\lambda_T$</td>
<td>0.021</td>
<td>Weight on taxes</td>
</tr>
<tr>
<td>$r^e_S$</td>
<td>-0.01</td>
<td>Shock size</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.88</td>
<td>Shock persistence</td>
</tr>
<tr>
<td>$\rho(\text{pre-QE}2)$</td>
<td>0.9502</td>
<td>Baseline debt maturity</td>
</tr>
<tr>
<td>$\Delta \rho(\text{QE}2)$</td>
<td>0.0024</td>
<td>Change in debt maturity</td>
</tr>
<tr>
<td>$\Delta \rho(\text{QE}3)$</td>
<td>0.0072</td>
<td>Change in debt maturity</td>
</tr>
<tr>
<td>$\psi$</td>
<td>7.2</td>
<td>Debt-to-taxes ratio</td>
</tr>
<tr>
<td>$\hat{b}_S$</td>
<td>0.30</td>
<td>Initial debt</td>
</tr>
</tbody>
</table>

The parameters $\beta$, $\kappa$, and $\sigma$ are taken from Eggertsson and Woodford (2003), as shown in Table 1. The model implies that $\lambda_y = \kappa/\theta$, where $\theta$ is the elasticity of substitution across varieties of goods. $\theta$ is chosen to be 11, such that it is consistent with a 10% markup. This target for markup is within the range of estimated values in De Loecker and Warzynski (2012). The size of the shock $r^e_i$ and its persistence are chosen to match the fall in real GDP and inflation at the start of the 2008 crisis in the US using the values for these targets from Del Negro et al. (2017) (-7.5% and -2.5% respectively). The resulting values are in Table 1.

The parameter that remains to be chosen is $\lambda_T$. This parameter is central to the results, and there is little guidance in the literature on how to choose it. Accordingly, it is the main focus of the calibration and the sensitivity analysis.

3.4.2 Calibration of Tax Distortions ($\lambda_T$)

The fundamental role $\lambda_T$ plays in the model is to determine the relative response of inflation and taxes to fiscal shocks that increase debt. To illustrate how we use this model property, together with an estimated empirical counterpart, to calibrate $\lambda_T$, we proceed in four steps.

First, Figure 8 shows the model’s impulse responses of inflation and taxes to a fiscal shock for the benchmark parameterization of $\lambda_T$. In response to a fiscal shock that increases debt, both inflation and taxes increase and then gradually return back to steady state. The shaded areas below the impulse responses represent the total response of inflation and taxes up to

---

34 For instance, Table 2 in De Loecker and Warzynski (2012) shows that depending on the specification, estimated markups vary from 3% to 28%.

35 Our interpretation of a fiscal shock here is a one-time increase in fiscal outlays, which results in an increase in debt if we keep debt maturity fixed.
Figure 8: Impulse responses outside of the zero lower bound

Note: The figure shows the impulse responses of inflation and taxes outside the zero lower bound to a shock that increases debt. This is for the benchmark parameterization in Table 1.

sixteen quarters. This is the standard horizon over which impulse responses are cumulated in the empirical literature on fiscal policy (see Ramey and Zubairy (2018a), or RZ hereafter).

The calibration strategy of $\lambda_T$ relies on the model property that the higher is $\lambda_T$, the more the government responds to a fiscal shock by increasing inflation relative to taxes in the impulse responses shown above.  \[36\] This is shown in Figure 9, which plots the ratio

$$D_{\pi,T}(\lambda_T) = \frac{IRF_{\pi}(16)}{IRF_T(16)}$$

as a function of $\lambda_T$ while using the benchmark values for all other parameters. Here, $IRF_{\pi}(16)$ and $IRF_T(16)$ represent the shaded areas in Figure 8. Figure 9 shows that the higher is $\lambda_T$, the more the government relies on inflation relative to taxes to pay down debt. This theoretical moment is then a natural target with which to identify a value for $\lambda_T$. Point B corresponds to the benchmark calibration reported in Table 1 for $\lambda_T$, which is obtained by matching (44) to the empirical counterpart of this statistic, which we now turn to.

The empirical counterpart to (44) is generated by replicating and extending the empirical analysis of RZ. RZ collect data on inflation, debt, taxes, and government spending for the period 1889–2018 in the US. In contrast, we truncate our data to end in 2008 to focus on evidence outside of the 2008 crisis.  \[37\] RZ focus on the effect of fiscal shocks on output using local projection methods.  \[38\] We adapt their analysis, using data from Ramey and Zubairy

---

\[36\] This can be seen analytically in period 2 of the simple model, as (12) rewritten gives $(\Pi_2 - 1) \Pi_2 = \lambda_T T_2^2$.

\[37\] We truncated the data based on feedback from referees so that it strictly covers the pre-ZLB period. The results do not change much if we instead use the whole period.

\[38\] In their analysis, they use Ramey’s (2011) military-news variable and Blanchard and Perotti’s (2002) shock series as instruments for government spending in local projection regressions; see their Equation (3).
Figure 9: Identification and estimation of tax-smoothing parameter

Note: The figure shows how \( \lambda_T \) is identified and estimated using the response of inflation relative to taxes in the model and matching it to the empirical counterpart. Point B corresponds to the point estimate of \( \lambda_T \).

(2018b), to analyze cumulative impulse responses for taxes and inflation in response to a fiscal shock—that is, the empirical objects \( IRF_\pi(16) \) and \( IRF_T(16) \). We then take the ratio of these impulse responses to construct the empirical target \( \hat{D}_{\pi,T} \), the point estimate of which is 0.162 and the standard error (\( \hat{\sigma} \)) of which is 0.085. Appendices B.1-B.3 contains more details on the estimation, including how we construct the standard error for the ratio of impulse responses using the delta method.

