

NBER WORKING PAPER SERIES

INFREQUENT BUT LONG-LIVED ZERO-BOUND EPISODES AND THE OPTIMAL RATE
OF INFLATION

Marc Dordal-i-Carreras
Olivier Coibion
Yuriy Gorodnichenko
Johannes Wieland

Working Paper 22510
<http://www.nber.org/papers/w22510>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2016

Forthcoming in Annual Review of Economics, doi: 10.1146/annurev-economics-080315-015306. We are grateful to seminar participants at Berkeley and IMF for comments on an earlier version of the paper. Yuriy Gorodnichenko thanks the NSF and Sloan Foundation for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Marc Dordal-i-Carreras, Olivier Coibion, Yuriy Gorodnichenko, and Johannes Wieland. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Infrequent but Long-Lived Zero-Bound Episodes and the Optimal Rate of Inflation
Marc Dordal-i-Carreras, Olivier Coibion, Yuriy Gorodnichenko, and Johannes Wieland
NBER Working Paper No. 22510
August 2016
JEL No. E3,E4,E5

ABSTRACT

Countries rarely hit the zero-lower bound on interest rates, but when they do, these episodes tend to be very long-lived. These two features are difficult to jointly incorporate into macroeconomic models using typical representations of shock processes. We introduce a regime switching representation of risk premium shocks into an otherwise standard New Keynesian model to generate a realistic distribution of ZLB durations. We discuss what different calibrations of this model imply for optimal inflation rates.

Marc Dordal-i-Carreras
Department of Economics
University of California at Berkeley
Berkeley, CA 94720-3880
marcdordal@gmail.com

Olivier Coibion
2225 Speedway
Economics Department
University of Texas at Austin
Austin, TX 78712
and NBER
ocoibion@gmail.com

Yuriy Gorodnichenko
Department of Economics
530 Evans Hall #3880
University of California, Berkeley
Berkeley, CA 94720-3880
and IZA
and also NBER
ygorodni@econ.berkeley.edu

Johannes Wieland
Department of Economics
University of California at San Diego
9500 Gilman Drive #0508
La Jolla, CA 92093
and NBER
jfwieland@ucsd.edu

1 Introduction

When the U.S. Federal Reserve finally raised its target for the Federal Funds Rate in December 2015, this likely marked the end of the zero-bound on short-term nominal interest rates for the United States after a staggering *seven years*. Japan's zero-bound period will most likely exceed this duration under Abenomics, while the Bank of England has similarly had near-zero interest rates since March of 2009. The Euro Central Bank is also not expected to raise interest rates for years. Combined with the previous experiences with the zero-bound on interest rates that occurred during the Great Depression and in Japan during the 1990s-2000s, this suggests that the two most prominent empirical features of zero-bound episodes are that they are *rare* but *long-lived*.

The zero-bound on interest rates raises a number of profound problems for monetary policymakers, one of which is the traditional question of what the optimal inflation rate should be. While it is well-understood that even stable inflation has costs (such as those arising from price dispersion), higher average inflation is also associated with higher nominal interest rates, which can benefit policymakers by giving them extra room to avoid running into the zero-bound. Quantifying the optimal rate of inflation then requires balancing the costs of inflation against its benefits, such as minimizing the frequency and severity of zero lower bound (ZLB) episodes.

But quantifying this potential benefit of higher inflation is difficult because the paucity of ZLB episodes makes their frequency and duration hard to gauge. For example, Schmidt-Grohe and Uribe (2010) calibrated their model prior to the start of the Great Recession and had no post-WWII zero-bound episodes in the U.S. to guide their choice over the frequency of hitting the zero-bound. This resulted in a calibration with very rare and short-lived ZLB episodes. Coibion, Gorodnichenko and Wieland (2012) used the fact that the U.S. had spent 3 years at the ZLB at the time of their writing out of the post-WWII period to fix their frequency, yielding more frequent but still mostly short-lived episodes. In each case, these authors conclude that the optimal rate of inflation is unlikely to be much above 2% despite the zero-bound on interest rates. But given the actual durations of the most recent ZLB experienced by developed economies, each of these papers likely underestimated the average duration of ZLB episodes and therefore the potential benefits of higher levels of target inflation on the part of central banks.

In this paper, we revisit the topic of longer-lived ZLB episodes in two steps. First, following previous work, we generate longer-lived ZLB episodes by either increasing the persistence or the volatility of AR(1) risk-premium shocks which push the economy into the ZLB in our model. By doing so, we can generate a longer average duration of ZLB episodes consistent with the data. In our benchmark New Keynesian model, increasing the average duration of ZLB episodes (for a given steady-state level of inflation) through either more persistent or more volatile shocks can have large positive effects on the optimal inflation rate. For example, moving from an average duration of ZLB episodes of 5 quarters to just 6 quarters, holding the frequency of ZLB episodes fixed (by varying the persistence and volatility of shocks accordingly), can raise the optimal inflation rate from 1.3% to 2.2% in our baseline calibration, a very high sensitivity.

This sensitivity, however, reflects an unappealing characteristic common to all standard models of the ZLB in which normally distributed shocks drive the economy into the ZLB: the vast majority of ZLB episodes in the model are extremely short-lived. Significantly raising the average duration therefore requires large tail events, and these episodes have disproportionately large welfare costs. Policymakers become very willing to tolerate higher average inflation rates to avoid these episodes, leading to significantly higher optimal rates of inflation for even small changes in average durations of ZLB

episodes. Hence, the high sensitivity of the optimal rate of inflation to the average duration of ZLB episodes in our benchmark model is a reflection of the counterfactual distribution of ZLB episodes, namely that they are too frequent and short-lived compared to the rare and long-lived episodes that we observe in the data.

Our second step is then to incorporate an alternative modeling strategy for the shocks that drive the economy into the ZLB which generates an empirically realistic distribution of ZLB episodes, namely that they tend to be rare but long-lived. We assume that each period, risk premia follow a regime-switching process. In each period when the economy is not at the ZLB, there is a fixed probability that the risk premium will rise sharply for a set number of periods. If the increase in the premium is large enough, this shock will give rise to a distribution with very long-lived ZLB durations, thereby more closely representing the empirical distribution of ZLB episodes. As a result, we can more carefully assess whether (or how much) raising the rate of inflation is optimal, in a welfare sense, to offset the presence of the zero bound on interest rates.

Unlike the AR(1) model, the regime switching approach does not display an excessive sensitivity of the optimal inflation rate to the average duration of ZLB episodes. Nonetheless, long-lived ZLB episodes generate large welfare costs in the model, which higher levels of steady state inflation can help avoid by reducing their frequency. We find that depending on our calibration of the average duration and the unconditional frequency of ZLB episodes, the optimal inflation rate can range from 1.5% to 4%. This uncertainty stems ultimately from the paucity of historical experience with ZLB episodes, which makes pinning down these parameters with any degree of confidence very difficult. A key conclusion of the paper is therefore that much humility is called for when making recommendations about the optimal rate of inflation since this fundamental data constraint is unlikely to be relaxed anytime soon.

Our paper builds on a broad literature on the optimal rate of inflation. This literature has covered a wide range of costs and benefits, with the zero bound on interests only recently coming to the forefront as a plausible source for positive optimal rates of inflation. In a survey of pre-Great Recession work, Schmidt-Grohe and Uribe (2010) highlighted that, although the quantitative conclusions about the optimal rate of inflation were potentially sensitive to the choice of the model used to assess the costs and benefits of inflation (or deflation) in the steady-state, one generally found that it was optimal to have a small amount of deflation. For example, using a standard model with demand for money, Schmidt-Grohe and Uribe (2010) estimated the optimal inflation rate at -0.6 percent per year with the Ramsey optimal policy. While this rate of deflation is considerably smaller than the rate of deflation originally suggested by Milton Friedman (approximately equal to the real interest rate), the optimality of deflation in steady state is inconsistent with the 1-3%/year inflation rates currently targeted by most central banks.

Even when one moves to cashless economies, it is difficult to push the optimal rate of inflation near the levels commonly targeted by modern central targets using traditional arguments for positive levels of inflation such as downward wage rigidity. Indeed, New Keynesian models generally suggest that the optimal rate of inflation should be close to zero because price dispersion generated by non-zero trend inflation is costly (see Benigno and Woodford 2005).¹ However, one may be able to raise the optimal rate of inflation by departing from the workhorse specifications to incorporate e.g. foreign demand for currency (Schmidt-Grohe and Uribe 2012), firm-specific productivity growth (Weber 2012), occasionally binding financial constraints (Abo-Zaid 2015), or tax evasion (Schmidt-Grohe and Uribe 2010).

¹ Wolman (2011) and others show that even in the absence of money demand considerations, the optimal rate of inflation in New Keynesian models can be negative. For example, in Wolman's model and calibration the optimal deflation is 0.4% per year.

Pre-Great Recession work (e.g., Summers 1991) discussed the zero lower bound (ZLB) on nominal interest rates as a potential reason for positive inflation but generally considered the ZLB as an improbable event. As a result, the models were calibrated to generate infrequent and short-lived ZLB episodes. For example, Schmidt-Grohe and Uribe (2010) indicate that to violate the zero bound “...the nominal interest rate ... must fall more than 4 standard deviations below its target level” thus making ZLB an extremely rare event.

Many others found similar results. For example, Reifschneider and Williams (2000) and Chung et al. (2012) document that the frequency of ZLB for three popular dynamic stochastic general equilibrium (DSGE) models estimated on the post-WWII, pre-2007 data is typically less than 5 percent.² Furthermore, ZLB episodes longer than 8 quarters can be observed less than 1 percent of the time. If one uses data from the Great Moderation period to assign parameters in DSGE models, ZLB episodes are even shorter and less frequent. Similarly, Adam and Billi (2007) find that, with optimal monetary policy, conditional on hitting ZLB, the likelihood of being at the ZLB for more than 4 quarters is a mere 1.8%. In Billi’s (2011) calibration, the ZLB binds 4 percent of the time and the average duration of ZLB is only 2 quarters. Using non-linear methods to solve and simulate calibrated DSGE models, Amano and Shukayev (2012) report that the probability of hitting the ZLB is 1.7% per quarter (i.e., a 4-quarter ZLB episode occurs once every 60 years). In other words, ZLBs in models used by researchers and policymakers were too short and too rare to matter.

As the welfare costs of short ZLB episodes tend to be small, the ZLB was found to have tiny effects on the estimated optimal rate of inflation. For example, the optimal rate of inflation in the Schmidt-Grohe and Uribe (2010) calibration increased modestly from -0.6 to -0.4 percent per year in light of the ZLB. In short, the consensus view before the Great Recession was that, although the ZLB was an interesting and curious possibility, one could treat it as remote and largely irrelevant.

With policy interest rates in major developed economies having spent years at the zero lower-bound during the Great Recession and its aftermath, there has of course been a shift in thinking about the frequency and nature of the ZLB. Examination of new cross-country evidence and long time series (i.e., series including the Great Depression) suggests that ZLB episodes are potentially costly (e.g., Williams (2009) estimated that four years at the ZLB can cost as much as \$1.8 trillion), more frequent (e.g., Chun et al. (2012) indicate that, based on pre-2010 data, one should double the probability of being at the ZLB in calibrated models), and more persistent. The latter point is particularly important as ZLB episodes in the U.S. and elsewhere are not characterized by a series of short intervals of constrained policy rates. Instead, the Great Depression and the Great Recession in the U.S. or the crash in Japan indicate that ZLB episodes can last for years if not decades.

Incorporating these changes in the way we model the ZLB can have dramatic effects on the optimal rate of inflation. Indeed, apart from ZLB episodes being modeled as more frequent and thus costlier, we know from Coibion et al. (2012) and others that the cost of ZLB is increasing steeply in its duration. That is, an 8-quarter ZLB is costlier than two 4-quarter ZLB episodes. Thus, the cost of ZLB in a new calibration can be considerably larger than in previous calibrations and can entail an optimal rate of inflation higher than the conventionally suggested 2 percent per year.

While the treatment of the ZLB is one important source of differences in the estimated optimal rate of inflation, there are other factors and modelling choices that can affect the optimal rate. For example, Coibion et al. (2012) show that how one models price stickiness can also influence results.

² Coibion et al. (2012) calibrate the frequency of ZLB in the basic New Keynesian model at 5 percent.

Using the Calvo (1983) approach tends to yield a lower optimal inflation rate because Calvo-style pricing generates a larger increase in cross-sectional price dispersion for a given increase in trend inflation than e.g. Taylor (1977) pricing.³ Intuitively, firms with Calvo pricing may be stuck at suboptimal prices for a long time while Taylor pricing guarantees that prices can be reset after a fixed number of periods which caps the size of departures from optimal levels of prices. Because cross-sectional price dispersion is the main cost of non-zero steady-state inflation in New Keynesian models, the choice of pricing assumptions can alter the point at which the cost of positive inflation balances the benefit of positive inflation (e.g., avoid ZLB). Consistent with this logic, Coibion et al. (2012) find the optimal rate of inflation to be 1.5% under Calvo pricing (when the probability of price adjustment is set at 0.55) and 1.8% under Taylor pricing (when the duration of contracts is set at 3 quarters).

In a similar spirit, menu-cost models limit the degree of cross-sectional price dispersion (since a firm can reset its price whenever it deviates too far from the optimal price) and thus could reduce the cost of non-zero steady-state inflation. As a result, it may be optimal in such models to target a higher rate of inflation which will reduce the probability of hitting the ZLB, but the exact magnitude depends on the details of menu cost models. While the optimal rate of inflation in the Dotsey et al. (1999) model is below 2 percent (see Coibion et al. 2012), Blanco (2015) found that in the Golosov and Lucas (2007) model welfare is maximized at approximately 5%/year inflation rate. Because the computational demands become exceedingly high for long-lasting ZLB periods even in linearized models, we will focus on the Calvo approach to model price stickiness.

The paper is organized as follows. Section 2 presents the model and the two ways of modeling shocks that drive the economy into the zero bound. Section 3 then presents the main results of the paper, including comparing the distribution of ZLB episodes under the two assumptions about shock processes and their implications for optimal inflation. Section 4 concludes.

2 Model

In our quantitative analyses, we use the standard New Keynesian model similar to the framework in Coibion, Gorodnichenko and Wieland (2012). To preserve space, we describe the main building blocks of the model and relegate derivations and various details to the Appendix.

2.1. Households

The representative consumer maximizes the present discounted value of the utility stream from consumption and leisure

$$\max E_t \sum_{j=0}^{\infty} \beta^j \left\{ \log(C_{t+j} - h \cdot GA_{t+j} \cdot C_{t+j-1}) - \frac{\eta}{\eta+1} \int_0^1 N_{t+j}(i)^{1+1/\eta} di \right\} \quad (1)$$

where C is consumption of the final good, $N(i)$ is labor supplied to individual industry i , GA is the gross growth rate of technology, η is the Frisch labor supply elasticity, h the internal habit parameter and β is the discount factor. The budget constraint in each period t is given by

$$\xi_t: C_t + \frac{S_t}{P_t} + T_t \leq \int_0^1 \left(\frac{N_t(i)W_t(i)}{P_t} \right) di + \frac{S_{t-1}q_{t-1}R_{t-1}}{P_t} + \Gamma_t \quad (2)$$

³ Using a medium-scale DSGE model, Ascari, Phaneuf, and Sims (2015) estimate that a consumption-equivalent welfare loss from of raising inflation from 2% to 4% can be as large as 7 percent.

where S is the stock of one-period bonds held by the consumer, R is the gross nominal interest rate, P is the price of the final good, $W(i)$ is the nominal wage earned from labor in industry i , T is real lump sum taxation (or transfers), Γ are real profits from ownership of firms, q is a risk premium shock, and ξ is the shadow value of wealth (i.e., the Lagrange multiplier on constraint (2)). As we discuss below, the risk premium shock plays a central role in generating binding ZLB.

