The Zero Lower Bound and Estimation Accuracy*

Tyler Atkinson   Alexander W. Richter   Nathaniel A. Throckmorton

First Draft: May 7, 2018

This Draft: June 3, 2019

Abstract

During the Great Recession, many central banks lowered their policy rate to its zero lower bound (ZLB), creating a kink in the policy rule and calling into question linear estimation methods. There are two promising alternatives: estimate a fully nonlinear model that accounts for precautionary savings effects of the ZLB or a piecewise linear model that is much faster but ignores the precautionary savings effects. This paper compares the accuracy of the two methods using artificial datasets. We find the predictions of the nonlinear model are typically more accurate than the piecewise linear model, but the differences are usually small. There are far larger gains in accuracy from estimating a richer, less misspecified piecewise linear model.

Keywords: Bayesian Estimation; Projection Methods; Particle Filter; OccBin; Inversion Filter
JEL Classifications: C11; C32; C51; E43

*Atkinson and Richter, Research Department, Federal Reserve Bank of Dallas, 2200 N Pearl Street, Dallas, TX 75201 (tyler.atkinson@dal.frb.org; alex.richter@dal.frb.org); Throckmorton, Department of Economics, William & Mary, P.O. Box 8795, Williamsburg, VA 23187 (nat@wm.edu). We thank Boragan Aruoba, Alex Chudik, Marc Giannoni, Matthias Hartmann, Rob Hicks, Ben Johannsen, Giorgio Primiceri, Sanjay Singh, Kei-Mu Yi, and an anonymous referee for suggestions that improved the paper. We also thank Chris Stackpole and Eric Walter for supporting the supercomputers at our institutions. This research was completed with supercomputing resources provided by the Federal Reserve banks of Kansas City and Dallas, William & Mary, Auburn University, the University of Texas at Dallas, the Texas Advanced Computing Center, and Southern Methodist University. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 Introduction

Using Bayesian methods to estimate linear dynamic stochastic general equilibrium (DSGE) models has become common practice in the literature over the last 20 years. Many central banks also use these models for forecasting and counterfactual simulations. The estimation procedure sequentially draws parameters from a proposal distribution, solves the model given that draw, and then evaluates the likelihood function. With linearity and normally distributed shocks, the model solves in a fraction of a second and it is easy to exactly evaluate the likelihood function with a Kalman filter.[1]

The financial crisis and subsequent recession compelled many central banks to take unprecedented action to reduce their policy rate to its zero lower bound (ZLB), calling into question linear estimation methods. The ZLB constraint presents a challenge for empirical work because it creates a kink in the central bank's policy rule. The constraint has always existed, but when policy rates were well above zero and the likelihood of hitting the constraint was negligible, it was reasonable to ignore it. The lengthy period of near zero policy rates over the last decade and the increased likelihood of future ZLB events due to estimates of a lower natural rate has forced researchers to think more carefully about the ZLB constraint and its implications (e.g., Laubach and Williams, 2016).

There are two promising estimation methods used in the literature that account for the ZLB constraint in DSGE models. The first method estimates a fully nonlinear model with an occasionally binding ZLB constraint (e.g., Gust et al., 2017; Plante et al., 2018; Richter and Throckmorton, 2016). This method provides the most comprehensive treatment of the ZLB constraint but is numerically intensive. It uses projection methods to solve the nonlinear model and a particle filter to evaluate the likelihood function for each draw from the posterior distribution (henceforth, NL-PF).[2]

The second method estimates a piecewise linear version of the nonlinear model (e.g., Guerrieri and Iacoviello, 2017). The model is solved using the OccBin toolbox developed by Guerrieri and Iacoviello (2015). The likelihood is evaluated using an inversion filter, which solves for the shocks that minimize the distance between the data and the model predictions. The benefit of this method (henceforth, OB-IF) is that it is nearly as fast as estimating a linear model with a Kalman filter while still capturing the kink in the decision rules created by the ZLB. However, OB-IF differs from NL-PF in a potentially important way. Households do not account for the possibility that the ZLB may bind in the future when it does not currently bind, which is inconsistent with survey data.[3]

[1] Schorfheide (2000) and Otrok (2001) were the first to use these methods to generate draws from the posterior distribution of a linear DSGE model. See An and Schorfheide (2007) and Herbst and Schorfheide (2016) for examples.