Then we choose \( \lambda_T \) so that (44) in the model matches this empirical target; that is, we back out the \( \lambda_T \) that solves \( D_{\pi,T}(\lambda_T) = \hat{D}_{\pi,T} \). As shown in Figure 9, there is a solution for \( \lambda_T \) that solves this equation. Point B is thus the point estimate for \( \lambda_T \), as reported in Table 1. Points A and C correspond to values for \( \lambda_T \) estimated by matching +/- one standard deviation from the point estimate of \( \hat{D}_{\pi,T} \), which we use in our main sensitivity analysis later.

3.5 Main Results

3.5.1 Dynamic Paths

The blue line in Figure 10 shows the evolution of the main variables of interest in the benchmark scenario in which government debt and its duration are fixed throughout the ZLB and an MPE is in place once the shock has subsided. In response to the shock, output and inflation fall by 7.5% and 2.5% respectively, as targeted in the calibration. The figure shows the realization of a shock that reverts back to steady state in period 8, the expected duration of the shock. Once the shock reverts to steady state, the government raises the interest rate to stabilize inflation and output at a level slightly above steady state because government debt is 30% above steady state.
Figure 10: Impulse responses under various maturities of debt at the zero lower bound

Note: The figure shows the impulse responses of output, inflation, the nominal rate, and the real rate when a negative demand shock makes the zero lower bound bind initially and the shock reverts back to steady state in eight periods. The model is solved for three different debt durations at the zero lower bound, indicated by different colors, holding fixed the value of debt issued. This illustrates quantitative easing as a comparative-static experiment. The figure also shows, for comparison, the model with optimal monetary and fiscal policy under commitment.

The red line in Figure 10 shows the optimal monetary and fiscal policy under commitment. In response to the shock, the government commits to a lower future interest rate. While the discretionary central bank raises the interest rate in period 8, the central bank under the optimal-commitment policy raises rates six periods later. As shown in the lower-right panel, this brings about a lower real interest rate during the period of the negative shock, which mitigates the recession greatly, with output only dropping by about 2%. Meanwhile, inflation stays within 50 bp of zero throughout the crisis.

While the optimal-commitment solution represents the best equilibrium, a discretionary

---

39It can be shown, using a similar argument as that in the simple model, that under monetary and fiscal commitment, the maturity structure of debt is irrelevant even in this quantitative model. The commitment solution was computed using the approximation-based toolkit of Eggertsson et al. (2019), which is based on Eggertsson and Woodford (2003). The properties of the approximation at the ZLB in this class of models are discussed in Eggertsson and Singh (2019).
government, as portrayed by the blue line, is unable to credibly commit to keeping the interest rate low once the shock subsides because it has an incentive to renege on its previous promises and raise interest rates. This is where QE comes in, as in the simple model.

The yellow line shows the solution if the government reduces the duration of debt as in QE2. This results in a commitment to lower the nominal interest rate in the future relative to the blue line, thus more closely corresponding to the optimal commitment. This is made clearer via the real interest rate (shown in the fourth panel), which is what determines output. By shortening the maturity of outstanding debt, the government is able to commit to lower real interest rates once the shock is over, which stimulates output during the period in which the ZLB is binding. Hence QE is an effective commitment device and renders low future interest rates credible. The green line in the figure shows the effect of QE3, which is larger.

To understand why a shorter maturity brings about a commitment to lower future interest rates, and to compare with the simple model, it is helpful to make a few additional assumptions, even if this glosses over some features of the solution discussed in the next paragraph. The endogenous state variables are \( l_{t-1} \) and \( \rho_{t-1} \). For simplicity, consider the case \( \rho_t = \rho_{t-1} \), so that \( S_t = Q_t \). Furthermore, contemplate a small deviation from steady state so that in period \( t + 1 \) the interest rate is back to steady state, in which case \( 1 + i_{t+1} = \beta^{-1} \) and \( S_t = \frac{1}{1 - \beta \rho_{t-1}} \). Making the simplification that \( \beta = 1 \), the budget constraint (35) can then be written as

\[
l_t = \frac{l_{t-1}}{\Pi_t} + i_t (1 - \rho_{t-1}) \frac{l_{t-1}}{\Pi_t} + \frac{1}{S_t} [F - T_t].
\]  

(45)

That is analogous to (23) in the simple model, where the balance sheet and rollover incentives are defined. The first term on the right-hand side of (45) corresponds to the balance sheet effect, which is independent of the duration of government debt, while the second term corresponds to the rollover incentive, but \( x_0 \) has been replaced by \( \rho_{t-1} \). The last piece corresponds to the tax-smoothing incentive as before. The logic is thus the same in the two models.

Finally, Figure 11 shows how the transition dynamics of debt and duration are affected by QE once the shock is over. As the figure reveals, once out of the ZLB, the government reduces the debt back to its steady state. Note that as the debt decreases, the government lengthens the duration of its debt, an issue discussed towards the end of Section 4.4.

3.5.2 Output Effects of Quantitative Easing

As a summary of the macroeconomic effects of QE shown in Figure 10, Table 2 shows the output effects of QE in the benchmark calibration. Output at the ZLB increases by about 1.65 percentage points for QE2 and 4.85 percentage points for QE3.
Figure 11: Transition dynamics of the government balance sheet

Note: The figure shows the dynamics of debt and debt duration once the shock is over.