2.2. Firms

For each intermediate good $i \in [0,1]$, a monopolist generates output using a production function linear in labor

$$Y_t(i) = A_t N_t(i) \quad (3)$$

where A denotes the level of technology, common across firms. The time series of technology is described by a random walk process: $A_t = \exp(u_t^A)$, $u_t^A = \mu + u_{t-1}^A + \varepsilon_t^A$ with $\varepsilon_t^A \sim iid N(0, \sigma_A^2)$. Parameter μ sets the average growth rate of technology in the model.

A perfectly competitive sector combines intermediate goods into a final good using the Dixit-Stiglitz aggregator

$$Y_t = \left[\int_0^1 Y_t(i)^{(\theta-1)/\theta} di \right]^{\theta/(\theta-1)} \quad (4)$$

where Y is the final good and θ denotes the elasticity of substitution across intermediate goods, yielding the following demand curve for goods of intermediate sector i :

$$Y_t(i) = Y_t (P_t(i)/P_t)^{-\theta} \quad (5)$$

and the following expression for the aggregate price level

$$P_t = \left[\int_0^1 P_t(i)^{(1-\theta)} di \right]^{1/(1-\theta)}. \quad (6)$$

Each intermediate good producer has sticky prices, modeled as in Calvo (1983) where $1 - \lambda$ is the probability that each firm will be able to reoptimize its price each period. Denoting the optimal reset price of firm i by B (all firms choose the same rest price), re-optimizing firms solve the following profit-maximization problem

$$\max_{B_t(i)} E_t \sum_{j=0}^{\infty} \lambda^j Q_{t,t+j} [Y_{t+j}(i) B_t(i) - W_{t+j}(i) N_{t+j}(i)] \quad (7)$$

where $Q_{t,t+j} = \beta^j E_t \left\{ \frac{\xi_{t+j} P_t}{\xi_t P_{t+j}} \right\}$ is the stochastic discount factor. The optimal reset price B_t is then given by

$$\frac{B_t}{P_t} = \frac{E_t \sum_{j=0}^{\infty} \lambda^j Q_{t,t+j} Y_{t+j} \left(\frac{P_{t+j}}{P_t} \right)^{\theta+1} \left(\frac{\theta}{\theta-1} \right) (MC_{t+j}(i)/P_{t+j})}{E_t \sum_{j=0}^{\infty} \lambda^j Q_{t,t+j} Y_{t+j} (P_{t+j}/P_t)^\theta} \quad (8)$$

where $MC_t(i) = \frac{W_t(i)}{A_t}$ is the marginal cost of firm i .⁴

Given these price-setting assumptions and price index in (6), the dynamics of the price level are governed by

$$P_t^{1-\theta} = (1-\lambda)B_t^{1-\theta} + \lambda P_{t-1}^{1-\theta}. \quad (9)$$

Firms' aggregate real profits are

$$\begin{aligned} \Gamma_t &= \int_0^1 \Gamma_t(i) di = \frac{1}{P_t} \int_0^1 [P_t(i)Y_t(i) - N_t(i)W_t(i)] di \\ &= Y_t - \int_0^1 \left(\frac{N_t(i)W_t(i)}{P_t} \right) di. \end{aligned} \quad (10)$$

We define the aggregate labor input as

$$N_t = \left[\int_0^1 N_t(i)^{(\theta-1)/\theta} di \right]^{\theta/(\theta-1)} = \left[\int_0^1 \left(\frac{Y_t(i)}{A_t} \right)^{(\theta-1)/\theta} di \right]^{\theta/(\theta-1)} = \frac{Y_t}{A_t}. \quad (11)$$

2.3. Government

We allow for government consumption of final goods (G) with the good market clearing condition $Y_t = C_t + G_t$. The government budget constraint is defined as

$$T_t + S_t/P_t = G_t + \frac{S_{t-1}q_{t-1}R_{t-1}}{P_t}, \quad (12)$$

where $G_t = \bar{G}_t \exp(u_t^G)$, \bar{G}_t is the path of government spending such that the share of government spending in the economy is fixed when prices are flexible, and u_t^G is an exogenous, forcing variable: $u_t^G = \rho_G u_{t-1}^G + \varepsilon_t^G$ with $\varepsilon_t^G \sim iid N(0, \sigma_G^2)$.

The policy rule followed by the central bank is

$$R_t = \max\{1, R_t^*\} \quad (13)$$

$$R_t^* = \bar{R} \left(\frac{R_{t-1}^*}{\bar{R}} \right)^{\rho_1} \left(\frac{R_{t-2}^*}{\bar{R}} \right)^{\rho_2} \left[\left(\frac{\Pi_t}{\bar{\Pi}} \right)^{\phi_\pi} \left(\frac{Y_t}{\bar{Y}_t} \right)^{\phi_Y} \left(\frac{GY_t}{\bar{G}\bar{Y}} \right)^{\phi_{GY}} \right]^{(1-\rho_1-\rho_2)} \exp(\varepsilon_t^R) \quad (14)$$

where R is the *realized* gross interest rate, R^* is the *desired* gross interest rate, GY is the gross growth rate of output, $\bar{\Pi}$ is the gross, steady-state level of inflation, $\bar{G}\bar{Y}$ is the steady state growth rate of output, \bar{Y}_t is the flexible-price level of output, \bar{R} is the steady state nominal interest rate, and ε^R is an i.i.d policy shock. Equation (13) is responsible for introducing the zero lower bound to the model. We abstract from alternative monetary policy actions during ZLB episodes, such as quantitative easing. While these could

⁴ Labor employed by firm i each period is obtained through the minimization of production costs.

potentially lower the costs of ZLB episodes, there is little evidence suggesting that these policies have had large economic effects.⁵

2.4. Risk premium shocks

As discussed in Amano and Shukayev (2012), the risk premium shock is the main “tool” that can generate a binding ZLB in standard New Keynesian models. To be clear, this shock should be interpreted broadly as capturing a variety of forces that bring interest rates to ultra-low levels. We consider two general approaches to model the dynamics of the shock.

The first approach is to describe the time series of the shock as an AR(1) process similar to what is usually assumed for other forcing variables in DSGE models (e.g., Coibion et al. 2012):

$$q_t = \exp(u_t^q), u_t^q = \rho_q u_{t-1}^q + \varepsilon_t^q \text{ with } \varepsilon_t^q \sim iid N(0, \sigma_q^2). \quad (15)$$

By adjusting ρ_q and σ_q^2 , one can regulate the frequency and duration of ZLB episodes. As we will show later, a major shortcoming of this approach to modeling the risk premium is that it cannot replicate the main qualitative empirical properties of ZLB episodes, namely that they are rare but long-lived. Instead, AR(1) shocks primarily deliver frequent and short-lived ZLB episodes.

As a result, we also consider a second approach which allows for two regimes of risk premia. For example, Christiano, Eichenbaum and Rebelo (2011), Eggertsson and Woodford (2003), and Guerrieri and Lorenzoni (2009) assume that the ZLB is binding for a fixed number of periods or that, conditional on being at the ZLB, every period there is a random, i.i.d. draw determining exit from the ZLB; that is, with some probability the risk premium declines from a high level (ZLB is binding) to a low level (ZLB is not binding). This line of work typically assumes that after exiting ZLB the economy does not return to it.

To permit recurrent ZLB episodes, we consider the following regime-switching process. The risk premium can take two values: zero and $\Delta > 0$. Each period when the risk premium is zero, there is a random, i.i.d. draw such that with probability p_{12} the risk premium switches from zero to Δ and stays at this elevated level for T_q periods. After T_q periods with low interest rates, the risk premium returns to zero. By varying Δ, p_{12}, T_q , we can obtain variation in the frequency and duration of ZLB. Note that $\Delta > 0$ does not guarantee that the interest rate will be literally stuck at zero: other shocks (e.g., productivity) can lift the economy off the ZLB. However, by making Δ large enough, we can reduce the incidence of

⁵ Coibion, Gorodnichenko and Wieland (2012, CGW henceforth) examine how the optimal rate of inflation varies if the central bank can implement an optimal stabilization policy with commitment. One can think of the commitment policy as introducing a very powerful form of forward guidance. CGW find that in this case the optimal rate of inflation shrinks to zero considerably. Intuitively, with a strong form of forward guidance delivered by fully credible commitment to keep low interest rates far into the future, the stabilization powers of monetary policy remain large unaffected by the ZLB. As a result, there is no need for a “cushion” created by positive trend inflation. CGW also show that if a Taylor rule includes an element of price level targeting, the central bank can nearly achieve the welfare one can obtain under the optimal policy with commitment because current below-target inflation is compensated with above-target inflation in the future. Since our objective is to consider scenarios that should push up the optimal rate of inflation (most importantly, increase in the duration of ZLB episodes), we do not cover the optimal policy with commitment as these move the optimal rate of inflation in the opposite direction.

such lift-offs. We solve the model by adapting the solution algorithm in Coibion et al. (2012) to these deterministic regime-switching processes.⁶

While the difference in modeling the risk premium shock may seem subtle, these two approaches can generate different distributions for ZLB durations with important implications for calculating welfare losses arising from binding ZLB. As we demonstrate below, the AR(1) approach tends to yield frequent, short-lived ZLB episodes. Such a distribution of ZLB episodes appears to be inconsistent with the experience of the U.S. and other developed economies during the Great Recession or in other instances. In contrast, the regime-switching approach can produce long-lived ZLB episodes, similar to what we observe in the data.

2.5. Log-linearized system

Using lower-case letters with hats to denote variables log-linearized around the stochastic trend in technology, we can summarize the system of optimality conditions and budget constraints by the familiar equations.

Phillips curve:

$$\begin{aligned} \left(1 + \frac{\theta}{\eta}\right) \left(\frac{\lambda \bar{\Pi}^{(\theta-1)}}{1 - \lambda \bar{\Pi}^{(\theta-1)}}\right) \hat{\pi}_t &= \sum_{j=0}^{\infty} [\gamma_2^j (1 - \gamma_2) - \gamma_1^j (1 - \gamma_1)] [\hat{y}_{t+j} + \hat{\xi}_{t+j}] \\ &+ (1 - \gamma_2) \sum_{j=0}^{\infty} \gamma_2^j \left[\frac{1}{\eta} \hat{y}_{t+j} - \hat{\xi}_{t+j}\right] \\ &+ \sum_{j=0}^{\infty} \left[\gamma_2^{j+1} \theta \left(1 + \frac{1}{\eta}\right) - \gamma_1^{j+1} (\theta - 1)\right] E_t[\hat{\pi}_{t+j+1}] + \hat{u}_t^m, \end{aligned} \quad (16)$$

where $\gamma_1 = \lambda \beta \bar{\Pi}^{(\theta-1)}$ and $\gamma_2 = \gamma_1 \bar{\Pi}^{(1+\theta/\eta)}$ and \hat{u}_t^m is an ad hoc cost-push shock such that $\hat{u}_t^m = \rho_m \hat{u}_{t-1}^m + \varepsilon_t^m$ and $\varepsilon_t^m \sim iid N(0, \sigma_m^2)$.

IS curve (consumption Euler equation):

$$\hat{\xi}_t = E_t[\hat{\xi}_{t+1} + \hat{r}_t - \hat{\pi}_{t+1} + \hat{u}_t^q], \quad (17)$$

where $\hat{\xi}_t = \frac{h}{(1-h)(1-\beta h)} \hat{c}_{t-1} - \frac{1+\beta h^2}{(1-h)(1-\beta h)} \hat{c}_t + \frac{\beta h}{(1-h)(1-\beta h)} E_t \hat{c}_{t+1}$.

Taylor rule:

$$\hat{r}_t = \max\{\hat{r}_t^*, -\bar{r}\}, \quad (18)$$

$$\hat{r}_t^* = \rho_1 \hat{r}_{t-1}^* + \rho_2 \hat{r}_{t-2}^* + (1 - \rho_1 - \rho_2) [\phi_\pi \hat{\pi}_t + \phi_y \hat{y}_t + \phi_{gy} \widehat{gy}_t] + \varepsilon_t^r,$$

where $\widehat{gy}_t = \hat{y}_t - \hat{y}_{t-1} + \varepsilon_t^A$ is the log-linearized growth rate of output.

Market clearing:

⁶ We fix the duration T_q so we only have to solve backward once from period $t + T_q$. By contrast, if the exit were stochastic we would have to solve backward from every possibly realization and weigh these paths by their probability.

$$\hat{y}_t = (1 - s_G)\hat{c}_t - s_G\hat{g}_t, \quad (19)$$

where $s_G = \bar{G}_t/\bar{Y}_t$.

2.6. Welfare

Proposition 1 in Coibion et al. (2012) derives the second order approximation to expected per period utility in eq. (1) when steady state inflation is different from zero:

$$\Theta_0 + \Theta_1\text{var}(\hat{y}_t) + \Theta_2\text{var}(\hat{\pi}_t) + \Theta_3\text{var}(\hat{c}_t) \quad (19)$$

where parameters $\Theta_k, k = \{0,1,2,3\}$ depend on the steady state inflation $\bar{\pi}$. As discussed in Coibion et al. (2012), this approximation has an intuitive interpretation and properties. The term Θ_0 captures the cost of cross-sectional price dispersion arising from positive trend inflation. For quantitatively relevant inflation rates, Θ_0 becomes more negative as steady-state inflation increases. Because of the functional assumption about the household's utility, $\Theta_1 < 0$ but Θ_1 does not depend directly on steady-state inflation. The coefficient on the variance of inflation $\Theta_2 < 0$, which is the main cost of business cycle in the standard New Keynesian model like ours, is decreasing in steady state inflation. Finally, the coefficient on the variance of consumption $\Theta_3 < 0$ captures the desire of habit-driven consumers to smooth consumption.

2.7. Calibration

We calibrate the model as in Coibion et al. (2012), see Table 1. This parametrization uses values standard in the literature. Parameter values governing the frequency and duration of ZLB (that is, ρ_q, σ_q^2 for the AR(1) model and Δ, p_{12}, T_q for the regime switching model) are harder to pin down because ZLB episodes are rare. Consequently, we will consider combinations of parameter values that yield a spectrum of durations and unconditional frequencies of ZLB episodes. As a baseline, we will focus on parameter values that generate an unconditional frequency of the ZLB equal to 10%, which corresponds to the U.S. post-WWII experience (seven years at the ZLB over seventy years), although we relax this assumption later on. In the case of the regime switching model, we have an extra free parameter. As a baseline, we choose to set $\Delta = 0.090$ to ensure that the risk premium shock almost always yields a binding ZLB. For robustness, we will also consider two additional values of Δ . One is based on setting $\Delta = \bar{R} - 1 = \frac{1}{\beta}\bar{\Pi}(1 + \mu) - 1$. That is, the size of the premium is equal to the steady-state level of the nominal rate, which in turn depends on the time preference parameter β , the steady state level of inflation $\bar{\Pi}$, and the growth rate of output (and technology) in the economy μ . Given the calibration of other parameters, we have $\Delta \approx 6\%$ per year in this case. Note that because R_t may be larger than \bar{R} , the risk premium $\Delta = \bar{R} - 1$ may be not large enough to push interest rates all the way to zero. Even when they do, the duration of the ZLB episode may be very short-lived if some other positive shocks hit the economy. The alternative calibration is to set a much higher value of $\Delta = 0.120$. This value will ensure that ZLB episodes are almost always long-lived.

3 Results

For each calibration, we simulate the model for 10,000 periods to calculate welfare and various statistics such as the frequency and duration of ZLB episodes. Because Coibion et al. (2012) provide an exhaustive

description of mechanisms and results for the conventionally calibrated model, we focus our analysis on the effects of alternative calibrations of risk premium shocks that govern the properties of ZLB.