[2] Several papers examine the effects of the ZLB constraint in a *calibrated* nonlinear model using projection methods similar to ours (e.g., Aruoba et al., 2018; Fernández-Villaverde et al., 2015; Gavin et al., 2015; Keen et al., 2017; Mertens and Ravn, 2014; Nakata, 2017; Nakov, 2008; Ngo, 2014; Richter and Throckmorton, 2015; Wolman, 2005).

[3] The inversion filter also removes the interest rate as an observable and sets the monetary policy shock to zero when the ZLB binds, whereas the particle filter estimates those shocks given the data. This difference is important to the extent that monetary policy shocks impact the economy at the ZLB. However, in practice, the particle filter typically estimates monetary policy shocks close to zero when the ZLB binds, suggesting there is little it can identify.

This paper compares the accuracy of the two estimation methods. We specify a true parameterization of a medium-scale nonlinear model with an occasionally binding ZLB constraint, solve the model with a projection method, and generate a large sample of datasets. The datasets either contain no ZLB events or a single event with various durations to understand the influence of the ZLB on the posterior estimates. For each dataset, we use NL-PF and OB-IF to estimate a small-scale, but nested, version of the medium-scale model that generates the data. We also estimate the linear model with a Kalman filter (henceforth, Lin-KF), since it was the most common method before the Great Recession. The small-scale model excludes features of the medium-scale model that others have shown are empirically important. The difference between the two models—referred to as misspecification—account for the practical reality that all models are misspecified. It also sheds light on the merits of estimating a simpler, more misspecified, model with NL-PF, versus a richer, less misspecified, model with OB-IF that is numerically very costly with fully nonlinear methods.

We find NL-PF and OB-IF produce similar parameter estimates. In contrast, the predictions and forecasting performance of NL-PF are typically more accurate than OB-IF. For example, the estimates of the notional interest rate (the rate the central bank would set in the absence of the ZLB constraint), the expected ZLB duration, the probability of a 4 quarter or longer ZLB event, and forecasts of the policy rate are closer to their actual values. The increase in accuracy, however, is often small because the precautionary savings effects of the ZLB and the effects of other nonlinearities are weak in canonical models. The benefits also come with a steep increase in estimation time. The model takes roughly a week to estimate with NL-PF versus a couple hours with OB-IF.

These results suggest that OB-IF may provide an adequate substitute for NL-PF, but there are two important caveats. One, our analysis focuses exclusively on the ZLB constraint. Other constraints could create inaccuracies that provide a stronger justification for the computational burden of NL-PF. Two, OB-IF only captures nonlinearities from occasionally binding constraints. OB-IF could not account for nonlinear features such as stochastic volatility, non-convex adjustment costs, endogenous regime-switching, default, Bayesian learning, and non-Gaussian shock distributions. Our results will provide a useful benchmark for future work that examines these nonlinear features.

Model misspecification has a much larger impact on accuracy than the estimation method. It biases many of the parameter estimates and often creates significant differences between the predictions of the estimated models and the data generating process (DGP). These results suggest researchers are better off reducing misspecification by estimating a richer piecewise linear model than a simpler but computationally less intensive nonlinear model when the ZLB binds in the data. This important finding could open the door to promising new work on the implications of the ZLB.

Our paper is the first to compare different estimation methods that account for the ZLB constraint. Others compare nonlinear estimation methods to linear methods. Fernández-Villaverde and Rubio-Ramírez (2005) show that a neoclassical growth model estimated with NL-PF predicts moments closer to the true moments than the estimates from Lin-KF using two artificial datasets and actual data. The primary source of nonlinearity in their model is high risk aversion. Hirose and Inoue (2016) generate artificial datasets from a linear model where the ZLB constraint is imposed using anticipated monetary policy shocks and then apply Lin-KF to estimate the model without the constraint. They find the estimated parameters, impulse responses, and structural shocks become less accurate as the frequency and duration of ZLB events increase in the data. Hirose and Sunakawa (2015) extend that work by generating data from a nonlinear model and re-examine the bias. None of these papers introduce misspecification, which is an important aspect of our analysis.