Table 2: Output effects of QE2 and QE3 under baseline calibration

<table>
<thead>
<tr>
<th>Model Specification</th>
<th>ΔOutput, QE2</th>
<th>ΔOutput, QE3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Calibration</td>
<td>1.65%</td>
<td>4.85%</td>
</tr>
</tbody>
</table>

3.5.3 Untargeted Moments

We can now ask: how well does the model account for empirically estimated financial-market behavior in response to QE? To assess this, we compare the model prediction of the changes in the long-term interest rate and in expected inflation to well-known empirical estimates. This is an out-of-sample test, as our calibration did not target any estimates of financial-market effects of QE taken from the literature.

The main comparison is to the estimated effects of QE2 in Krishnamurthy and Vissing-Jorgensen (2011, or KVJ hereafter). Using fed funds futures, they estimate a reduction of $\Delta i(8) = -16$ bp in the short-term interest rate’s eight-quarters-ahead yield and an increase of $\Delta \pi (40) = 5$ bp in expected inflation over ten years. This empirical estimate is particularly attractive for our analysis since KVJ argue that the changes in the fed funds futures are due to signaling by the Fed about future interest rate policy. This is the mechanism the model is designed to capture. These out-of-sample predictions are summarized in Table 3. Since none of the parameters were chosen with this in mind, the match is less than perfect.

A few points are worth highlighting. The way the shock is chosen (that is, its persistence) implies an expected ZLB episode of 8.4 quarters, which is in the range of surveyed expectations of the duration of the ZLB in this period (see, for example, Del Negro et al. (2017)). As already stressed, we pick the shock parameters to generate a recession and drop in inflation and not
Table 3: Untargeted data and model moments under baseline calibration

<table>
<thead>
<tr>
<th>Untargeted Moments</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero-lower-bound duration</td>
<td>6 – 8 qt</td>
<td>8.4 qt</td>
</tr>
<tr>
<td>$\Delta i(8), QE2$</td>
<td>−16 bp</td>
<td>−10.24 bp</td>
</tr>
<tr>
<td>$\Delta i(40), QE2$</td>
<td>−30 bp</td>
<td>−11.2 bp</td>
</tr>
<tr>
<td>$\Delta i(40), QE3$</td>
<td>−46 bp</td>
<td>−34.9 bp</td>
</tr>
<tr>
<td>$\Delta \pi(40), QE2$</td>
<td>5 bp</td>
<td>14.34 bp</td>
</tr>
</tbody>
</table>

to match the expected duration of ZLB.

The model accounts for about two-thirds of KVJ’s estimated fall in two-year expected future short rates in response to QE2. In addition, the model accounts for about a third of the fall in ten-year Treasury yields as a result of QE2 and about two-thirds of the fall in ten-year Treasury yields that Ehlers (2012) estimates resulted from QE3. Finally, the model overshoots KVJ’s estimated ten-year-ahead increase in inflation expectations due to QE2.

To summarize, the model accounts fully for the rise in inflation expectations due to QE2 and accounts for roughly a third to two-thirds of the fall in long-term yields due to QE2 and QE3. Taken at face value, this suggests that alternative theories, such as those based on portfolio rebalancing or financial frictions, may account for the rest of the effects on yields.

3.6 Sensitivity Analyses

3.6.1 Bounds on Output Effects

We consider sensitivity of the results to reasonable variation in $\lambda_T$ by reporting the results of assuming that $\lambda_T$ is +/- one standard deviation from the estimated value used in the baseline calibration. That is, it corresponds to points A (labeled $\lambda_T^{high}$ below) and C (labeled $\lambda_T^{low}$ below) in Figure 9. The output effects implied from this analysis are shown in Table 4.

Table 4: Output effects of QE2 and QE3 with different degrees of tax smoothing

<table>
<thead>
<tr>
<th>Model Specification</th>
<th>$\Delta$Output, QE2</th>
<th>$\Delta$Output, QE3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline $\lambda_T$</td>
<td>1.65%</td>
<td>4.85%</td>
</tr>
<tr>
<td>$\lambda_T^{high}$</td>
<td>2.16%</td>
<td>5.73%</td>
</tr>
<tr>
<td>$\lambda_T^{low}$</td>
<td>0.26%</td>
<td>0.89%</td>
</tr>
</tbody>
</table>

Table 4 shows that for the lower value of $\lambda_T$, the output effects of QE2 and QE3 are smaller than in the benchmark, at 0.26 and 0.89 percentage points respectively. Meanwhile, for the

---

40The ten-year-yield effects of QE2 estimated by Ehlers (2012) are not pure signaling effects; they include a term-premia effect. Thus, it was to be expected that our model would not match that moment closely.
higher value of $\lambda_T$, the output effects of QE2 and QE3 are higher, at 2.16 and 5.73 percentage points respectively. The intuition for why QE has a smaller effect for a lower value of $\lambda_T$ is that for low values of this parameter, as we saw in Figure 9, the government relies relatively less on inflation and more on tax increases in order to pay down public debt.

This sensitivity analysis chooses $r^e_S$ to generate an output drop of 7.5%, but leaves inflation free. The value of the persistence of the shock $\mu$ is unchanged as $\lambda_T$ is changed, but recall that $\mu$ maps onto the expected duration of the ZLB. This procedure then implies that as $\lambda_T$ changes, the model generates different inflation drops at the ZLB than that in the data. Appendix C.9.2 documents this and also shows the responses of other key empirical moments.

### 3.6.2 Bounds on Output Effects Matching Inflation at the Zero Lower Bound

An alternative way of doing the sensitivity analysis with respect to $\lambda_T$ is to recalibrate $\mu$ to hit the 2.5% drop in inflation observed in the data. Doing so yields very high estimates for the output effects of QE2 and QE3. In our view, however, these high estimates are hard to defend. High values of $\lambda_T$ imply, using this approach, that a very persistent expected slump is needed to generate a 2.5% drop in inflation (high $\mu$). The expected duration of the ZLB according to this procedure is more than double the expected duration in the benchmark calibration. The benchmark calibration implies an expected duration of the ZLB that is very close to market expectations following 2008, while the alternative calibration does not, casting doubt on the plausibility of the large estimated effects of QE using this procedure.