3.1 Parameters of Risk Premium Shock and the Properties of ZLB Episodes

We first consider how different parameter values in each representation of the risk premium shock process affect the properties of ZLB episodes. Panel A of Figure 1 illustrates how ρ_q and σ_q^2 in the AR(1) models affect the unconditional frequency of the economy being at ZLB (that is, the fraction of periods when $R_t = 1$). By raising either ρ_q and σ_q^2 , we increase the unconditional frequency of the ZLB. This is intuitive: more persistent shocks (higher ρ_q) naturally tend to leave the economy depressed longer and more volatile shocks (higher σ_q^2) imply that large enough shocks that push the economy into the ZLB happen relatively more frequently. At the same time, there is a clear trade-off between ρ_q and σ_q^2 : one can sustain a given level of the unconditional frequency of ZLB episodes by reducing σ_q^2 (i.e., making the risk premium shocks less volatile) and increasing ρ_q (i.e., making the shocks more persistent) or vice versa. Hence, one can in principle achieve a target frequency of ZLB episodes through different combinations of σ_q^2 and ρ_q .

However, changing the parameter values of σ_q^2 and ρ_q in such a way that the unconditional frequency of ZLB episodes is unchanged still changes the nature of ZLB episodes. When ρ_q is relatively high for a given unconditional frequency of ZLB episodes (and σ_q^2 is therefore relatively low), ZLB episodes will tend to be rare but longer-lived, as suggested by the historical experience. Panel A of Figure 2 demonstrates this result: as ρ_q rises and we move along an isoquant for a given frequency of ZLB episodes (so σ_q^2 falls by the necessary amount), the average duration of ZLB episodes also rises. This suggests that, within the context of AR(1) risk-premium shocks, we can model the notion of rare but long-lived ZLB episodes by raising ρ_q and lowering σ_q^2 , thereby changing the distribution of ZLB episodes from being frequent and short-lived to being rare and long-lived.

However, Panel A of Figure 2 also reveals that the average duration of ZLB episodes is fairly insensitive to changes in ρ_q when these are offset by corresponding changes in σ_q^2 that leave the ZLB frequency unchanged. It takes very large changes in ρ_q to raise the duration of ZLB by a quarter. For example, if we focus on the unconditional probability of 0.1, one has to increase ρ_q from 0.97 to 0.985 (that is, increase the half-life of the risk premium shock from ≈ 23 quarters to ≈ 46 quarters) to raise the average ZLB duration by just one quarter.⁷

To further explore why the average ZLB duration is relatively unresponsive to changes in ρ_q , we examine the distribution of ZLB durations in the AR(1) model for different calibrations of σ_q^2 and ρ_q in Figure 3. In each case, we choose σ_q^2 and ρ_q such that the unconditional frequency of ZLB episodes is 0.10 but the average duration of ZLB episodes varies from a little over two quarters to almost seven quarters in duration. A striking feature common to all calibrations is that the distribution of ZLB episodes has a very heavy left tail: most ZLB episodes are just one- or two-quarters long while the share of ZLB episodes longer than 12 quarters is less than 20%. Similar results have been found in other studies (e.g., Cheng et al. 2012) using an AR(1) process for shocks akin to our risk premium shock. This characteristic of the ZLB distribution is largely invariant to the average duration. As ρ_q increases, there are relatively

⁷ The half life is given by $\ln(0.5) / \ln(\rho_q)$.

more very long-lived episodes. But higher values of ρ_q also require lower values of σ_q^2 , so the share of 1-quarter ZLB episodes falls only gradually. These two nearly off-setting effects explain the pattern noted in Panel A of Figure 2 that even large increases in ρ_q have very modest effects on average ZLB durations.⁸ In short, it is very difficult to generate an empirically realistic pattern of ZLB episodes using AR(1) shocks to the risk premium.

As a result, we also consider an alternative modeling strategy of regime switching risk-premia, as described in section 2.4. There are now three parameters of interest: p_{12} (the probability of a risk-premium increase when the economy is outside the ZLB), T_q (the duration of the high risk premium period), and Δ (the size of the risk premium shock). In Panel B of Figure 1, we illustrate that, for a fixed value of $\Delta = 9\%$, by changing p_{12} and T_q we can maintain a given unconditional frequency of ZLB, which is qualitatively similar to the AR(1) case. Increasing p_{12} means raising the probability the risk premium going up when the economy is outside the ZLB, which is similar to raising σ_q^2 in the AR(1) case. Increasing T_q makes the length of the risk premium shock longer, which is akin to increasing ρ_q in the AR(1) case. Hence, raising either parameter serves to increase the frequency of ZLB episodes and there is a tradeoff between the two parameters that can be utilized to maintain a fixed unconditional frequency of ZLB episodes, as in the AR(1) case. In this respect, the two ways of modeling risk premia appear similar.

However, the regime switching model is much more successful at allowing us to change average durations of ZLB episodes. Panel B of Figure 2 plots, again for a fixed value of $\Delta = 9\%$, how the average duration of ZLB episodes changes as one increases T_q (the length of risk premium shocks) while changing p_{12} by just enough to maintain a fixed unconditional frequency of ZLB episodes (as indicated by isoquants in the Figure). In contrast to the very flat slopes obtained with the AR(1) model, the regime switching model yields an approximately linear increase with a slope just above one in the average duration of ZLB episodes.

The reason for this difference lies in the distribution of ZLB episodes generated by the regime switching model. Figure 4 plots these distributions for four different values of Δ : 6%, 9%, 12%, and 18%. In each case, T_q is held fixed at 12 quarters while p_{12} is chosen to generate an unconditional frequency of ZLB episodes of 0.10. When the size of the risk premium shock is low ($\Delta = 6\%$), the distribution of ZLB episodes is very similar to the AR(1) case. Even though the risk-premium shocks are long-lived, they are not large enough to keep the economy in the ZLB for extended periods because other shocks tend to quickly push the economy out. As a result, ZLB episodes end up being frequent and short-lived, as in the AR(1) case. But as the size of the risk-premium shock goes up, the distribution of ZLB durations shifts away from short durations and toward longer-lived episodes. In part, this increase in the duration of ZLB episodes is generated by eliminating short breaks in periods with low interest rates. For example, a risk premium shock lasting 8 quarters can push the nominal interest rate towards zero but an expansionary demand can interrupt the spell of low interest rates. As a result, the simulated path may have three periods at the ZLB, then one period outside the ZLB, and then another four periods at the ZLB even though these eight periods are effectively the same episode. A sufficiently high Δ ensures that such interruptions are minimized which raises the average duration of ZLB episodes. In contrast, the AR(1) model does not allow for a straightforward treatment of such breaks. Once Δ is large enough, we see almost no short-

⁸ In principle, it is possible to push ρ_q close to one and make ZLB episodes potentially very long. In this case, however, we start to face numerical issues. Once we have very long periods with the Taylor principle being violated, the model generates indeterminacy and thus can break down.

lived ZLB episodes because the size of the risk premium shock is too large to be offset by other economic shocks and the duration of ZLB episodes is generally close to, albeit somewhat less than, the duration of the risk premium shock. Hence, this alternative modeling strategy is much more successful at replicating the empirical pattern of ZLB episodes being both rare and long-lived.

It's also worth noting that as Δ becomes large, the distribution of ZLB episodes becomes increasingly tight around T_q , a feature which may seem unrealistic given that ZLB episodes have been varied in duration across countries and time. This reflects our assumption that T_q is deterministic and constant. One could readily assume a stochastic process for T_q , which would generate much more variation in the distribution of durations of ZLB episodes. Unfortunately, because of the lack of historical data on ZLB episodes, it is not clear a priori how one might best characterize this distribution. As a result, and because our baseline calibration of $\Delta = 9\%$ already seems to yield a reasonable distribution of ZLB episodes, we prefer to treat T_q as a constant.

3.2 ZLB Duration, Welfare, and Optimal Inflation

We now consider how changes in the duration of ZLB episodes affect welfare. To do so, we first illustrate how welfare changes with different levels of steady state inflation under different calibrations of the risk premium process. Parameters for the risk premium are chosen to achieve different average durations of ZLB episodes but a fixed unconditional frequency of the ZLB equal to 0.1 when the steady state level of inflation in the model is equal to 3.5% (the historical average for the U.S.). We then simulate the model for each set of parameter values under different levels of steady state inflation to quantify changes in welfare.

The results for the AR(1) assumption for risk premia are plotted in Panel A of Figure 5 for average ZLB durations ranging from a little over two quarters to almost seven quarters. When the average ZLB duration is very low (about two quarters), welfare losses are very high at all levels of inflation. This is because achieving short durations of ZLB episodes for this fixed frequency requires very volatile risk premium shocks, and this volatility generates a very high level of welfare losses. These losses decline as average durations rise to around five quarters because the latter requires much less volatile shocks to the risk premium.

As ZLB durations get much higher, the welfare losses experienced at low levels of steady state inflation become extremely high, the welfare curves start to shift down, and the peaks of the curves start to move to the right. The first and second observations suggest that the cost of ZLB episodes increases in the duration of ZLB episodes. As a result, it is optimal to trade off some steady-state inflation for a reduced incidence of the ZLB.

To confirm this intuition, we plot the cost of the ZLB per quarter for the same combination of parameters in Panel A of Figure 6. As the duration of ZLB episodes increases, the welfare cost per period of ZLB rises. Furthermore, the increase in the cost is non-linear and rapid. If steady-state inflation is zero, then for the combination of ρ_q and σ_q^2 with the implied average duration of ZLB episodes equal to approximately 7 quarters, the *permanent* consumption-equivalent cost of a quarter at ZLB is a whopping 13%. This cost, however, rapidly declines with the average duration of ZLB episodes. For example, with the same unconditional frequency of binding ZLB but an average duration equal 4 quarters, the cost is around 1.3%. These costs also decline sharply with higher levels of trend inflation, since the latter reduce the duration of ZLB episodes. For example, the same calibration that yields a 13% cost of a quarter at the

ZLB when steady-state inflation is zero yields a much smaller ZLB cost per quarter of just over 2% when steady-state inflation is 3%.

With AR(1) shocks, a small increase in the duration of ZLB from around 5.5 quarters to almost 7 quarters is associated with an increase in the optimal steady-state level of inflation from 1.5% per year to around 2.5% per year. From a policy point of view, this is a dramatic difference in the inflation rate coming from a relatively small change in the average duration of ZLB episodes. This sensitivity of the optimal rate of inflation reflects the fact that while the average duration of the ZLB may be rising only little, engineering this change with AR(1) shocks requires generating some dramatically longer-lived ZLB episodes in the tail of the distribution of ZLB durations to make up for the fact that most episodes remain very short-lived, as illustrated in Figure 3. Because long-lasting episodes are extremely costly in the model, even a very rare occurrence of such episodes translates into a non-trivial unconditional cost of the ZLB. These episodes are extremely costly because the Taylor principle is not satisfied for a long time and thus a large volatility of output, inflation and consumption is possible. Because the cost of the ZLB is convex in ZLB duration, the welfare loss essentially explodes with these very long-lived ZLB episodes. As a result, raising the steady state inflation rate becomes worthwhile to offset these otherwise extremely rare and costly events.

The very high sensitivity of the optimal inflation rate to the average duration of the ZLB therefore appears to be an artefact of the empirically unrealistic distribution of ZLB episodes generated by AR(1) shocks, making it an unreliable guide to policy. We therefore turn to the predictions of the regime switching approach, which can generate more empirically realistic distributions of ZLB episodes. First, the shapes of the welfare curves in the regime-switching model (Panel B of Figure 5) are qualitatively similar to those of the AR(1) model. When average ZLB durations are relatively high, the welfare losses of low trend inflation are particularly large. This again reflects the disproportionately high cost of ZLB episodes when average durations are higher, as illustrated in Panel B of Figure 6. Second, the optimal inflation rate is rising with the average duration of ZLB episodes (once these are sufficiently high) as higher levels of inflation work to reduce the incidence of these episodes that induce such high welfare costs.

However, there are also some important differences between the results for the AR(1) and regime-switching models. One is that the curvature in Panel B is weaker than that in Panel A, especially at higher durations of ZLB episodes. Another is that the peaks of the curves in Panel B are closer to zero than in Panel A, such that welfare is generally higher in the regime-switching model than in the AR(1) model. The latter reflects the fact that ZLB episodes are less costly in the regime-switching model than in the AR(1) model even when we use high values of Δ . Panel B of Figure 6 confirms this conjecture: the costs of the ZLB per quarter of hit are more compressed and flatter in the regime switching model than in the AR(1) model. For example, at $\bar{\pi} = 0$, an average duration of ZLB episodes of 7 quarters yields a welfare cost of 13% in consumption equivalent per quarter of binding ZLB under AR(1) shocks but only around 3.5% with regime switching in the risk premium. This much smaller cost suggests that raising steady-state inflation levels might be less effective at combatting ZLB in the regime-switching model than in the AR(1) model. Indeed, in Panel B of Figure 5, we see that raising the average duration of ZLB episodes by a full year raises the optimal inflation rate by less than a percentage point, a significantly reduced sensitivity relative to the AR(1) case.

Since the cost of the ZLB is lower in the regime-switching model than in the AR(1) model, the implied optimal steady-state rate of inflation rate is also lower in the regime-switching model. For example, when average durations of ZLB episodes are around 5-5.5 quarters, the optimal inflation rate is

1.4% with regime switching risk premia but approximately 1.7% with AR(1) shocks. When average durations are higher, the differences are even more pronounced: the optimal inflation rate with AR(1) shocks is nearly 3% when ZLB episodes have an average duration of 6.8 quarters whereas it is only 1.8% with regime switching in risk premia.

In short, these results highlight the pitfalls associated with relying on AR(1) shocks to study how economies hit the ZLB. Because this approach necessarily implies the existence of many very short-lived ZLB episodes, generating longer *average* durations requires hitting the economy with extremely long-lived and disproportionately costly episodes that drive welfare and policy results. In contrast, the regime switching approach can deliver a more realistic distribution of ZLB episodes and this distribution implies a smaller sensitivity of the optimal inflation rate to the average duration of ZLB episodes.

3.3 Optimal Inflation Rates for Different Durations and Frequencies of the Zero Bound

In Figure 5, we provided some results on optimal inflation rates for a few average durations of ZLB episodes and a single unconditional frequency of ZLB episodes. But as discussed earlier, the paucity of historical experience with this type of episode should make anyone wary of taking a strong stand on the precise values of these parameters. As a result, we now consider a much wider range of both frequencies and durations of ZLB episodes and characterize optimal inflation rates in each case.

Our results for AR(1) shocks are presented in Panel A of Figure 7 while analogous results for regime switching model are in Panel B of Figure 7. In each case, we plot optimal inflation rates (vertical axis) associated with different average durations of ZLB episodes (horizontal axis) and unconditional frequencies of the ZLB (captured by isoquants), where the latter two are measured at a 3.5% steady state inflation rate. The key result in the case of AR(1) shocks is, regardless of the specific frequency of the ZLB, the optimal inflation rate rises extremely rapidly with the average duration, as indicated by the slope of the isoquants. For example, going from an average duration of the ZLB of five quarters at an unconditional frequency of ZLB episodes equal to 7% to an average duration of eight quarters raises the optimal inflation rate from about 2% to almost 4.5%. But as discussed earlier, this excessive sensitivity reflects the unrealistic distribution of ZLB episodes generated by AR(1) shocks to risk premia.

A second unappealing feature of the AR(1) approach to modeling shocks is the fact that low frequency isoquants are to the left of higher frequency isoclines. This implies that for a given average duration of ZLB episodes, a higher frequency of the ZLB is associated with a *lower* optimal rate of inflation. The reason is that we cannot separately calibrate the volatility of the risk premium and the frequency and duration of ZLB episodes with only two parameters for the shock process, which is yet another undesirable property of AR(1) shocks.