We also build on recent empirical work that analyzes the implications of the ZLB constraint (e.g., Gust et al., 2017; Iiboshi et al., 2018; Plante et al., 2018; Richter and Throckmorton, 2016). These papers use NL-PF to estimate a nonlinear model similar to ours using actual data from the U.S. or Japan that includes the ZLB period. Our contribution is to examine the accuracy of these nonlinear estimation methods and show under what conditions they outperform other approaches.

The measurement error (ME) in the observation equation of the filter is a key aspect of the estimation procedure that could potentially affect the accuracy of the parameter estimates. Unlike the inversion filter, the particle filter requires positive ME variances to prevent degeneracy—a situation when the likelihood is inaccurate. The literature has used a wide range of different values, with limited investigation on how they impact accuracy. Canova et al. (2014) show the downside of introducing ME is that the posterior distributions of some parameters do not contain the truth in a DSGE model estimated with Lin-KF. Cuba-Borda et al. (2017) show that ME in the particle filter reduces the accuracy of the likelihood function using a calibrated model with an occasionally binding borrowing constraint. Our analysis provides a potentially important role for ME because it includes model misspecification. We find larger ME variances improve the accuracy of some parameters, but the benefits are more than offset by decreases in the accuracy of other parameters.[4]

The paper proceeds as follows. Section 2 describes our DGP and how we construct our artificial datasets. Section 3 outlines the estimated model and estimation methods. Section 4 shows our posterior estimates and several measures of accuracy for each estimation method. Section 5 concludes.

2 DATA GENERATING PROCESS

To test the accuracy of recent estimation methods that account for the ZLB constraint, we generate a large number of artificial datasets from a medium-scale New Keynesian model with capital and an occasionally binding ZLB constraint. Our model is the same as the one in Gust et al. (2017), except it removes government spending, inflation indexation, and the investment efficiency shock.[5]

[4] Herbst and Schorfheide (2018) develop a tempered particle filter that sequentially reduces the ME variances. They assess accuracy against the Kalman filter on U.S. data with a linear model and find it outperforms the untempered filter.

[5] Appendix E.7 shows how the addition of government spending to the DGP and estimated model affects our results.

2.1 FIRMS The production sector consists of a continuum of monopolistically competitive intermediate goods firms and a final goods firm. Intermediate firm $f \in [0,1]$ produces a differentiated good, $y(f)$, according to $y_t(f) = (\upsilon_t k_{t-1}(f))^{\alpha}(a_t n_t(f))^{1-\alpha}$, where $n(f)$ is the labor hired by firm f and $k(f)$ is the capital rented by firm f. $a_t = z_t a_{t-1}$ is productivity and υ is the capital utilization rate, which are both common across firms. Deviations from the steady-state growth rate, $\bar{z}$, follow

$$z_t = \bar{z} + \sigma_z \varepsilon_{z,t},\ \varepsilon_z \sim \mathbb{N}(0,1). \tag{1}$$

The final goods firm purchases output from each intermediate firm to produce the final good, $y_t \equiv [\int_0^1 y_t(f)^{(\theta_p-1)/\theta_p} df]^{\theta_p/(\theta_p-1)}$, where $\theta_p > 1$ is the elasticity of substitution. Dividend maximization determines the demand for intermediate good f, $y_t(f) = (p_t(f)/p_t)^{-\theta_p} y_t$, where $p_t = [\int_0^1 p_t(f)^{1-\theta_p} df]^{1/(1-\theta_p)}$ is the price level. Following Rotemberg (1982), intermediate firms pay a price adjustment cost, $adj_t^p(f) \equiv \varphi_p(p_t(f)/(\bar{\pi} p_{t-1}(f))-1)^2 y_t/2$, where $\varphi_p > 0$ scales the cost and $\bar{\pi}$ is the steady-state gross inflation rate. Given this cost, firm f chooses $n_t(f)$, $k_{t-1}(f)$, and $p_t(f)$ to maximize the expected discounted present value of future dividends, $E_t \sum_{k=t}^{\infty} q_{t,k} d_k(f)$, subject to its production function and the demand for its product, where $q_{t,t} \equiv 1$, $q_{t,t+1} \equiv \beta(\lambda_t/\lambda_{t+1})$ is the pricing kernel between periods t and $t+1$, $q_{t,k} \equiv \prod_{j=t+1}^{k>t} q_{j-1,j}$, and $d_t(f) = p_t(f) y_t(f)/p_t - w_t n_t(f) - r_t^k \upsilon_t k_{t-1}(f) - adj_t^p(f)$. In symmetric equilibrium, the optimality conditions reduce to