A more informative alternative is to keep $\mu$ fixed in line with market expectations and vary $\lambda_T$ around the point estimate, but recalibrate $\kappa$ so as to match the drop in inflation at the ZLB. Appendix C.9.3 reports the results of this procedure. The bottom-line is that this alternative procedure results in higher output effects of QE for both high and low $\lambda_T$, but they are less extreme than the previous procedure.

### 3.6.3 Bayesian Approach to Estimating $\lambda_T$

In Section 3.4.2, $\lambda_T$ is estimated by equating the theoretical object $D_{\pi,T}(\lambda_T)$ to $\hat{D}_{\pi,T}$, and then computing the 68% confidence interval for $\lambda_T$ by matching $\hat{D}_{\pi,T} +/\- \hat{\sigma}$, the estimated standard error of $\hat{D}_{\pi,T}$. The object $D_{\pi,T}(\lambda_T)$, however, is highly nonlinear and a complicated function of $\lambda_T$. It is not obvious a priori that if we characterize the entire distribution of $\hat{D}_{\pi,T}$ and back out the implied $\lambda_T$ for each point in the distribution, that the resulting confidence interval corresponds closely to the confidence interval we used for $\lambda_T$ in the previous sections.

We show below, however, that this alternative approach leads to almost identical inference.

---

41 We thank the editor for suggesting this alternative estimation procedure.
Note: The figure shows the posterior distribution of $\lambda_T$ using a Bayesian estimation approach.

Drawing from the distribution of $D_{\pi,T}$ and computing the implied distribution of $\lambda_T$ has a Bayesian interpretation described in detail in Appendix B.4. Figure 12 shows the estimation results in terms of the posterior density of $\lambda_T$, where we depict the mean and the 68% credible sets with vertical lines. The mean is 0.022 while the 68% probability interval is (0.007, 0.032). These estimates are almost identical to those from our baseline approach reported in Figure 9, where the mean is 0.021 and the 68% confidence interval is (0.00804, 0.03684). These new estimates’ implied confidence interval for output effects of QE2 is also essentially the same as that in Table 4, see Appendix B.4.

3.6.4 Alternative Calibration Strategies

Recalibrating to match financial-market effects. So far the strategy in the numerical experiment has been to pick the parameters based on existing literature or the data and then to pay special attention to the estimation of $\lambda_T$. We then asked, given the parameterization, how well the model could account for financial-market estimates of QE, which we did not directly target. An alternative way of parameterizing the model is to estimate directly more parameters and ask, can the model fully account for the financial-market effects of QE for parameter values that look reasonable? We address this question by calibrating $(\lambda_y, \kappa)$ to

\[ A \text{ small fraction of draws, approximately 3.1%, imply negative values of } D_{\pi,T}. \text{ For those draws, we impose the assumption that the estimated } \lambda_T \text{ is zero. In the posterior density plot of } \lambda_T, \text{ we do not depict the zeros because it is continuous. In computing the mean and probability intervals, however, we do count the zeros.} \]
match KVJ’s estimates that QE2 led to \( \Delta i^* (8) = -16 \) bp and \( \Delta \pi^* (40) = 5 \) bp. In other words, we estimate \((\lambda_y, \kappa, \lambda_T, r_S, \mu)\) to match five moments exactly: the drop in inflation and output at the ZLB, the ratio of the inflation response to the tax response to match RZ’s evidence, and the effects of QE2 on yields and expected inflation to match KVJ’s evidence.

Appendix C.9.4 contains both the parameter values from this exercise as well as the model’s predictions for the output effects of QE2 and QE3. Observe, that while we choose the parameters to match the change in financial-market prices as result of QE2 and QE3, the parameterized model is needed to draw any inference about the effect these policies had on output. The bottom-line is that parameter values are similar to our baseline calibration reported earlier in Table 1 and can thus be interpreted to be relatively reasonable, since the parameters were motivated by the outside literature. The output effects of QE are also very similar to those in the baseline experiment where we did not match the empirically estimated effects of QE2 on yields and expected inflation (the effect of QE2 is 1.63% while that of QE3 is 5.12%) \(^{44}\).

### Sensitivity to structural model: A discounted Euler equation

A recent criticism of the standard New Keynesian model is that current variables are overly sensitive to expectations of future interest rate policy. This is dubbed the “forward guidance puzzle” by Del Negro et al. (2015). McKay et al. (2016), Gabaix (2016), and Michaillat and Saez (2019) have suggested a resolution to this puzzle: the consumption Euler equation has additional discounting, which we denote by \( \alpha \), with \( \alpha = 1 \) corresponding to the standard model. McKay et al. (2016) suggest that the forward guidance puzzle is resolved with a value of \( \alpha = 0.97 \). Using this as an input in the model equation directly, we re-estimate the current model using the same procedure as in last subsection. We find that the output effect of QE2 is now 1.33% while that of QE3 is 4.18%. Thus, the output effects of QE are only modestly reduced by this extension.\(^{45}\)

---

\(^{43}\)Note that \( \lambda_y = \kappa/\theta \), and thus we are estimating \( \theta \) when estimating \( \lambda_y \).

\(^{44}\)The reason for similar output effects is that in the baseline exercise, the model generated smaller effects on long-term yields but bigger effects on expected inflation compared with KVJ’s estimates. These two forces essentially have equal but opposite effects for output in the model.