Panel B of Figure 7 presents the analogous results from the regime switching approach to modeling shocks that push the economy into the zero bound on interest rates. The first difference to note is that, as expected, the slopes of the isoquants are now much flatter: optimal inflation rates rise less rapidly with average ZLB durations. This reflects the fact that one does not need to introduce extremely long-lived ZLB periods to change the average duration as is the case with AR(1) shocks. Nonetheless, high inflation rates can be sustained as optimal if one believes that average durations of ZLB episodes are sufficiently high or sufficiently frequent.

A second difference to note is that the regime switching approach now implies that, holding the average duration constant, higher frequencies of the ZLB would be associated with higher optimal rates of inflation. This result holds even at lower levels of Δ , as illustrated in Appendix Figure 1. Hence, the

alternative modeling strategy of regime switching shocks can fix this additional undesirable property of AR(1) shocks when it comes to characterizing the tradeoffs faced by policymakers.

More broadly, the results in Panel B of Figure 7 suggest that a wide range of optimal inflation rates can potentially be defended, depending on what one perceives to be the correct values for the average duration and frequency of ZLB episodes are. For example, relying only on the U.S. post-war experience of a single ZLB episode lasting seven years over a seventy year period points to an unconditional frequency of 10% and an average duration of twenty-eight quarters. This duration would justify an optimal inflation rate of approximately five percent, well above the Federal Reserve's current objective of two percent but in line with recommendations made by economists like Olivier Blanchard and Paul Krugman. Of course, other countries such as Canada experienced much shorter ZLB episodes during the Great Recession and other advanced economies such as Australia and New Zealand did not reach the ZLB at all. This suggests that the U.S. experience is likely not an average experience.

To get a better sense of the cross-country experience with the zero bound on interest rates since World War II, we summarize the experience of a range of advanced economies since 1950 in Table 2. Only Switzerland has experienced one-quarter long ZLB experiences, in 1972Q1 and again in 1972Q3. All other durations have been of at least one year. There are several episodes of approximately one year in length, although some of these are what one might consider interrupted sequences of longer underlying periods of economic weakness, such as the Euro-zone and Sweden from 2009Q3-2010Q3. For countries still at the ZLB, we assume that they will remain at the ZLB until 2015Q4 and measure durations and frequencies using this final date.

Figure 8 plots the distribution of the duration of ZLB episodes from Table 2 as well as a kernel estimate of the density. Despite the small number of observations, we can see that this distribution resembles, at least loosely, the distribution of ZLB episodes of the regime switching model with $\Delta = 9\%$. There is a small but non-zero share of very short ZLB episodes, but the vast majority of ZLB episodes are longer-lived, with significant dispersion in terms of actual durations of ZLB episodes. We interpret this as suggesting that our calibration of the model is indeed well suited to characterize the empirical distribution of ZLB episodes and also that the U.S. experience with the ZLB is in the upper ranges, at least in terms of durations. The mean duration across all episodes, for comparison, is fourteen quarters, less than half what the U.S. will have spent at the ZLB since the start of the Great Recession. If Japan is excluded from the sample, or if we include all of the Euro-zone countries as separate observations, the mean durations are lower still, at 11.5 and 12.3 quarters respectively. Hence, we interpret a reasonable estimate for average durations at the ZLB as somewhere between ten and fifteen quarters.

For each country in this sample, we also estimate the frequency of the ZLB as the share of quarters spent at the ZLB between 1950Q1 and 2015Q4. These frequencies range from a high of 23% for Japan to a low of 0% for Norway, Australia, and New Zealand, the three countries who experienced no ZLB periods at all over this time period. The U.S. frequency, at 11%, is again well above the average frequency across all countries in the sample, which is given by 7.5%. Excluding Japan lowers this frequency further to just under 6%, although it can go as high as 11% if we exclude those countries which have never experienced a ZLB episode from our estimate. Again, the U.S. is on the high end of estimates, and the cross-country experience suggests a plausible range of values for the frequency of ZLB estimates ranging anywhere from 6% to 11%.⁹

⁹ Dropping the high-inflation period (e.g. 1968-1982), when reaching the ZLB was less likely given higher nominal interest rates, reduces the length of the sample and therefore raises measured frequencies somewhat. Monetary

With these ranges of values for the frequency and duration of ZLB episodes in mind, we can then reexamine optimal inflation rates in Panel B of Figure 7. If we consider the very low end of our estimates for both frequency and duration of ZLB episodes (a duration of around 10 quarters and a frequency of 6%), then our model points toward an optimal level of the annual inflation rate of 1.5%. The high end of our estimates for both frequency and duration of ZLB episodes (a duration of around 15 quarters and a frequency of 11%) points instead toward an optimal rate of inflation of 3%. Hence, these results suggest that reasonable bounds for the optimal rate of inflation are 1.5-3%.

The results vary little if we consider alternative calibrations of our regime switching model. In Appendix Figure 1, we plot equivalent results as those in Figure 7 using $\Delta = 6\%$ and $\Delta = 12\%$. With the former, ZLB episodes are less costly, so optimal inflation rates are somewhat lower. But the equivalent range is only slightly changed, to 1.5-2.8%. With a higher value of Δ , ZLB episodes become more costly and optimal inflation rates rise. In this case, the range of optimal inflation rates becomes 1.6-4.0%. But as illustrated in Figure 4, neither calibration is as successful in replicating the historical distribution of ZLB durations as our baseline calibration.

4 Conclusion

Economies rarely hit the zero bound on interest rates, but when they do, these episodes tend to be long-lived. This simple empirical pattern is one that is not replicated by traditional models of optimal inflation that incorporate the zero bound on interest rates. We show that the most common approach to modeling shocks that drive the economy into the zero bound yields a distribution of ZLB episodes that is counterfactual: ZLB episodes are frequent and short-lived rather than rare and long-lived. And this counterfactual distribution is not innocuous. The fact that AR(1) shocks have to generate a large share of very short-lived ZLB episodes implies that longer average durations of ZLB episodes can only occur if there are also extremely long-lived episodes that generate disproportionately large welfare costs. The latter make policymakers very willing to raise inflation rates to avoid the ZLB, making the optimal inflation rate exceedingly sensitive to the average duration rate of ZLB episodes.

In contrast, we show that a regime switching approach to modeling the shocks that push the economy into the ZLB can generate an empirically realistic distribution of ZLB episodes and that this approach does not generate the same sensitivity of the optimal inflation rate to the average duration of ZLB episodes. The optimal rate of inflation is still sensitive to the ZLB since the latter is costly, and more frequent or long-lived episodes increase the incentives of policymakers to raise the target rate of inflation, but this incentive is much reduced relative to what is implied by the standard approach to modeling shocks.

The specific optimal rate of inflation implied by the model remains very sensitive to one's beliefs about the frequency and duration of ZLB episodes, values for which history provides only limited guidance. Using only the U.S. post-WWII experience, for example, our model would imply an optimal rate of inflation above 4%. From a broader cross-section of countries' historical experiences with the ZLB, one could just as readily conclude that the optimal rate of inflation is 2%, with a plausible range of values running from 1.5% to at least 3%. Given the uncertainty associated with measuring historical frequencies and durations associated with ZLB episodes, the range of plausible outcomes for optimal inflation rates implies that profound humility is called for by anyone advocating a specific inflation target.

policy prior to this period appears to have been made similarly to that of post-Volcker era (Romer and Romer 2002), so there is little reason to disregard the earlier time period.

References

- Abo-Zaid, Salem, 2015. "Optimal long-run inflation with occasionally binding financial constraints," *European Economic Review* 75, 18-42.
- Amano, Robert, and Malik Shukayev, 2012. "Risk Premium Shocks and the Zero Bound on Nominal Interest Rates," *Journal of Money, Credit and Banking* 44(8), 1475-1505.
- Ascari, Guido, Louis Phaneuf, and Eric Sims, 2015. "On the Welfare and Cyclical Implications of Moderate Trend Inflation," NBER Working Paper No. 21392.
- Ascari, Guido, and Argia M. Sbordone. 2014. "The Macroeconomics of Trend Inflation." *Journal of Economic Literature*, 52(3): 679-739.
- Benigno, Pierpaolo, and Michael Woodford, 2005. "Inflation Stabilization And Welfare: The Case Of A Distorted Steady State," *Journal of the European Economic Association* 3(6), 1185-1236.
- Billi, Roberto M., 2011. "Optimal Inflation for the US Economy," *American Economic Journal: Macroeconomics* 3(3), 29-52.
- Blanco, Andres, 2015. "Optimal Inflation Target in an Economy with Menu Costs and an Occasionally Binding Zero Lower Bound," manuscript.
- Calvo, Guillermo, 1983. "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics* 12, 983-998.
- Christiano, Lawrence J., Martin Eichenbaum, and Sergio Rebelo, 2011. "When is the Government Spending Multiplier Large?" *Journal Of Political Economy* 119(1), 78-121.
- Chung, Hess, Jean-Philippe Laforte, David Reifschneider, and John C. Williams, 2012. "Have We Underestimated the Likelihood and Severity of Zero Lower Bound Events?" *Journal of Money, Credit and Banking* 44, 47-82.
- Dotsey, Michael, Robert G. King, and Alexander L. Wolman, 1999. "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *Quarterly Journal of Economics* 114, 655-690.
- Eggertsson, Gauti B. and Michael Woodford, 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity* 34(1), 139-235.
- Golosov, Mikhail, and Robert E. Lucas Jr., 2007. "Menu Costs and Phillips Curves," *Journal of Political Economy* 115, 171-199.
- Guerrieri, Veronica and Guido Lorenzoni, 2009. "Liquidity and Trading Dynamics," *Econometrica* 77(6), 1751-1790.
- Klaus, Adam, and Roberto Billi, 2007. "Discretionary monetary policy and the zero lower bound on nominal interest rates," *Journal of Monetary Economics* 54(3), 728-752.
- Reifschneider, David, and John C. Williams, 2000. "Three Lessons for Monetary Policy in a Low-Inflation Era," *Journal of Money, Credit and Banking* 32(4), 936-966.

Romer, Christina D. and David H. Romer, 2002. "A Rehabilitation of Monetary Policy in the 1950s." *American Economic Review* 92(2), 121-127.

Schmitt-Grohé, Stephanie, and Martin Uribe. 2010. "The Optimal Rate of Inflation," Handbook of Monetary Economics, in: Benjamin M. Friedman & Michael Woodford (ed.), *Handbook of Monetary Economics*, edition 1, volume 3, chapter 13, pages 653-722 Elsevier.

Schmitt-Grohé, Stephanie, and Martín Uribe, 2012. "Foreign Demand for Domestic Currency and the Optimal Rate of Inflation," *Journal of Money, Credit and Banking* 44(6), 1207-1224.

Summers, Lawrence, 1991. "Price Stability: How Should Long-Term Monetary Policy Be Determined?" *Journal of Money, Credit and Banking* 23(3), 625-631.

Taylor, John, 1977. "Staggered Wage Setting in a Macro Model," *American Economic Review* 69(2), 108-113.

Weber, Henning, 2012. "The Optimal Inflation Rate and Firm-Level Productivity Growth," Kiel Working Papers 1773, Kiel Institute for the World Economy.

Williams, John, 2009. "Heeding Daedalus: Optimal Inflation and the Zero Lower Bound," *Brookings Papers on Economic Activity*, Fall.

Wolman, Alexander L., 2011. "The Optimal Rate of Inflation with Trending Relative Prices," *Journal of Money, Credit and Banking* 43, 355-384.

TABLE 1: BASELINE PARAMETER VALUES

<u>Parameters of Utility Function</u>		<u>Steady-State Values</u>	
η : Frisch Labor Elasticity	1.00	μ : Growth Rate of RGDP/cap	1.5% p.a.
β : Discount factor	0.998	\bar{c}_y : Consumption Share of GDP	0.80
h : Habit in consumption	0.7	\bar{g}_y : Government Share of GDP	0.20
<u>Pricing Parameters</u>		<u>Shock Persistence</u>	
θ : Elasticity of substitution	7	ρ_g : Government Spending Shocks	0.97
λ : Degree of Price Stickiness	0.55	ρ_m : Cost-Push Shocks	0.90
ω : Price indexation	0.00		
<u>Taylor Rule Parameters</u>		<u>Shock Volatility</u>	
ϕ_π : Long run response to inflation	2.50	σ_g : Government Spending Shocks	0.0052
ϕ_{gy} : Long run response to output growth	1.50	σ_m : Cost-Push Shocks	0.0014
ϕ_x : Long run response to output gap	0.11	σ_a : Technology Shocks	0.0090
ρ_1 : Interest smoothing	1.05	σ_r : Monetary Policy Shocks	0.0024
ρ_2 : Interest smoothing	-0.13		

Notes: The table presents the baseline parameter values assigned to the model in section 2.1 and used for solving for the optimal inflation rate in section 2.2. “p.a.” means per annum.

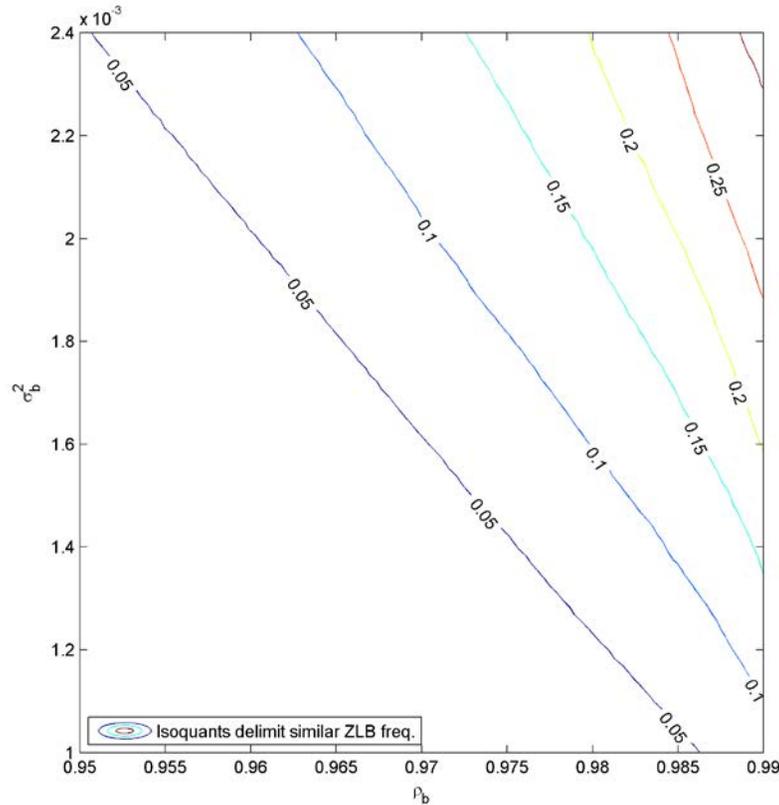
TABLE 2: POST-WAR EXPERIENCES WITH THE ZLB

	Start of ZLB Episode	End of ZLB Episode	Duration (quarters)	Duration (years)	Unconditional Frequency of ZLB for Country
Australia	N/A	N/A	N/A	N/A	0.00
Canada	2009Q2	2010Q2	5	1.25	0.02
Germany/ECB	2009Q3	2010Q4	6	1.50	0.08
Germany/ECB	2012Q1	2015Q4*	16	4	0.08
Japan	1998Q4	2006Q3	32	8	0.23
Japan	2008Q4	2015Q4*	29	7.25	0.23
New Zealand	N/A	N/A	N/A	N/A	0.00
Norway	N/A	N/A	N/A	N/A	0.00
Sweden	2009Q3	2010Q3	5	1.25	0.04
Sweden	2014Q3	2015Q4*	6	1.50	0.04
Switzerland	1972Q1	1972Q1	1	0.25	0.16
Switzerland	1972Q3	1972Q3	1	0.25	0.16
Switzerland	1978Q1	1979Q1	5	1.25	0.16
Switzerland	2003Q1	2004Q3	7	1.75	0.16
Switzerland	2008Q4	2015Q4*	29	7.25	0.16
United States	2008Q4	2015Q4*	29	7.25	0.11
United Kingdom	2009Q1	2015Q4*	28	7	0.11
Average:			14.2	3.6	0.075
Average with all Euro countries:			12.3	3.1	0.085
Average w/o Japan:			11.5	2.9	0.058
Average w/o Norway, Australia and New Zealand:			14.2	3.6	0.108
Average without the 1968-1984 period			17.5	4.4	0.098