$$y_t = (\upsilon_t k_{t-1})^{\alpha}(a_t n_t)^{1-\alpha}, \tag{2}$$

$$w_t = (1-\alpha) mc_t y_t / n_t, \tag{3}$$

$$r_t^k = \alpha mc_t y_t / (\upsilon_t k_{t-1}), \tag{4}$$

$$\varphi_p(\pi_t/\bar{\pi} - 1)(\pi_t/\bar{\pi}) = 1 - \theta_p + \theta_p mc_t + \beta \varphi_p E_t[(\lambda_t/\lambda_{t+1})(\pi_{t+1}/\bar{\pi} - 1)(\pi_{t+1}/\bar{\pi}) y_{t+1}/y_t], \tag{5}$$

where $\pi_t = p_t/p_{t-1}$ is the gross inflation rate. If $\varphi_p = 0$, the real marginal cost of producing a unit of output (mc_t) equals $(\theta_p - 1)/\theta_p$, which is the inverse of the markup of price over marginal cost.

2.2 HOUSEHOLDS Each household consists of a unit mass of members who supply differentiated types of labor, $n(\ell)$, at real wage rate $w(\ell)$. A perfectly competitive labor union bundles the labor types to produce an aggregate labor product, $n_t \equiv [\int_0^1 n_t(\ell)^{(\theta_w-1)/\theta_w} d\ell]^{\theta_w/(\theta_w-1)}$, where $\theta_w > 1$ is the elasticity of substitution. Dividend maximization determines the demand for labor type ℓ, $n_t(\ell) = (w_t(\ell)/w_t)^{-\theta_w} n_t$, where $w_t = [\int_0^1 w_t(\ell)^{1-\theta_w} d\ell]^{1/(1-\theta_w)}$ is the aggregate real wage.

The households choose $\{c_t, n_t, b_t, x_t, k_t, \upsilon_t\}_{t=0}^{\infty}$ to maximize expected lifetime utility given by $E_0 \sum_{t=0}^{\infty} \beta^t [\log(c_t - h c_{t-1}^a) - \chi \int_0^1 n_t(\ell)^{1+\eta} d\ell/(1+\eta)]$, where β is the discount factor, χ determines steady-state labor, $1/\eta$ is the Frisch labor supply elasticity, c is consumption, c^a is aggregate consumption, h is the degree of external habit persistence, b is the real value of a privately-issued 1-period nominal bond, x is investment, and E_0 is an expectation operator conditional on information available in period 0. Following Chugh (2006), the nominal wage rate for each labor type is subject to an adjustment cost, $adj_t^w(\ell) = \varphi_w(w_t^g(\ell) - 1)^2 y_t/2$, where $w_t^g(\ell) = \pi_t w_t(\ell)/(\bar{\pi}\bar{z}w_{t-1}(\ell))$ is nominal wage growth relative its steady-state. The cost of utilizing the capital shock, u, is given by

$$u_t = \bar{r}^k(\exp(\sigma_\upsilon(\upsilon_t - 1)) - 1)/\sigma_\upsilon, \tag{6}$$

where $\sigma_\upsilon \geq 0$ scales the cost. Given the two costs, the household's budget constraint is given by

$$c_t + x_t + b_t/(s_t i_t) + u_t k_{t-1} + \int_0^1 adj_t^w(\ell)d\ell = \int_0^1 w_t(\ell)n_t(\ell)d\ell + r_t^k \upsilon_t k_{t-1} + b_{t-1}/\pi_t + d_t,$$

where i is the gross nominal interest rate, r^k is the capital rental rate, and d is a real dividend from ownership of intermediate firms. The nominal bond, b, is subject to a risk premium, s, that follows

$$s_t = (1 - \rho_s)\bar{s} + \rho_s s_{t-1} + \sigma_s \varepsilon_{s,t},\ 0 \leq \rho_s < 1,\ \varepsilon_s \sim \mathbb{N}(0, 1), \tag{7}$$

where $\bar{s}$ is the steady-state value. An increase in s_t boosts saving, which lowers period-t demand.