\(^{45}\)It may seem surprising that the impact of QE did not decrease more. The reason is that this alternative specification leads to different values for all of the five parameters that were chosen to match empirical moments. Importantly, the experiment is thus not to change \( \alpha \) from 1 to 0.97, holding all other parameters fixed, for in this case the model would no longer be matching the empirical moments targeted, such as the drop in output and inflation seen in 2008.
4 Extensions and Discussion

4.1 Irrelevance of Quantitative Easing under Policy Commitment

We have shown that QE has a natural interpretation when the government cannot fully commit to future policy, i.e., in an MPE. The first fundamental assumption driving this result is that taxation carries some cost, as we made explicit in Section 2.2 by considering the case in which $\lambda_T = 0$. A second fundamental assumption is that the government is unable to commit to future policy apart from being able to commit to pay back the nominal value of outstanding debt. The role of the second assumption to our results is the focus of this subsection.

If the government can fully commit to future policy, then the maturity structure of government debt is irrelevant. It follows, according to our theory of QE, that QE is irrelevant as well. That full commitment renders the maturity structure of debt irrelevant is a central result in the classic paper of Lucas and Stokey (1983). It is especially simple to show the logic of their argument in the 3 period model of Section 2 but we relegate its derivation to Appendix A.3 along with some additional discussion. What we show there, formally, is that the flow budget constraints of the government, together with the ex-ante asset pricing conditions, allows us to write the government budget constraints as a single intertemporal budget constraint (IBC). The IBC, which is the relevant constraint under the optimal commitment plan, involves neither $b_0$ nor $l_0$ (and hence maturity of debt), but instead only the initial condition $b_{-1}$. Accordingly, the term structure imposes no constraints on the government’s optimal plan under commitment.

The bottom-line, then, is that the assumption of taxation costs as well as the assumption that the government cannot commit to future policy are the two most important ingredients driving the results.\[46\]

4.2 An Independent Central Bank and Its Balance Sheet

A natural objection to our model of QE is that fiscal considerations are often not considered to be part of a central bank’s mandate. This subsection shows an equivalence result in the model of Section 2. Suppose an independent central bank cares about its own balance sheet gains/losses in addition to the traditional dual mandate of inflation and output gap. Then, under certain conditions, the same equilibrium can be derived as if the government as a whole

\[46\] For completeness, we show in Figure A.1 in Appendix A.7.1, the solution for optimal joint monetary and fiscal policy under commitment vs. optimal monetary policy alone under commitment for the numerical example of Section 2. The simulation shows that the two are close. A similar result also applies in the quantitative model. The figure highlights that not much was missed in the discussion of our paper when we focused on optimal monetary policy alone under commitment. The optimal-policy-commitment solution was also shown in Figure 6 earlier in comparison with the optimal-discretion solution when $\lambda_T > 0$. 

38
sets monetary and fiscal policy jointly, which is the benchmark set-up of this paper.

Consider an independent central bank with the objective:

\[- \sum_{t=0}^{2} \beta^t \{(\Pi_t - 1)^2 + \lambda_Y (Y_t - \bar{Y})^2 + \lambda_v (T^c_t - \bar{T})^2\}, \]  

(46)

where \(T^c_t\) is its transfer to the Treasury and \(\bar{T}\) is its target transfer. There are several examples of objectives of this form in the literature.\(^{47}\) One motivation is political economy. In the US context, it seems reasonable that Congress would start asking questions if the Federal Reserve were to sustain large capital losses and not contribute the same amount to the Treasury as before.\(^{48}\) Another is that the central bank should care about capital losses, as they eventually must be followed by taxation (or higher inflation) which in turn affects social welfare.\(^{49}\)

Recently declassified minutes and memoranda from the Federal Reserve provide evidence suggesting its staff was concerned with balance sheet considerations following the crisis of 2008. One example is a memorandum prepared by Joseph Gagnon, David Lucca, Jonathan McCarthy, Julie Remeche, and Jennifer Roush for the Federal Reserve Open Market Committee (FOMC) on March 11, 2009.\(^{50}\) On page 9, the authors state:

However, to the extent that holding a large volume of long-term assets necessitates carrying a large volume of interest-earning liabilities, future Federal Reserve net income is subject to increased interest rate risk. In the example raised above, the extra Federal Reserve net income from holding $1 trillion of MBS [mortgage-backed securities] would shrink to zero if the rate of interest on reserve balances (or on reverse repos) were to rise to 4.25 percent.

Another example is a memorandum to the FOMC on April 21, 2009.\(^{51}\) Its authors, Eileen Mauskopf and Jae Sim, simulate the Federal Reserve Board’s FRB/US model to compute trade-offs related to balance sheet exposure and possible losses. On page 2, the authors state:

\(^{47}\)For example, Sims (2005) introduces the constraint that central bank transfers to the Treasury cannot fall below a target. Jeanne and Svensson (2007) instead directly introduce into the objective of the central bank one-time losses if the capital of the central bank falls below a certain level. Berriel and Bhattachary (2009) use a quadratic term in central bank net worth in the loss function. See also Del Negro and Sims (2015). Perhaps a more realistic version is an asymmetric objective. That is, we can posit that the central bank is happy to pay more to the Treasury than \(\bar{T}\) but does not like \(T^c_t\) to be below the target. This extension is possible but not particularly important for the QE question, as one can assume a \(\bar{T}\) sufficiently high that \(T^c_t < \bar{T}\) in all three periods or, alternatively, that the initial net worth of the central bank is sufficiently low.

\(^{48}\)Not to mention if, as in the case of Iceland in 2008, the central bank would need a large capital injection in response to enormous capital losses, which led the government to fire the entire board of the central bank.

\(^{49}\)Jeanne and Svensson (2007) and Berriel and Bhattachary (2009) provide anecdotal evidence from statements from several central bankers before the financial crisis, and they point to some recent central bank legislation.