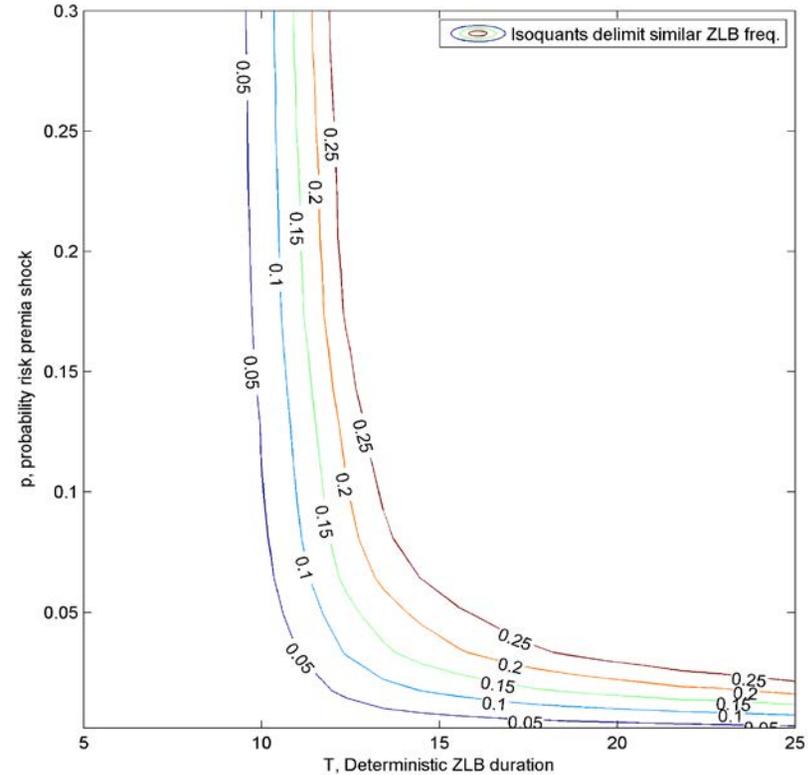
Notes: The table presents a summary of ZLB episodes for advanced economies since World War II. We use Germany as representative of Euro-zone economies. For countries currently at the ZLB and expected to remain so through the end of 2015, we list the end of the episode as 2015Q4. Frequencies are measured using data starting in 1950Q1 going until 2015Q4. “Average with all Euro countries” incorporates eleven additional countries with the same ZLB experience as Germany.

FIGURE 1. UNCONDITIONAL FREQUENCY OF ZLB EPISODES

Panel A: AR(1) model



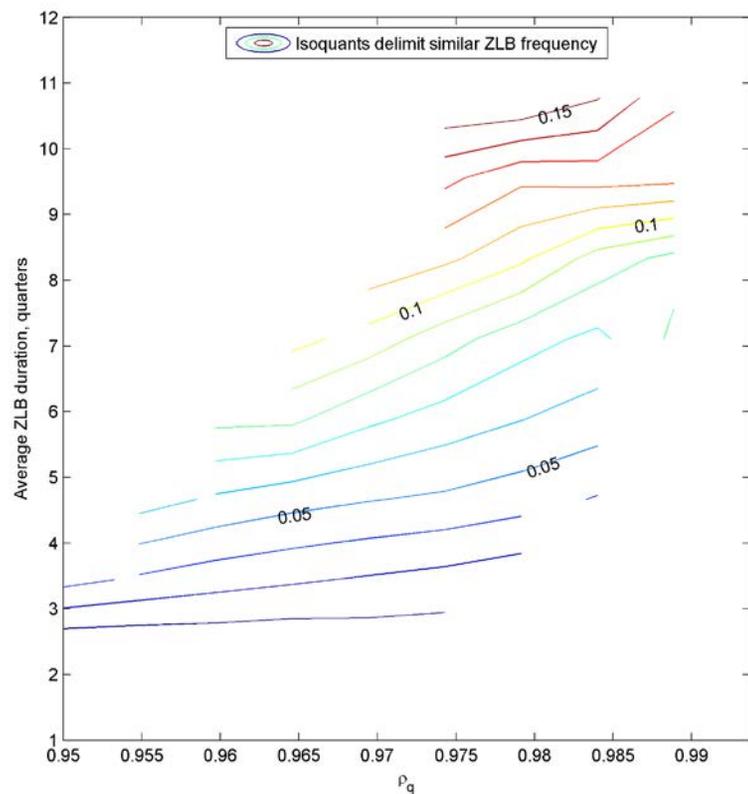
Panel B: Regime-switching model



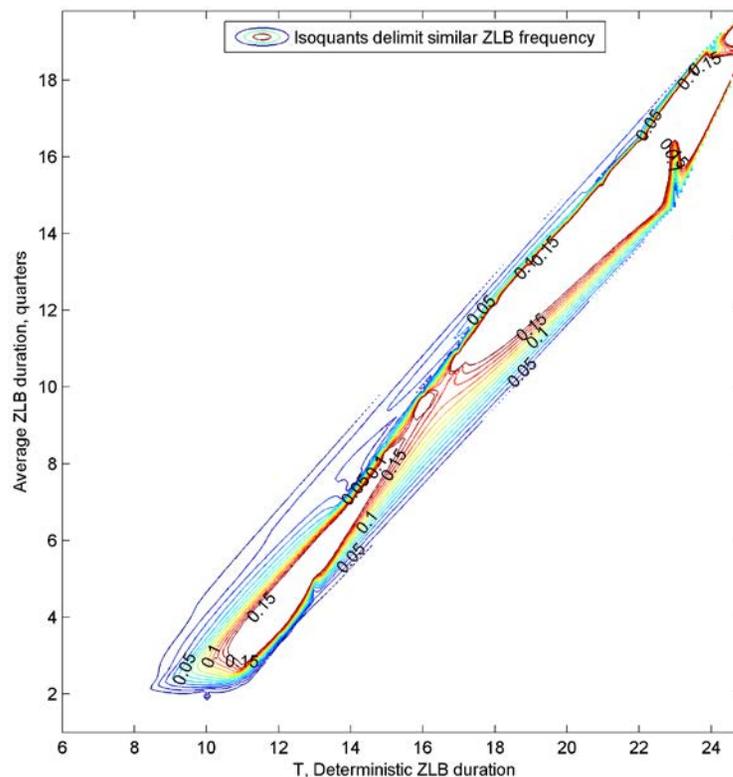
Notes: Each figure plots combinations of parameters that yield specific frequencies of ZLB episodes in simulated data, as indicated by isoquants. In the left panel, the two parameters are the persistence of the AR(1) process for the risk premium (x-axis) and its volatility (y-axis). In the right panel, the two parameters are the duration of the risk premium shock (x-axis) and the probability that a risk premium shock will occur in periods when the economy is not in the ZLB (y-axis). The size of the risk premium shock in Panel B is 9%. See section 3.1 for details.

FIGURE 2. DURATION OF ZLB EPISODES

Panel A: AR(1) model

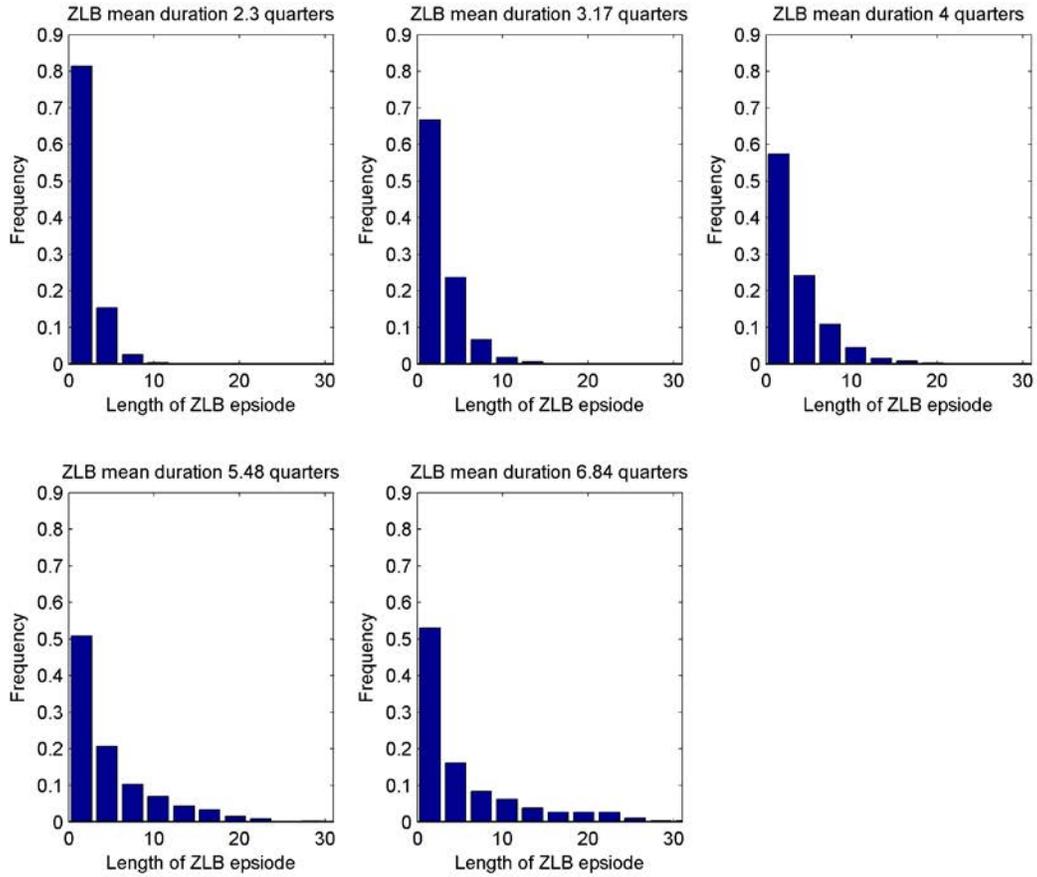


Panel B: Regime-switching model



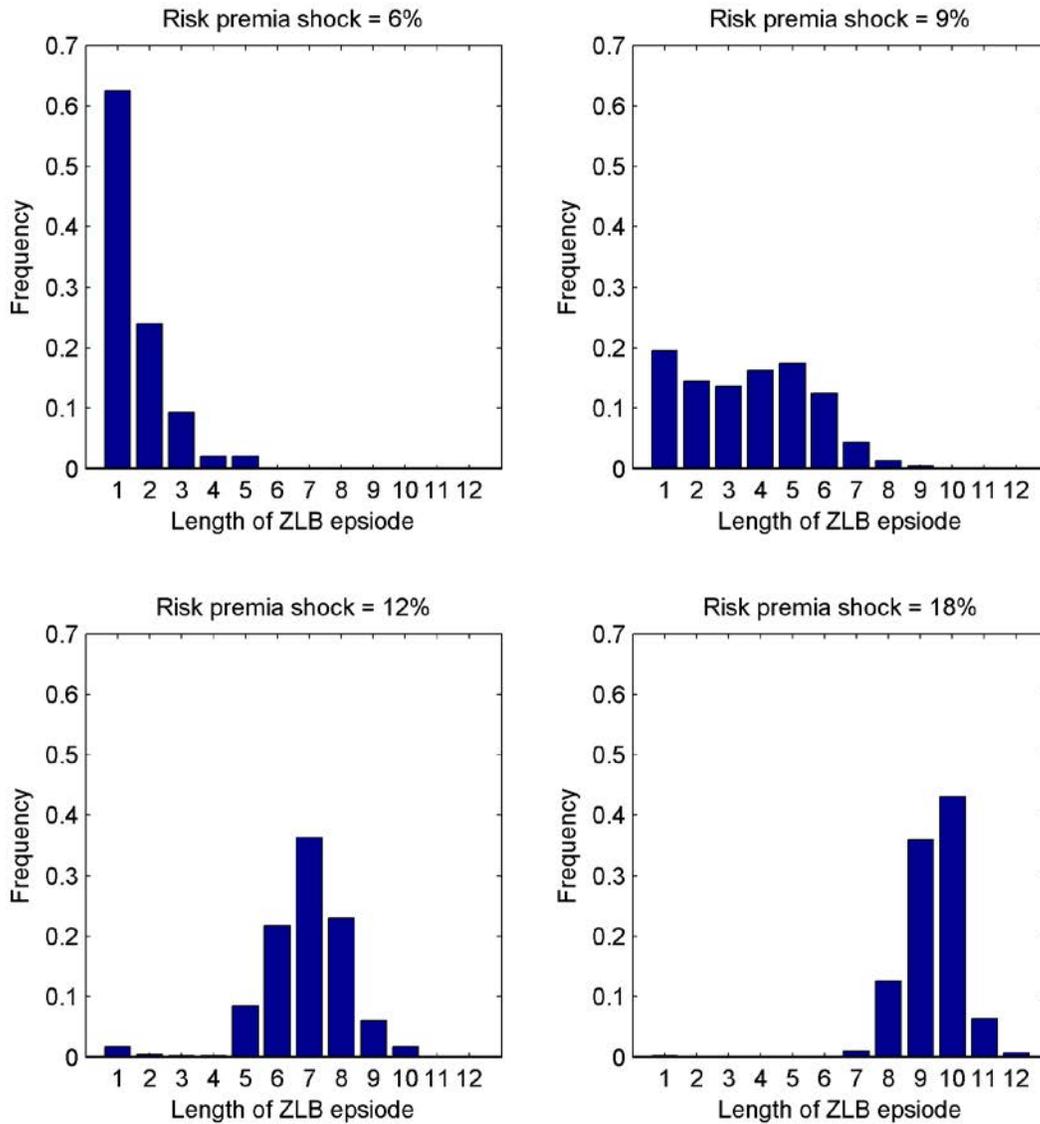
Notes: Each figure plots how average durations of ZLB episodes change as persistence of risk premium shocks varies holding constant the unconditional frequency of ZLB episodes. In panel A, shocks are AR(1), so persistence depends on ρ_q , with σ_q^2 being changed in offsetting way to hold frequency of ZLB constant along isoquants. In panel B, shocks follow regime-switching, so T_q determines the duration of the risk premium shock while p_{12} is changed in offsetting manner to hold frequency of ZLB constant along isoquants. The size of the risk premium shock in Panel B is 9%. See section 3.1 for details.

FIGURE 3. DURATION OF ZLB EPISODES WITH AR(1) SHOCKS



Notes: The figure plots the distribution of durations of ZLB episodes from simulating the model with AR(1) shocks for different average durations of ZLB episodes but a constant unconditional frequency of ZLB episodes. This is done by changing the persistence and volatility of the shock process. See section 3.1 for details.