Households also face an investment adjustment cost, so the law of motion for capital is given by

$$k_t = (1 - \delta)k_{t-1} + x_t(1 - \nu(x_t^g - 1)^2/2),\ 0 \leq \delta \leq 1, \tag{8}$$

where $x_t^g = x_t/(\bar{z}x_{t-1})$ is investment growth relative to its steady-state and $\nu \geq 0$ scales the cost.

The first order conditions to each household's constrained optimization problem are given by

$$r_t^k = \bar{r}^k \exp(\sigma_\upsilon(\upsilon_t - 1)), \tag{9}$$

$$\lambda_t = c_t - hc_{t-1}^a, \tag{10}$$

$$w_t^f = \chi n_t^\eta \lambda_t, \tag{11}$$

$$1 = \beta E_t[(\lambda_t/\lambda_{t+1})(s_t i_t/\pi_{t+1})], \tag{12}$$

$$q_t = \beta E_t[(\lambda_t/\lambda_{t+1})(r_{t+1}^k \upsilon_{t+1} - u_{t+1} + (1 - \delta)q_{t+1})], \tag{13}$$

$$1 = q_t[1 - \nu(x_t^g - 1)^2/2 - \nu(x_t^g - 1)x_t^g] + \beta\nu\bar{z}E_t[(\lambda_t/\lambda_{t+1})q_{t+1}(x_{t+1}^g)^2(x_{t+1}^g - 1)], \tag{14}$$

$$\varphi_w(w_t^g - 1)w_t^g = [(1 - \theta_w)w_t + \theta_w w_t^f]n_t/y_t + \beta\varphi_w E_t[(\lambda_t/\lambda_{t+1})(w_{t+1}^g - 1)w_{t+1}^g y_{t+1}/y_t], \tag{15}$$

where $1/\lambda$ is the marginal utility of consumption, q is Tobin's q, and w^f is the flexible wage rate.

Monetary Policy The central bank sets the gross nominal interest rate, i, according to

$$i_t = \max\{1, i_t^n\}, \tag{16}$$

$$i_t^n = (i_{t-1}^n)^{\rho_i}(\bar{\imath}(\pi_t/\bar{\pi})^{\phi_\pi}(y_t^{gdp}/(y_{t-1}^{gdp}\bar{z}))^{\phi_y})^{1-\rho_i}\exp(\sigma_i\varepsilon_{i,t}),\ 0 \leq \rho_i < 1,\ \varepsilon_i \sim \mathbb{N}(0, 1), \tag{17}$$

where y^{gdp} is real GDP (i.e., output, y, minus the resources lost due to adjustment costs, adj^p and

adj^w, and utilization costs), i^n is the gross notional interest rate, $\bar{\imath}$ and $\bar{\pi}$ are the target values of the inflation and nominal interest rates, and ϕ_π and ϕ_y are the responses to the inflation and output growth gaps. A more negative net notional rate indicates that the central bank is more constrained.

Competitive Equilibrium The aggregate resource constraint and real GDP definition are given by

$$c_t + x_t = y_t^{gdp}, \tag{18}$$

$$y_t^{gdp} = [1 - \varphi_p(\pi_t/\bar{\pi} - 1)^2/2 - \varphi_w(w_t^g - 1)^2/2]y_t - u_t k_{t-1}. \tag{19}$$

The model does not have a steady-state due to the unit root in productivity, a_t. Therefore, we define the variables with a trend in terms of productivity (i.e., $\tilde{x}_t \equiv x_t/a_t$). The detrended equilibrium system is provided in Appendix A. A competitive equilibrium consists of sequences of quantities, $\{\tilde{c}_t, \tilde{y}_t, \tilde{y}_t^{gdp}, x_t^g, y_t^g, n_t, \tilde{k}_t, \tilde{x}_t\}_{t=0}^\infty$, prices, $\{\tilde{w}_t, \tilde{w}_t^f, \tilde{w}_t^g, i_t, i_t^n, \pi_t, \tilde{\lambda}_t, \upsilon_t, u_t, q_t, r_t^k, mc_t\}_{t=0}^\infty$, and exogenous variables, $\{s_t, z_t\}_{t=0}^\infty$, that satisfy the detrended equilibrium system, given the initial conditions, $\{\tilde{c}_{-1}, i_{-1}^n, \tilde{k}_{-1}, \tilde{x}_{-1}, \tilde{w}_{-1}, s_0, z_0, \varepsilon_{i,0}\}$, and three sequences of shocks, $\{\varepsilon_{z,t}, \varepsilon_{s,t}, \varepsilon_{i,t}\}_{t=1}^\infty$.