\(^{50}\)https://www.federalreserve.gov/monetarypolicy/files/FOMC20090311memo02.pdf

Among the costs that we assume that the Committee will consider are those related to the volume of assets being brought onto the Federal Reserve’s balance sheet. For example, holding a large portfolio of long-term securities exposes the Federal Reserve (and thus taxpayers) to appreciable capital losses if interest rates rise quickly as the economy recovers.

Moreover, later in the QE period, other memoranda to the FOMC discussing the effects of further asset purchases not only consider macroeconomic effects, but also include simulation results concerning effects on the Federal Reserve’s balance sheet and income. An example is a memorandum to the FOMC on November 30, 2012. This memorandum discusses in detail how, given the simulations’ assumptions of interest rate increases in the future, additional asset purchases by the Federal Reserve could lead to capital losses, an increase in interest expense, and a decline in cumulative remittances to the Treasury between 2012 and 2015. We interpret this evidence as being consistent with the idea that avoiding large balance sheet losses was part of the the objective function of the FOMC during the QE period.

Motivated by this evidence, we model the balance sheet of an independent central bank in our 3 period model. In period 0, it has one-period liabilities \(d_0\) (interest-bearing reserves) and both one-period \((b_0)\) and two-period \((l_0)\) assets. In period 1, it issues one-period liabilities, \(d_1\), and buys one-period assets, \(b_1\). The initial net asset position of the central bank is \(l_{-1}\).

As in the analysis of a consolidated government, the budget constraints can be rewritten in terms of the minimum set of state variables in interest-inclusive terms. The state variable in period 0 is \(w_{-1}\); in period 1, the state variables are \(w_0\) and \(x_0\); and in period 2, the state variable is \(W_1\) defined as:

\[
W_1^c \equiv w_1^c + x_0 \frac{w_0^c}{\Pi_1},
\]

The budget constraint of the independent central bank can now be written as:

\[
0 = \frac{W_1^c}{\Pi_2} - T_2^c,
\]

\[
W_1^c = \frac{w_0^c}{\Pi_1} + x_0 (1 - x_0) \frac{w_0^c}{\Pi_1} - (1 + i_1) T_1^c,
\]

\[
x_0 \frac{w_0^c}{1 + R_0} + (1 - x_0) \frac{w_0^c}{1 + i_0} = W_{-1} + T_0^c,
\]

where \(w_0^c \equiv (1 + R_0) b_0 + (1 + i_0) (b_0^c - d_0^c)\), \(w_1^c \equiv (1 + i_1) (b_1^c - d_1^c)\), \(w_{-1}^c = l_{-1}\), \(x_0 \equiv \frac{(1 + R_0) b_0 + (1 + i_0) (b_0^c - d_0^c)}{(1 + R_0) b_0^c + (1 + i_0) (b_0^c - d_0^c)}\), and \((1 - x_0) \equiv \frac{(1 + i_0) (b_0^c - d_0^c)}{(1 + R_0) b_0^c + (1 + i_0) (b_0^c - d_0^c)}\). In this notation, \(w_0^c\) measures the net

---

52 https://www.federalreserve.gov/monetarypolicy/files/FOMC20121130memo05.pdf
asset position of the central bank in period 0, written in terms of maturity value. Meanwhile, \( x_0 \) measures the maturity composition of the central bank’s net asset position in period 0—the ratio of long-term assets to the net asset position. It thus measures the maturity mismatch in the central bank’s balance sheet in period 0. Detailed derivations are in Appendix D.

This leads to the following proposition:

**Proposition 4** Under the assumption that an independent central bank maximizes the objective \( J(\mathbf{y}) \) subject to (47)-(49), the problems of the consolidated government and independent central bank are identical if \( l_{-1} = -b_{-1} \) and \( \lambda_v = \lambda_T \).

The bottom-line, then, is that our model of a consolidated government budget constraint can alternatively be interpreted from the perspective of an independent central bank. In Appendix D.3 we discuss in more detail the condition in Proposition 4 that makes the two models exactly equivalent, and how this alternative model should lead researchers to interpret the data in somewhat different way than we have done here.\(^{53}\)

### 4.3 Forward Guidance

A policy of forward guidance is usually defined as verbal commitment by a central bank about future interest rate policy. In an MPE, however, verbal commitments have no effect. While it is extreme to assume that central bank statements have no effect on expectations, the MPE formalizes a common argument by policy makers against using forward guidance because the credibility problem reduces its effectiveness.\(^{54}\)

Even if one does not accept the extreme position that central banks’ words carry no weight, the analysis of the MPE is still of interest as it helps isolate what *actions* can be taken to enhance the credibility of forward guidance. There is considerable evidence that policy makers believe their statements alone are not fully credible.\(^{55}\) This suggests the importance of identifying such a commitment technology.

---

\(^{53}\)We make an initial attempt to clarify the relevant data for an independent central bank in our NBER WP version of this paper (No. 21336), but have omitted this extension here to conserve on space.

\(^{54}\)Nakata (2014) is an analysis of credible plans which incorporates reputation considerations. There is a related literature estimating the effect of forward guidance highlighting that it can either signal change in policy or information about future state of the economy (Campbell et al. (2012) and Andrade and Ferroni (2016)). Since our model assumes full information, it does not capture this interesting feature of forward guidance. If there was incomplete information, our result would still apply under the assumption that the government is committed to revealing the true state of the economy.