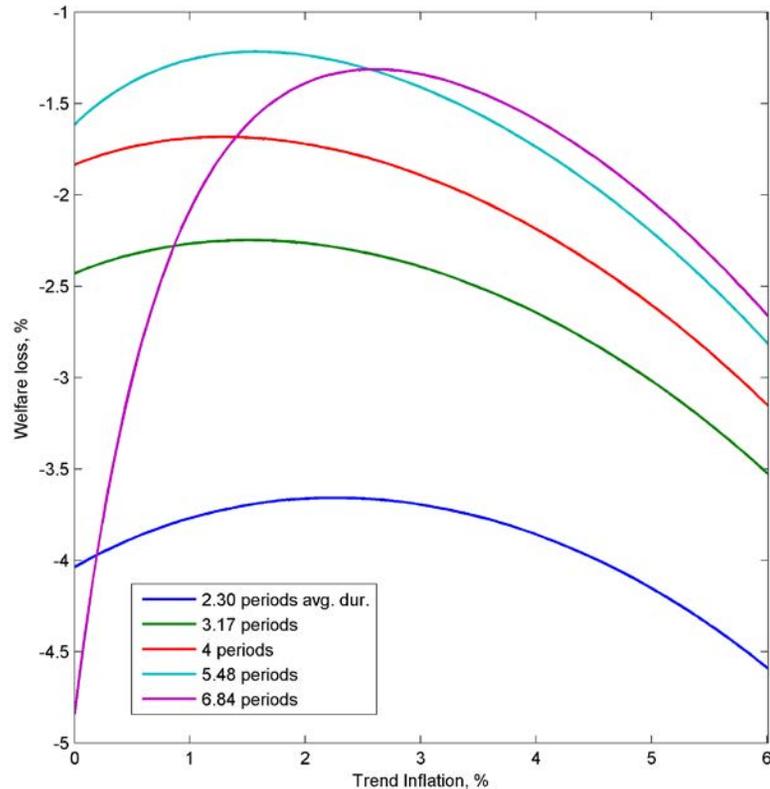
FIGURE 4. DURATION OF ZLB EPISODES WITH REGIME-SWITCHING SHOCKS



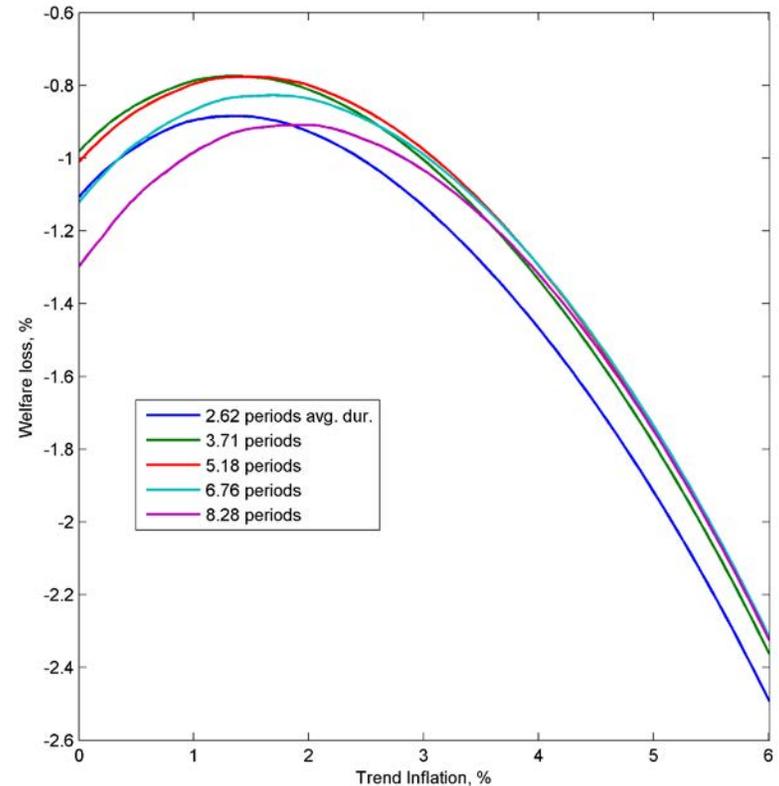
Notes: The figure plots the distribution of durations of ZLB episodes from simulating the model with a regime switching risk premium shock for different average durations of ZLB episodes but a constant unconditional frequency of ZLB episodes. The duration of the risk premium shock (T_q) is held constant but the size of the risk premium shock (Δ) and the probability of a high risk premium occurring outside the ZLB (p_{12}) are allowed to vary to achieve the changing average duration of ZLB episodes with fixed unconditional frequency. See section 3.1 for details.

FIGURE 5. WELFARE LOSSES AT FIXED FREQUENCY OF ZLB

Panel A: AR(1) model



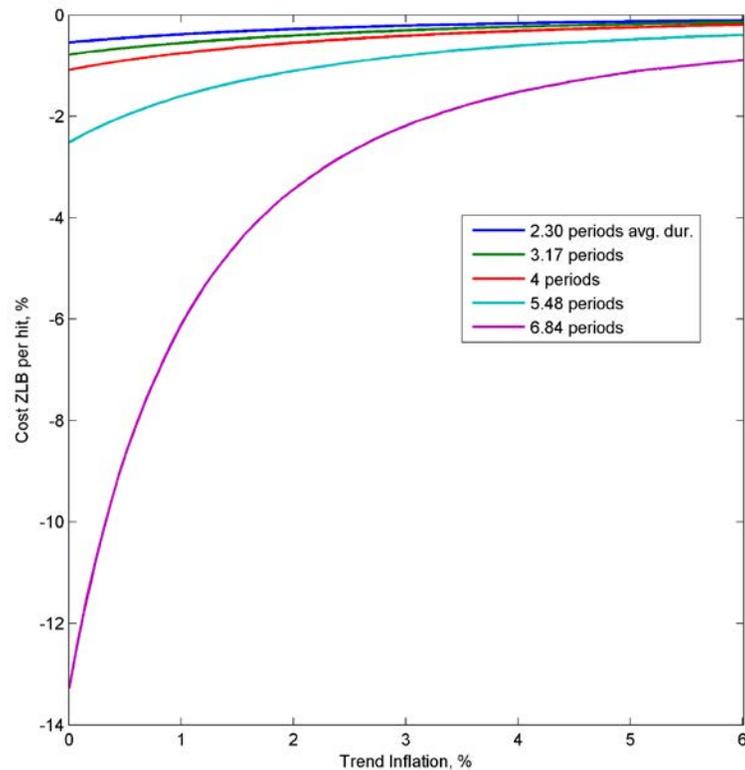
Panel B: Regime-switching model



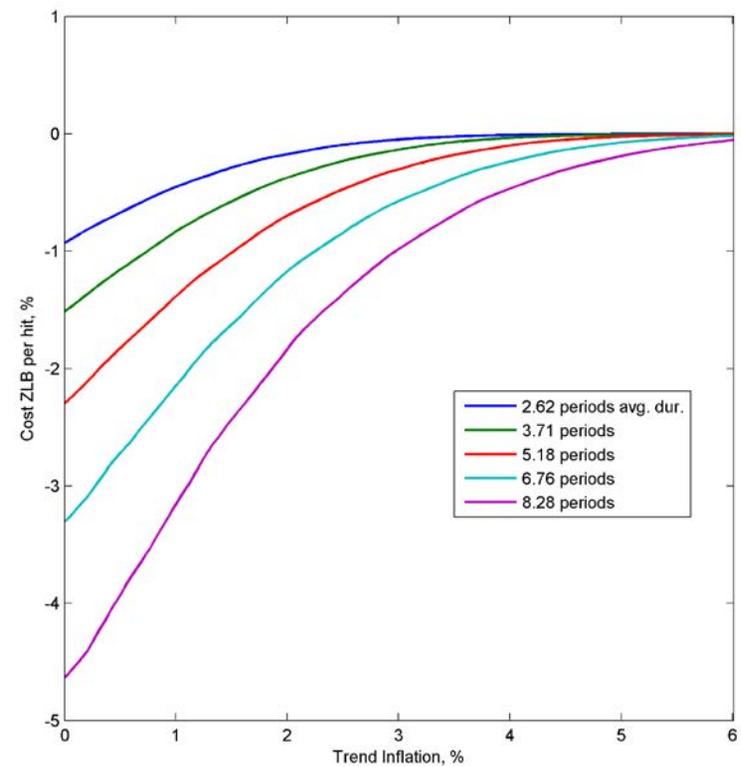
Notes: The figures plots welfare losses (y -axis) associated with different levels of annual trend inflation (x -axis) for different calibrations of the average duration of ZLB episodes (different colored lines) holding the unconditional frequency of ZLB episodes fixed at 0.10 for a trend inflation level of 3.5%. Panel A is done for AR(1) shocks, in which case ρ_q and σ_q^2 are varying to change the ZLB durations. Panel B is done for regime switching risk premia, with $\Delta=9\%$ while T_q and p_{12} are varied to change the average durations of ZLB episodes. See section 3.2 for details.

FIGURE 6. COST OF ZLB PER HIT

Panel A: AR(1) model



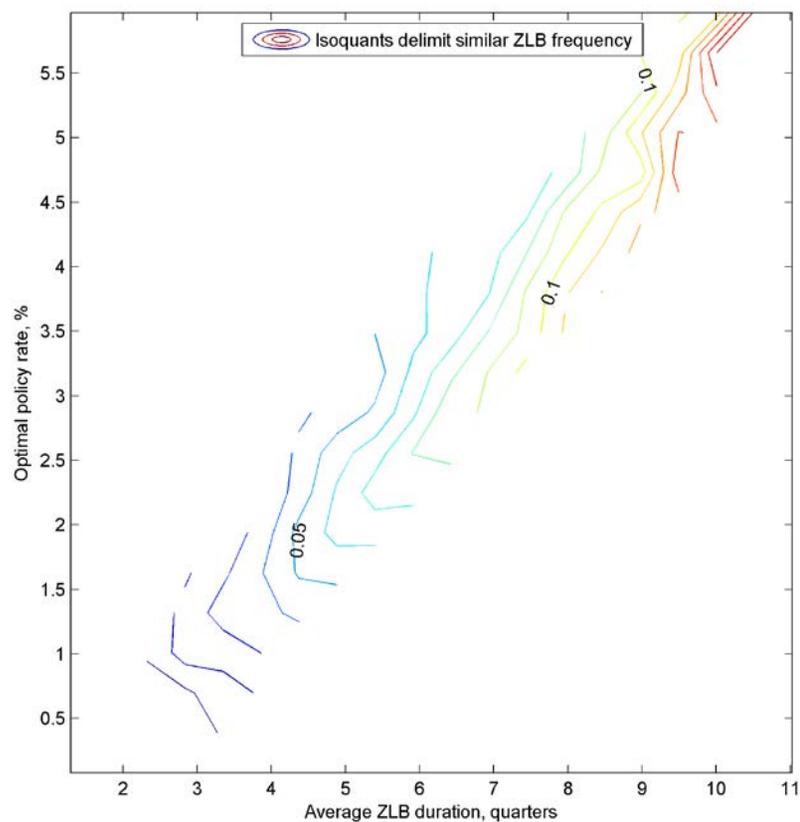
Panel B: Regime-switching model



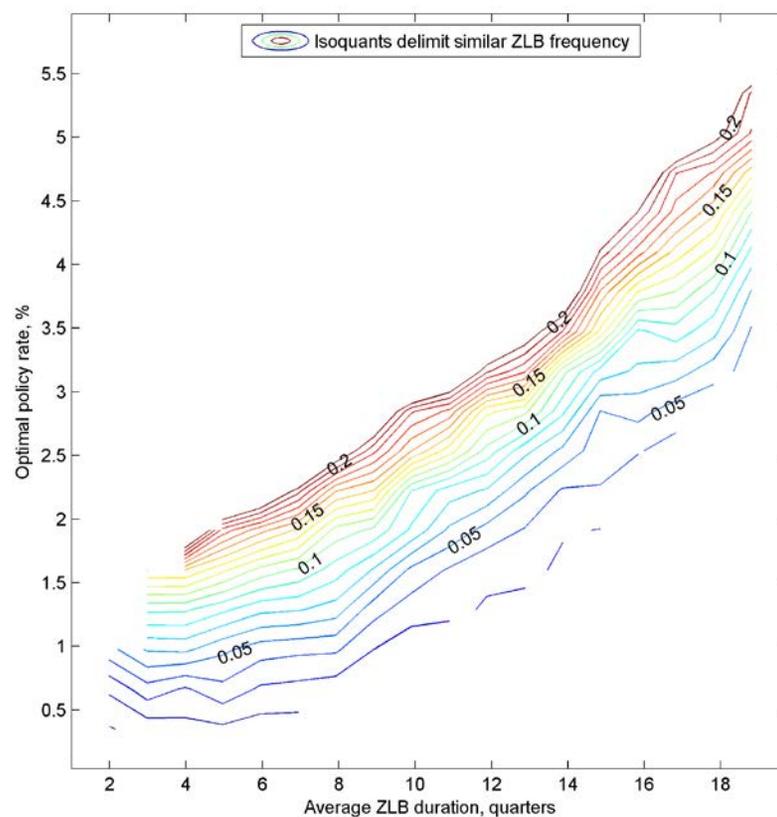
Notes: The figures plots the cost of each ZLB period (y-axis) associated with different levels of annual trend inflation (x-axis) for different calibrations of the average duration of ZLB episodes (different colored lines) holding the unconditional frequency of ZLB episodes fixed at 0.10 for a trend inflation level of 3.5%. Panel A is done for AR(1) shocks, in which case ρ_q and σ_q^2 are varying to change the ZLB durations. Panel B is done for regime switching risk premia, with $\Delta=9\%$ while T_q and p_{12} are varied to change the average durations of ZLB episodes. See section 3.2 for details.

FIGURE 7. OPTIMAL INFLATION FOR DIFFERENT FREQUENCIES AND DURATIONS OF ZLB EPISODES

Panel A: AR(1) model

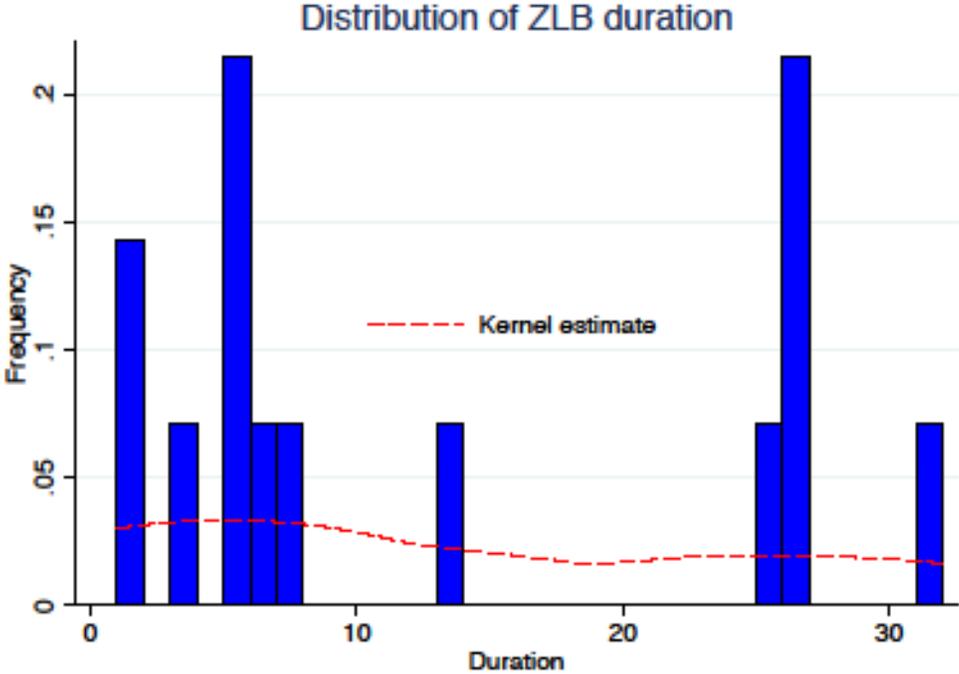


Panel B: Regime-switching model



Notes: The figures plots the optimal annualized inflation rate (y -axis) associated with different levels of average ZLB durations (x -axis) and unconditional frequencies of the ZLB (indicated by isoquants). Panel A is done for AR(1) shocks, in which case ρ_q and σ_q^2 are varying to change the ZLB durations. Panel B is done for regime switching risk premia, with $\Delta=9\%$ while T_q and p_{12} are varied to change the average durations of ZLB episodes. See section 3.2 for details.