\(^{55}\)For example, John Williams, then president of the Federal Reserve Bank of San Francisco, noted this in an FOMC meeting in 2011 when discussing the possibility of adopting more aggressive forward-guidance language: “In the jargon of academics, our commitment technology is very limited. It is simply impossible for us to set a predetermined course of policy that will bind future Committees.” Similarly, the current chair of the Federal Reserve, Jay Powell, noted in the context of policies that involved committing to a future expansion at the ZLB: “Part of the problem is that when the time comes to deliver the inflationary stimulus, that policy is likely to be unpopular, what is known as the time consistency problem in economics.”
From a practical point of view, then, the analysis of the MPE in this paper can be interpreted as a study of how QE can be used as a commitment technology to supplement forward guidance. A central bank can make a particular statement about future interest rate policy. If this forward guidance is not believed, then the central bank can engage in QE until market expectations of future interest rates adjust according to the central bank’s intent.

4.4 Empirical Effects of Changing Debt Maturity

The empirical evidence on the effect of changes in the maturity of government debt on interest rates (outside of the financial crisis or ZLB) is relatively small. The existing evidence, however, is largely consistent with the model, even in rough orders of magnitude. A direct mapping to our model, however, is challenging. Previously, we focused on the degree to which the model matched the evidence at the ZLB presented by KVJ, because the mapping of their estimate can be directly translated to the model at the ZLB. Here we focus on the degree to which the model can account for the empirical evidence outside of the ZLB.

This evidence is summarized in Tables 5 and 6. Table 5 compares the implied reduction in long-term rates in response to a one-year maturity reduction, comparing the output of our model to the estimate by Chada et al. (2013a), who use a regression analysis to estimate the effect of debt-maturity reduction on interest rates. As the table reveals, the implied reduction in long rates according to our model (at positive interest rates) is less than one-eight of that predicted by their estimation, suggesting that our calibration is relatively conservative.

The evidence in Table 5 maps relatively cleanly into our analysis as we can compare the reduction in debt maturity in the study in question to our model. Comparison to other studies requires an additional assumption, since they typically only report the size of policy intervention in long-term bonds as a fraction of GDP. Thus, it is only under the assumption that the composition of long-term bonds purchased in the intervention is exactly the same as in the model that the magnitudes are comparable.

With this caveat in mind, Table 6 suggests, again, that the reduction in 10-year yields is a bit more conservative in our model relative to empirical studies we consider. Hamilton and Wu (2012) use an affine term structure model to estimate the effect of purchasing long-term...
Table 5: Effects of debt-maturity reduction on long-term rates outside the zero lower bound

<table>
<thead>
<tr>
<th>Source</th>
<th>Policy intervention</th>
<th>Change in 5-year forward 10-year rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chada et al. (2013a)</td>
<td>1-year maturity reduction</td>
<td>−130 to −150 bp</td>
</tr>
<tr>
<td>Out of ZLB model prediction</td>
<td>1-year maturity reduction</td>
<td>−17.6 bp</td>
</tr>
</tbody>
</table>

debt on long-term rates using pre-crisis data. As the table suggests, each percentage point of intervention (as a fraction of GDP) leads to a reduction in 10-year yields of about 4.8 bp. Swanson (2011) uses an event-study approach focusing on the Federal Reserve’s Operation Twist in 1961. As the table reveals, each percentage point of intervention (as a fraction of GDP) leads to a drop in 10-year yields of 7.6-9.4 bp. The last row in Table 6 considers the effect of QE2 in the case when interest rates are positive in our model. As it indicates, an intervention corresponding to 1 percent of GDP reduces 10-year yields by 3 basis points. Thus, the evidence reported in Table 6 is largely consistent with the one presented in Table 5. Our benchmark calibration yields a response of long-term yields that is lower than what these studies document.

Table 6: Effects of Quantitative Easing on long-term yields

<table>
<thead>
<tr>
<th>Source</th>
<th>Policy intervention</th>
<th>Change in 10-year yields</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamilton and Wu (2012)</td>
<td>2.9% of GDP, 2006</td>
<td>−14 bp</td>
</tr>
<tr>
<td>Swanson (2011)</td>
<td>1.7% of GDP, 1961</td>
<td>−13-16 bp</td>
</tr>
<tr>
<td>Model implied QE2 out of ZLB</td>
<td>4% of GDP, 2010</td>
<td>−12 bp</td>
</tr>
</tbody>
</table>

There is also empirical literature related to Figure 11, which shows transition dynamics for debt maturity and debt in the quantitative model. As the figure reveals, the government increases the maturity of its debt as it reduces its debt position. One interpretation of this result is that in a cross-section of countries, those with higher government debt tend to have lower debt maturity and higher inflation, assuming they are in a transition to steady-state. This is the empirical pattern documented in Blanchard and Missale (1994). Finally, Greenwood and Vayanos (2010) consider the effect of the US Treasury buyback program in 1999–2000, which was on the order of $63.5 billion. They argue that this intervention led to a substantial drop in long-term rates at various maturities, even for maturities of Treasuries that the government did not directly target.

58 Bordo and Sinha (2016) study an intervention by the Federal Reserve during the Great Depression which included the purchase of both medium- and long-term government debt and equaled 2% of GDP. They report a reduction in yield for Treasury bonds (with an average maturity of eighteen years) of 19 to 42 bp. As our focus here is on evidence away from the ZLB, we do not include this estimate in Tables 5 and 6.

59 This is, however, a different interpretation from the one emphasized by Missale and Blanchard (1994). A similar finding is reported by Rose and Spiegel (2015), who show that countries that have a bond market for long-term debt (and thus presumably have a longer debt maturity) tend to experience lower inflation.
did not buy back. Such an effect can be rationalized by our model, as QE leads to a decrease in expected short rates.

4.5 Related Theoretical Literature

A common theme in the literature on inflation and government debt is that more long-term debt gives the government a stronger incentive to inflate, not a weaker incentive as this paper shows (see e.g., Missale and Blanchard (1994)). Similarly, empirical analyses often stress that the fiscal benefits of inflation are higher the longer the maturity of debt (see e.g., Aizenman and Marion (2011), Doepke and Schneider (2008), and Hilscher et. al. (2018)). This line of research does not contradict our result. The reason for the seemingly different findings is that these studies consider the fiscal benefits of inflation over an extended period—longer than the duration of the short-term debt and often of similar duration as the long-term debt. This results in a fundamentally different game than the one in this paper, where the government re-optimizes period by period. In Appendix E we flesh out this argument in detail.

The dynamic inconsistency problem created by the existence of government debt when taxation is costly is a theme of Lucas and Stokey (1983) and the rich subsequent literature. Lucas and Stokey’s solution to the problem is a careful manipulation of the maturity structure of debt, i.e., going long, making the optimal solution time consistent. Persson et al. (1987, 2006), Calvo and Obstfeld (1990), and Alvarez et al. (2004) show under what conditions a similar strategy can be implemented (and when it cannot) in a flexible price monetary economy. Calvo and Guidotti (1990,1992) are two other examples that touch on the same themes. Broadly speaking, the key idea in this paper can be interpreted as the Lucas and Stokey (1983) solution in reverse. While their strategy is to lengthen the duration of government debt in order to eliminate the government’s incentive to reduce the real interest rate in the future, QE accomplishes the opposite in order to give the government the incentive to keep real interest rates low in the future, as needed at the ZLB. Unlike this literature, this paper considers both the ZLB and price rigidities – these two frictions are fundamental for generating the results.

This paper follows the work cited above by assuming the government never defaults on its debt. This is a reasonable assumption for large countries that issue debt in their own

---

60 Missale and Blanchard (1994) directly assume a reduced-form sensitivity of debt to inflation that is decreasing in the maturity of debt, and the actions of the government in the game they present are best interpreted as corresponding to the actions of a government choosing an inflation policy regime over an extended period (see Appendix E for more details). In Aizenman and Marion (2011), the duration of the inflation policy is the same as that of long-term debt. Hilscher et al. (2018) consider the empirically estimated expected-inflation scenarios extracted from asset markets, while Doepke and Schneider (2008) consider an unexpected increase in inflation for ten years.

61 Calvo (1978) represents an important early paper in this literature.

62 There is also a literature emphasizing that, in the absence of state-contingent bonds, the maturity structure of debt can be used to hedge against fiscal shocks; see Angeletos (2002) and Leeper and Zhou (2013).
currencies. This assumption, however, is less plausible for small open economies that contract their debt in foreign currency. There is a considerable literature that focuses on small open-economy models which allow for the possibility that the government defaults in certain states of the world; see, e.g., Aguiar et al. (2019) and Arellano and Ramanarayanan (2012). This branch of the literature shows that the possibility of default also leads to a dynamic-inconsistency problem for the government. Moreover, it shows that shortening the maturity structure of government debt alleviates the dynamic inconsistency problem by creating fiscal discipline for future governments. This may seem to contradict the closed economy literature which instead suggests that the government should lengthen its debt maturity to avoid the incentive to inflate. In followup work, Bhattarai et al. (2022), we show how the two literatures can be reconciled. The incentive mechanism explored in the open economy literature is closely related to what we term the \textit{rollover incentive} in Section 2.3.2; see Bhattarai et al. (2022) for details.

Another argument for the effect of QE is that it works through a portfolio balance channel. The most common way of modeling the portfolio balance channel is via preferred habitat motives and market segmentation. Chen et al. (2012) is an example that incorporates such a friction, and finds a small role for the portfolio balance effect, with most effects coming via a commitment to hold future interest rates low, a mechanism formalized here.\footnote{There is no portfolio balance channel in this paper. This is not, however, because long- and short-term bonds are perfect substitutes. As explained in Eggertsson and Woodford (2003)(see p. 159), the model can be extended to incorporate different risk characteristics of long- and short-term bonds and, thus, a term- and risk-premia. That extension will not change the results in this paper for the same reasons discussed there.}

Other papers, such as Gertler and Karadi (2011, 2012) and Del Negro et al. (2019), provide frameworks in which asset purchases by the central bank have an effect because of financial imperfections and limits to arbitrage. Those mechanisms can account for a substantial effect of QE1, which included purchases of private securities at a time of considerable market dysfunction. It is more difficult, however, to make such a case for QE2 and QE3, when the market dysfunction was much smaller and the policy mostly involved buying long-term government bonds. This is why QE2 is the benchmark for the model calibration in this paper.

5 Conclusion

This paper develops a model of QE policy in which the maturity of government debt decreases. Faced with a binding ZLB and ensuing recession and deflation, a government policy that shortens government debt maturity improves on outcomes, as QE generates expectations of low future real interest rates. The key ingredient for the mechanism is that when policy is set in a time-consistent manner, there exists a role for manipulating the maturity of government
debt to generate a credible commitment to future expansionary policy. We first showed these results theoretically in a simple setting and then assessed their quantitative relevance.

Future work can extend ours both empirically and theoretically. While the focus in this paper was on QE episodes, the mechanism also holds in reverse. For instance, during the “taper tantrum” episode in 2013, when Ben Bernanke announced that the Fed would taper QE in the future, federal funds futures suggested markets expected future short-term rates to rise.\footnote{We are indebted to Jim Bullard, president of the Federal Reserve Bank of St. Louis, for educating us on this empirical pattern in an insightful discussion of our paper.} We leave a further analysis of such episodes for future research.

Data Availability Statement: The data and code underlying this research is available on Zenodo at \url{https://dx.doi.org/10.5281/zenodo.7007875}. 
References


50