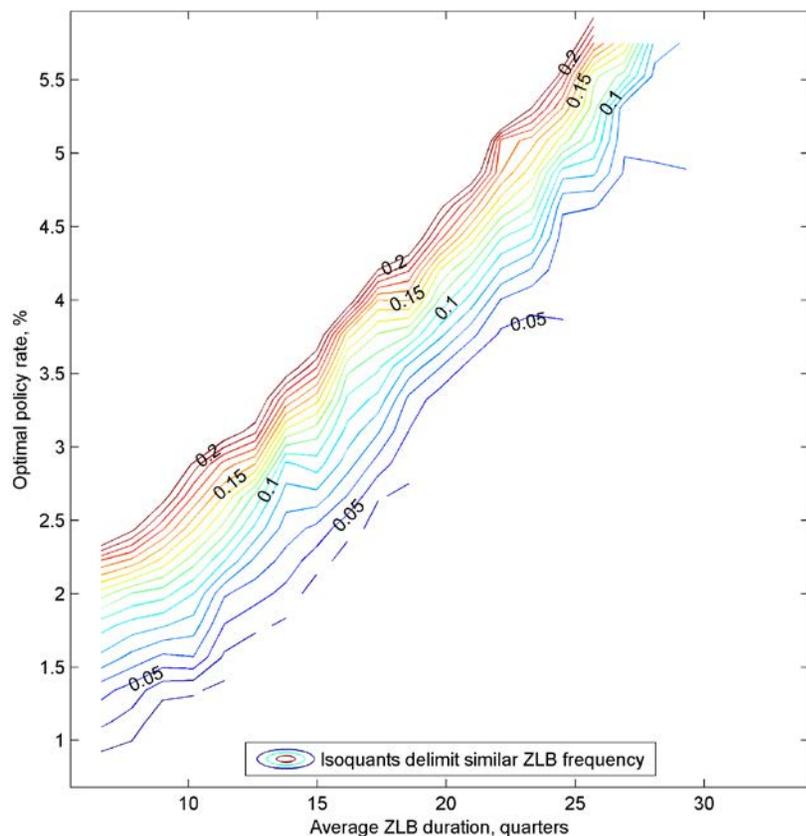
FIGURE 8. DISTRIBUTION OF HISTORICAL ZLB DURATIONS



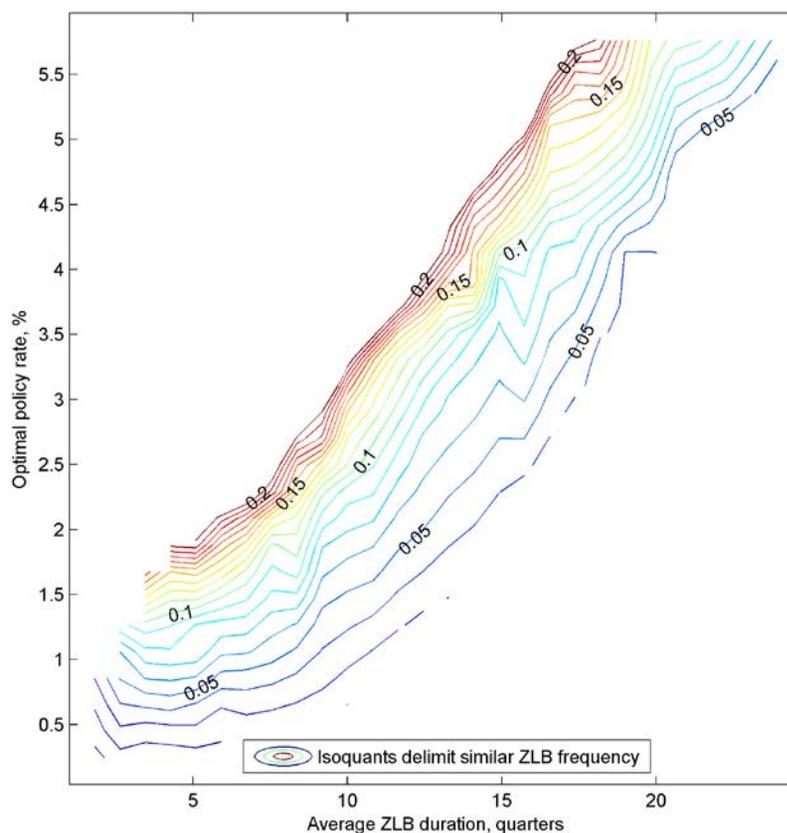
Notes: The figure plots the distribution of durations of historical ZLB episodes from Table 2. Durations are in quarters.

APPENDIX FIGURE 1: OPTIMAL INFLATION WITH DIFFERENT SIZES OF SHOCKS TO RISK PREMIUM

Panel A: $\Delta = 6\%$



Panel B: $\Delta = 12\%$



Notes: The figures plot the optimal annualized inflation rate (y-axis) associated with different levels of average ZLB durations (x-axis) and unconditional frequencies of the ZLB (indicated by isoquants). Panel A is done for regime switching risk premia, with $\Delta = 6\%$ while T_q and p_{12} are varied to change the average durations of ZLB episodes. Panel B is done with $\Delta = 12\%$. See section 3.2 for details.

APPENDIX 1: MODEL

In this appendix, we present a more detailed version of the model from section 2 in the paper.

Household

The representative consumer maximizes the present discounted value of the utility stream from consumption and leisure

$$\max E_t \sum_{j=0}^{\infty} \beta^j \left\{ \log(C_{t+j} - hGA_{t+j}C_{t+j-1}) - \frac{\eta}{\eta+1} \int_0^1 N_{t+j}(i)^{1+1/\eta} di \right\} \quad (1)$$

where C is consumption of the final good, $N(i)$ is labor supplied to individual industry i , GA is the gross growth rate of technology, η is the Frisch labor supply elasticity, h the internal habit parameter and β is the discount factor.¹⁰ The budget constraint each period t is given by

$$\xi_t: C_t + \frac{S_t}{P_t} + T_t \leq \int_0^1 \left(\frac{N_t(i)W_t(i)}{P_t} \right) di + \frac{S_{t-1}q_{t-1}R_{t-1}}{P_t} + \Gamma_t \quad (2)$$

where S is the stock of one-period bonds held by the consumer, R is the gross nominal interest rate, P is the price of the final good, $W(i)$ is the nominal wage earned from labor in industry i , T is real lump sum taxation (or transfers), Γ are real profits from ownership of firms, q is a risk premium shock, and ξ is the shadow value of wealth.¹¹

The risk premium shock q is defined as follows:

$$q_t = \exp(u_t^q)$$

$$u_t^q = \rho_q u_{t-1}^q + \varepsilon_{t-1}^q$$

with ε_{t-1}^q iid normally distributed.

The first order conditions from this utility-maximization problem are then:

$$(C_t - hGA_t C_{t-1})^{-1} - \beta h E_t GA_{t+1} (C_{t+1} - hGA_{t+1} C_t)^{-1} = \xi_t, \quad (3)$$

$$N_t(i)^{1/\eta} = \xi_t W_t(i) / P_t, \quad (4)$$

$$\xi_t / P_t = \beta E_t [\xi_{t+1} q_t R_t / P_{t+1}]. \quad (5)$$

¹⁰ We use internal habits rather than external habits because they more closely match the (lack of) persistence in consumption growth in the data. The gross growth rate of technology enters the habit term to simplify derivations.

¹¹ As discussed in Smets and Wouters (2007), a positive shock to q , which is the wedge between the interest rate controlled by the central bank and the return on assets held by the households, increases the required return on assets and reduces current consumption. The shock q has similar effects as net-worth shocks in models with financial accelerators. Such financial shocks have arguably played a major role in causing the zero lower bound to bind in practice. Amano and Shukayev (2010) also document that shocks like q are essential for generating a binding zero lower bound in the New Keynesian model.

Final Goods

Production of the final good is done by a perfectly competitive sector which combines a continuum of intermediate goods into a final good per the following aggregator

$$Y_t = \left[\int_0^1 Y_t(i)^{(\theta-1)/\theta} di \right]^{\theta/(\theta-1)} \quad (6)$$

where Y is the final good and $Y(i)$ is intermediate good i , while θ denotes the elasticity of substitution across intermediate goods, yielding the following demand curve for goods of intermediate sector i

$$Y_t(i) = Y_t (P_t(i)/P_t)^{-\theta} \quad (7)$$

and the following expression for the aggregate price level

$$P_t = \left[\int_0^1 P_t(i)^{(1-\theta)} di \right]^{1/(1-\theta)}. \quad (8)$$

Government

We allow for government consumption of final goods (G). Government budget constraint is defined as

$$T_t + S_t/P_t = G_t + \frac{S_{t-1}q_{t-1}R_{t-1}}{P_t} + \nu \int_0^1 \left(\frac{N_t(i)W_t(i)}{P_t} \right) di \quad (9)$$

$$G_t = \exp(u_t^G) \bar{G}_t \quad (10)$$

$$u_t^G = \rho_G u_{t-1}^G + \varepsilon_t^G$$

with ε_{t-1}^G iid normally distributed

\bar{G}_t is the path of government spending such that the share of government spending in the economy is fixed when prices are flexible

Substituting household's budget constraint (2) into the government budget constraint (9)

$$C_t + G_t = \Gamma_t + (1 - \nu) \int_0^1 (N_t(i)W_t(i)/P_t) di \quad (11)$$

Market Clearing

Firms' aggregate real profits are

$$\Gamma_t = \int_0^1 \Gamma_t(i) di = \frac{1}{P_t} \int_0^1 P_t(i)Y_t(i) - (1 - \nu)N_t(i)W_t(i) di = Y_t - (1 - \nu) \int_0^1 (N_t(i)W_t(i)/P_t) di \quad (12)$$

Plug (12) in (11), this gives us the goods market clearing condition for the economy

$$Y_t = C_t + G_t. \quad (13)$$

Intermediate Goods

The production of each intermediate good is done by a monopolist facing a production function linear in labor

$$Y_t(i) = A_t N_t(i) \quad (14)$$

$$A_t = \exp(u_t^A)$$

$$u_t^A = \mu + u_{t-1}^A + \varepsilon_{t-1}^A$$

with ε_{t-1}^A iid normally distributed, and A denotes the level of technology, common across firms. Each intermediate good producer has sticky prices, modeled as in Calvo (1983) where $1 - \lambda$ is the probability that each firm will be able to reoptimize its price each period. We allow for indexation of prices to steady-state inflation by firms who do not reoptimize their prices each period, with ω_s and ω_d respectively representing the degree of static and dynamic indexation (0 for no indexation to 1 for full indexation). Denoting the optimal reset price of firm i by $B(i)$, re-optimizing firms solve the following profit-maximization problem

$$\max_{B_t(i)} E_t \sum_{j=0}^{\infty} \lambda^j Q_{t,t+j} \left[Y_{t+j}(i) B_t(i) \bar{\Pi}^j \omega_s \left(\prod_{k=1}^j \Pi_{t+j-k} \right)^{\omega_d} - (1 - \nu) W_{t+j}(i) N_{t+j}(i) \right] \quad (15)$$

where $Q_{t,t+j} = \beta^j E_t \left\{ \frac{\xi_{t+j} P_t}{\xi_t P_{t+j}} \right\}$ is the stochastic discount factor, $\bar{\Pi}$ is the gross steady-state level of inflation and Π is gross level of inflation. The optimal relative reset price is then given by

$$\frac{B_t(i)}{P_t} = \frac{E_t \sum_{j=0}^{\infty} \lambda^j Q_{t,t+j} Y_{t+j} \left(\frac{P_{t+j}}{P_t} \right)^{\theta+1} \bar{\Pi}^{-j} \omega_s \theta \left(\prod_{k=1}^j \Pi_{t+j-k} \right)^{-\omega_d \theta} \left(\frac{\theta}{\theta-1} \right) (MC_{t+j}(i)/P_{t+j})}{E_t \sum_{j=0}^{\infty} \lambda^j Q_{t,t+j} Y_{t+j} (P_{t+j}/P_t)^{\theta} \bar{\Pi}^{-j} \omega_s (\theta-1) \left(\prod_{k=1}^j \Pi_{t+j-k} \right)^{-\omega_d (\theta-1)}} \quad (16)$$

Labor employed by firms each period is obtained through the minimization of production costs

$$\min_{N_t(i)} Costs_t(Y_t(i)) = (1 - \nu) W_t(i) N_t(i) \quad s.t. Y_t(i) = A_t N_t(i) \quad (17)$$

The FOC of problem (17) brings

$$MC_t(i) = \frac{(1-\nu)W_t(i)}{A_t} \quad (18)$$

Firm-specific marginal costs can be related to aggregate variables using

$$\frac{MC_{t+j}(i)}{P_{t+j}} = (1 - \nu) \left(\frac{\xi_{t+j}^{-1}}{A_{t+j}} \right) \left(\frac{Y_{t+j}}{A_{t+j}} \right)^{1/\eta} \left(\frac{B_t(i)}{P_t} \right)^{-\theta/\eta} \left(\frac{P_{t+j}}{\bar{\Pi}^j \omega_s \left(\prod_{k=1}^j \Pi_{t+j-k} \right)^{\omega_d} P_t} \right)^{\theta/\eta} \quad (19)$$

Note that in equilibrium all firms reoptimize to the same price, so $B_t/P_t = B_t(i)/P_t$.

Plugging (20) and the expression for $Q_{t,t+j}$ in (16) and rearranging

$$\left(\frac{B_t}{P_t} \right)^{\left(1 + \frac{\theta}{\eta}\right)} = \frac{E_t \sum_{j=0}^{\infty} \left(\frac{(1-\nu)\theta}{\theta-1} \right) \left(\frac{\beta\lambda}{\bar{\Pi}^j \omega_s \theta \left(1 + \frac{1}{\eta}\right)} \right)^j \left(\prod_{k=1}^j \Pi_{t+j-k} \right)^{-\omega_d \theta \left(1 + \frac{1}{\eta}\right)} \left(\frac{Y_{t+j}}{A_{t+j}} \right)^{\left(1 + \frac{1}{\eta}\right)} \left(\frac{P_{t+j}}{P_t} \right)^{\theta \left(1 + \frac{1}{\eta}\right)}}{E_t \sum_{j=0}^{\infty} \left(\frac{\beta\lambda}{\bar{\Pi}^j \omega_s (\theta-1)} \right)^j \left(\prod_{k=1}^j \Pi_{t+j-k} \right)^{-\omega_d (\theta-1)} (A_{t+j} \xi_{t+j}) \left(\frac{Y_{t+j}}{A_{t+j}} \right) \left(\frac{P_{t+j}}{P_t} \right)^{(\theta-1)}} \quad (20)$$

This equation can be log-linearized around the stochastic trend in technology as

$$\begin{aligned}
\left(1 + \frac{\theta}{\eta}\right) \hat{b}_t &= \sum_{j=0}^{\infty} [\gamma_2^j (1 - \gamma_2) - \gamma_1^j (1 - \gamma_1)] [\hat{y}_{t+j} + \hat{\xi}_{t+j}] + (1 - \gamma_2) \sum_{j=0}^{\infty} \gamma_2^j \left[\frac{1}{\eta} \hat{y}_{t+j} - \hat{\xi}_{t+j}\right] + \\
\sum_{j=0}^{\infty} \left[\gamma_2^{j+1} \theta \left(1 + \frac{1}{\eta}\right) - \gamma_1^{j+1} (\theta - 1)\right] E_t [\hat{\pi}_{t+j+1}] - \omega_d \sum_{j=0}^{\infty} \left[\gamma_2^{j+1} \theta \left(1 + \frac{1}{\eta}\right) - \gamma_1^{j+1} (\theta - 1)\right] \hat{\pi}_{t+j} + \hat{u}_t^m
\end{aligned} \tag{21}$$

Define F_t as the numerator of (20). It can be recursively expressed as

$$F_t = \left(\left(\frac{(1-\nu)\theta}{\theta-1} \right) \left(\frac{Y_t}{A_t} \right)^{\left(1+\frac{1}{\eta}\right)} + \left(\frac{\beta\lambda}{\bar{\Pi}^{\omega_s \theta \left(1+\frac{1}{\eta}\right)} \Pi_t^{\omega_d \theta \left(1+\frac{1}{\eta}\right)}} \right) E_t \left[\Pi_{t+1}^{\theta(1+1/\eta)} F_{t+1} \right] \right) \exp(u_t^m) \tag{22}$$

$$u_t^m = \rho_m u_{t-1}^m + \varepsilon_{t-1}^m$$

with ε_{t-1}^m iid normally distributed

where u_t^m is an ad-hoc cost push shock.

Define H_t as the denominator of (20). It can be recursively expressed as

$$H_t = (A_t \xi_t) \left(\frac{Y_t}{A_t} \right) + \left(\frac{\beta\lambda}{\bar{\Pi}^{\omega_s (\theta-1)} \Pi_t^{\omega_d (\theta-1)}} \right) E_t \left[\Pi_{t+1}^{(\theta-1)} H_{t+1} \right] \tag{23}$$

Therefore, (20) can be expressed as

$$\frac{B_t}{P_t} = \left(\frac{F_t}{H_t} \right)^{\left(\frac{1}{1+\theta/\eta} \right)} \tag{24}$$

Given these price-setting assumptions, the dynamics of the price level are governed by

$$P_t^{1-\theta} = (1 - \lambda) B_t^{1-\theta} + \lambda P_{t-1}^{1-\theta} \bar{\Pi}^{\omega_s (1-\theta)} \Pi_{t-1}^{\omega_d (1-\theta)}. \tag{25}$$

Dividing by $P_t^{1-\theta}$

$$1 = (1 - \lambda) \left(\frac{B_t}{P_t} \right)^{1-\theta} + \lambda \left(\frac{\bar{\Pi}^{\omega_s} \Pi_{t-1}^{\omega_d}}{\Pi_t} \right)^{1-\theta} \tag{26}$$

Plugging (24) on (26) and rearranging

$$\frac{F_t}{H_t} = \left(\frac{1 - \lambda \left(\frac{\bar{\Pi}^{\omega_s} \Pi_{t-1}^{\omega_d}}{\Pi_t} \right)^{1-\theta}}{1 - \lambda} \right)^{\frac{1+\theta/\eta}{1-\theta}} \tag{27}$$

We define the aggregate labor input as

$$N_t = \left[\int_0^1 N_t(i)^{(\theta-1)/\theta} di \right]^{\theta/(\theta-1)} \tag{28}$$

Plugging (15) on (30)

$$N_t = \left[\int_0^1 \left(\frac{Y_t(i)}{A_t} \right)^{(\theta-1)/\theta} di \right]^{\theta/(\theta-1)} \tag{29}$$

Plugging (7) on (31)

$$N_t = \frac{Y_t}{A_t} \left[\int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{1-\theta} di \right]^{-\left(\frac{\theta}{1-\theta}\right)} = \frac{Y_t}{A_t} \quad (30)$$

Monetary Policy

Finally, the policy rule followed by the monetary authority is

$$R_t = \max\{1, R_t^*\} \quad (31)$$

$$R_t^* = \bar{R} \left(\frac{R_{t-1}^*}{\bar{R}} \right)^{\rho_1} \left(\frac{R_{t-2}^*}{\bar{R}} \right)^{\rho_2} \left[\left(\frac{\Pi_t}{\bar{\Pi}} \right)^{\phi_\pi} \left(\frac{Y_t}{\bar{Y}_t} \right)^{\phi_Y} \left(\frac{GY_t}{\bar{G}\bar{Y}} \right)^{\phi_{GY}} \left(\frac{P_t}{\bar{P}_t} \right)^{\phi_P} \right]^{(1-\rho_1-\rho_2)} \exp(\varepsilon_t^R) \quad (32)$$

where R is *realized* gross interest rate, R^* is *desired* gross interest rate, GY is the gross growth rate of output and ε^R is an i.i.d policy shock. Note that equation (31) is responsible for introducing the zero lower bound to the model.

Equilibrium conditions

$$F_t = \left(\left(\frac{(1-\nu)\theta}{\theta-1} \right) \left(\frac{Y_t}{A_t} \right)^{\left(1+\frac{1}{\eta}\right)} + \left(\frac{\beta\lambda}{\bar{\Pi}\omega_s\theta\left(1+\frac{1}{\eta}\right)\Pi_t\omega_d\theta\left(1+\frac{1}{\eta}\right)} \right) E_t \left[\Pi_{t+1}^{\theta(1+1/\eta)} F_{t+1} \right] \right) \exp(u_t^m)$$

$$H_t = (A_t \xi_t) \left(\frac{Y_t}{A_t} \right) + \left(\frac{\beta\lambda}{\bar{\Pi}\omega_s(\theta-1)\Pi_t\omega_d(\theta-1)} \right) E_t \left[\Pi_{t+1}^{(\theta-1)} H_{t+1} \right]$$

$$\frac{F_t}{H_t} = \left(\frac{1 - \lambda \left(\frac{\bar{\Pi}\omega_s\Pi_{t-1}^{\omega_d}}{\Pi_t} \right)^{1-\theta}}{1 - \lambda} \right)^{\frac{1+\theta/\eta}{1-\theta}}$$

$$\xi_t = \beta E_t \left[\frac{\xi_{t+1} q_t R_t}{\Pi_{t+1}} \right]$$

$$\xi_t = (C_t - hGA_t C_{t-1})^{-1} - \beta h E_t GA_{t+1} (C_{t+1} - hGA_{t+1} C_t)^{-1}$$

$$Y_t = C_t + \exp(u_t^G) \bar{G}_t$$

$$R_t = \max\{1, R_t^*\}$$

$$R_t^* = \bar{R} \left(\frac{R_{t-1}^*}{\bar{R}} \right)^{\rho_1} \left(\frac{R_{t-2}^*}{\bar{R}} \right)^{\rho_2} \left[\left(\frac{\Pi_t}{\bar{\Pi}} \right)^{\phi_\pi} \left(\frac{Y_t}{\bar{Y}_t} \right)^{\phi_Y} \left(\frac{GY_t}{\bar{G}\bar{Y}} \right)^{\phi_{GY}} \left(\frac{P_t}{\bar{P}_t} \right)^{\phi_P} \right]^{(1-\rho_1-\rho_2)} \exp(\varepsilon_t^R)$$

$$GY_t = \frac{Y_t}{Y_{t-1}}$$

$$GC_t = \frac{C_t}{C_{t-1}}$$

$$\bar{P}_t = \bar{\Pi}^t$$

$$A_t = \exp(u_t^A)$$

$$q_t = \exp(u_t^q)$$

$$u_t^A = \mu + u_{t-1}^A + \varepsilon_t^A$$

$$u_t^q = \rho_q u_{t-1}^q + \varepsilon_t^q$$

$$u_t^G = \rho_G u_{t-1}^G + \varepsilon_t^G$$

$$u_t^m = \rho_m u_{t-1}^m + \varepsilon_t^m$$

Equilibrium conditions, stationary variables

$$F_t = \left(\left(\frac{(1-\nu)\theta}{\theta-1} \right) \tilde{Y}_t^{(1+\frac{1}{\eta})} + \left(\frac{\beta\lambda}{\bar{\Pi} \omega_s \theta^{(1+\frac{1}{\eta})} \Pi_t^{\omega_d \theta (1+\frac{1}{\eta})}} \right) E_t \left[\Pi_{t+1}^{\theta(1+1/\eta)} F_{t+1} \right] \right) \exp(u_t^m)$$

$$H_t = \tilde{\xi}_t \tilde{Y}_t + \left(\frac{\beta\lambda}{\bar{\Pi} \omega_s (\theta-1) \Pi_t^{\omega_d (\theta-1)}} \right) E_t \left[\Pi_{t+1}^{(\theta-1)} H_{t+1} \right]$$

$$\frac{F_t}{H_t} = \left(\frac{1 - \lambda \left(\frac{\bar{\Pi} \omega_s \Pi_{t-1}^{\omega_d}}{\Pi_t} \right)^{1-\theta}}{1 - \lambda} \right)^{\frac{1+\theta/\eta}{1-\theta}}$$

$$\tilde{\xi}_t = (\tilde{C}_t - h\tilde{C}_{t-1})^{-1} - \beta h E_t (\tilde{C}_{t+1} - h\tilde{C}_t)^{-1}$$

$$\tilde{\xi}_t = \beta E_t \left[\frac{\tilde{\xi}_{t+1} q_t R_t}{GA_{t+1} \Pi_{t+1}} \right]$$

$$\tilde{Y}_t = \tilde{C}_t + \exp(u_t^G) \bar{G}$$

$$GA_t = \exp(\mu + \varepsilon_t^A)$$

$$GY_t = \frac{\tilde{Y}_t}{\tilde{Y}_{t-1}} GA_t$$

$$GC_t = \frac{\tilde{C}_t}{\tilde{C}_{t-1}} GA_t$$

$$\bar{P}_t = \bar{P}_{t-1} \frac{\Pi_t}{\bar{\Pi}}$$

$$R_t^* = \bar{R} \left(\frac{R_{t-1}^*}{\bar{R}} \right)^{\rho_1} \left(\frac{R_{t-2}^*}{\bar{R}} \right)^{\rho_2} \left[\left(\frac{\Pi_t}{\bar{\Pi}} \right)^{\phi_\pi} \left(\frac{\tilde{Y}_t}{\bar{Y}} \right)^{\phi_Y} \left(\frac{GY_t}{\bar{G}\bar{Y}} \right)^{\phi_{GY}} \bar{P}_t^{\phi_P} \right]^{(1-\rho_1-\rho_2)} \exp(\varepsilon_t^R)$$

$$R_t = \max\{1, R_t^*\}$$

$$q_t = \exp(u_t^q)$$

$$u_t^q = \rho_q u_{t-1}^q + \varepsilon_t^q$$

$$u_t^G = \rho_G u_{t-1}^G + \varepsilon_t^G$$

$$u_t^m = \rho_m u_{t-1}^m + \varepsilon_t^m$$

where $\tilde{\xi}_t = A_t \xi_t$, $\tilde{C}_t = C_t/A_t$ and $\tilde{Y}_t = Y_t/A_t$

Steady state values

$$\bar{Y} = \left[\left(\frac{(\theta-1)(1-\beta h)(1-\beta \lambda \bar{\Pi}^{(1-\omega_s-\omega_d)\theta(1+1/\eta)})}{(1-\nu)\theta(1-h)(1-\bar{G})(1-\beta \lambda \bar{\Pi}^{(1-\omega_s-\omega_d)(\theta-1)})} \right) \left(\frac{1-\lambda \bar{\Pi}^{(1-\omega_s-\omega_d)(\theta-1)}}{1-\lambda} \right)^{\frac{1+\theta/\eta}{1-\theta}} \right]^{\frac{1}{1+1/\eta}}$$

$$\bar{F} = \left(\frac{(1-\nu)\theta}{(\theta-1)(1-\beta \lambda \bar{\Pi}^{(1-\omega_s-\omega_d)\theta(1+1/\eta)})} \right) \bar{Y}^{(1+1/\eta)}$$

$$\bar{H} = \frac{1-\beta h}{(1-h)(1-\bar{G})(1-\beta \lambda \bar{\Pi}^{(1-\omega_s-\omega_d)(\theta-1)})}$$

$$\bar{C} = (1-\bar{G})\bar{Y}$$

$$\bar{\xi} = \left(\frac{1-\beta h}{(1-h)(1-\bar{G})} \right) \bar{Y}^{-1}$$

$$\bar{R} = \frac{\bar{\Pi} \bar{G} \bar{Y}}{\beta}$$

$$\bar{R}^S = \bar{R}$$

$$\bar{G}\bar{Y} = \exp(\mu)$$

$$\bar{P} = 1$$

$$\bar{u}^q = 0$$

$$\bar{u}^G = 0$$

$$\bar{u}^m = 0$$

Log-linearized equilibrium conditions around stochastic trend in technology

$$\begin{aligned} \hat{f}_t &= \left(1 - \beta\lambda\bar{\Pi}^{(1-\omega_s-\omega_d)\theta\left(1+\frac{1}{\eta}\right)}\right)\left(1 + \frac{1}{\eta}\right)\hat{y}_t + \left(\beta\lambda\bar{\Pi}^{(1-\omega_s-\omega_d)\theta\left(1+\frac{1}{\eta}\right)}\right)\left[\theta\left(1 + \frac{1}{\eta}\right)\hat{\pi}_{t+1} - \omega_d\theta\left(1 + \frac{1}{\eta}\right)\hat{\pi}_t + \hat{f}_{t+1}\right] + \hat{u}_t^m \\ \hat{h}_t &= \left(1 - \beta\lambda\bar{\Pi}^{(1-\omega_s-\omega_d)(\theta-1)}\right)\left[\hat{\xi}_t + \hat{y}_t\right] + \left(\beta\lambda\bar{\Pi}^{(1-\omega_s-\omega_d)(\theta-1)}\right)\left[(\theta - 1)\hat{\pi}_{t+1} - \omega_d(\theta - 1)\hat{\pi}_t + \hat{h}_{t+1}\right] \\ \hat{f}_t - \hat{h}_t &= \left(1 + \frac{\theta}{\eta}\right)\left(\frac{\lambda\bar{\Pi}^{(1-\omega_s-\omega_d)(\theta-1)}}{1 - \lambda\bar{\Pi}^{(1-\omega_s-\omega_d)(\theta-1)}}\right)(\hat{\pi}_t - \omega_d\hat{\pi}_{t-1}) \\ \hat{\xi}_t &= \frac{h}{(1-h)(1-\beta h)}\hat{c}_{t-1} - \frac{1 + \beta h^2}{(1-h)(1-\beta h)}\hat{c}_t + \frac{\beta h}{(1-h)(1-\beta h)}E_t\hat{c}_{t+1} \\ \hat{\xi}_t &= E_t[\hat{\xi}_{t+1} + \hat{r}_t - \hat{\pi}_{t+1} + \hat{u}_t^q] \\ \hat{y}_t &= (1 - s_g)\hat{c}_t - s_g\hat{u}_t^G \\ \hat{r}_t &= \max\{\hat{r}_t^*, -\bar{r}\} \\ \hat{r}_t^* &= \rho_1\hat{r}_{t-1}^* + \rho_2\hat{r}_{t-2}^* + (1 - \rho_1 - \rho_2)[\phi_\pi\hat{\pi}_t + \phi_y\hat{y}_t + \phi_{gy}\widehat{gy}_t + \phi_p\hat{p}_t] + \varepsilon_t^r \\ \widehat{gy}_t &= \hat{y}_t - \hat{y}_{t-1} + \varepsilon_t^A \\ \hat{p}_t &= \hat{p}_{t-1} + \hat{\pi}_t \\ \hat{u}_t^q &= \rho_q\hat{u}_{t-1}^q + \varepsilon_t^q \\ \hat{u}_t^G &= \rho_G\hat{u}_{t-1}^G + \varepsilon_t^G \\ \hat{u}_t^m &= \rho_m\hat{u}_{t-1}^m + \varepsilon_t^m \end{aligned}$$