Subjective Discount Factor	β	0.9949	Rotemberg Price Adjustment Cost	φ_p	100
Frisch Labor Supply Elasticity	$1/\eta$	3	Rotemberg Wage Adjustment Cost	φ_w	100
Price Elasticity of Substitution	θ_p	6	Capital Utilization Curvature	σ_υ	5
Wage Elasticity of Substitution	θ_w	6	Inflation Gap Response	ϕ_π	2
Steady-State Labor Hours	$\bar{n}$	0.3333	Output Growth Gap Response	ϕ_y	0.5
Steady-State Risk Premium	$\bar{s}$	1.0058	Habit Persistence	h	0.8
Steady-State Growth Rate	$\bar{z}$	1.0034	Risk Premium Persistence	ρ_s	0.8
Steady-State Inflation Rate	$\bar{\pi}$	1.0053	Notional Rate Persistence	ρ_i	0.8
Capital Share of Income	α	0.35	Productivity Growth Shock SD	σ_z	0.005
Capital Depreciation Rate	δ	0.025	Risk Premium Shock SD	σ_s	0.005
Investment Adjustment Cost	ν	4	Notional Interest Rate Shock SD	σ_i	0.002

Table 1: Parameter values for the data generating process.

2.3 PARAMETER VALUES Table 1 shows the model parameters for the DGP. The parameters were chosen so our DGP is characteristic of recent U.S. data. The steady-state growth rate ($\bar{z}$), inflation rate ($\bar{\pi}$), risk-premium ($\bar{s}$), and capital share of income (α) are equal to the time averages of per capita real GDP growth, the percent change in the GDP implicit price deflator, the Baa corporate bond yield relative to the yield on the 10-Year Treasury rate, and the Fernald (2012) utilization-adjusted quarterly-TFP estimates of the capital share of income from 1988Q1-2017Q4.

The subjective discount factor, β, is set to 0.9949, which is the time average of the values implied by the steady-state Euler equation and the federal funds rate. The corresponding annualized steady-state nominal interest rate is 3.3%, which is consistent with the sample average and current long-run estimates of the federal funds rate. The leisure preference parameter, χ, is set so steady-state labor equals $1/3$ of the available time. The capital depreciation rate is set to 0.025. Both values are ubiquitous in the literature. The elasticities of substitution between intermediate goods and labor types, θ_p and θ_w, are set to 6, which correspond to a 20% average markup in each sector and match the values used in Gust et al. (2017). The Frisch elasticity of labor supply, $1/\eta$, is set to 3 to match the macro estimate in Peterman (2016). The investment adjustment cost parameter, ν, and capital utilization curvature, σ_υ, are consistent with the estimates in Gust et al. (2017). The price and wage adjustment cost parameters, φ_p and φ_w, are both set to 100, which correspond to Phillips curve slopes of 0.050 and 0.027. Estimates for the monetary responses to the inflation and output growth gaps, ϕ_π and ϕ_y vary in the literature, ranging from 1.5-2.5 and 0-1 (Aruoba et al., 2018; Gust et al., 2017). We set $\phi_\pi = 2.0$ and $\phi_y = 0.5$, which are in the middle of those ranges.

The persistence parameters and shock standard deviations are set to values that are in line with the estimates from Aruoba et al. (2018) and Gust et al. (2017). The most consequential parameters are the risk premium persistence and shock standard deviation because they have the largest impact on the expected frequency and duration of ZLB events. When either of those parameters increase, households place more weight on outcomes where the central bank cannot respond to adverse shocks by lowering the nominal interest rate, which increases the downward bias from the ZLB.

2.4 SOLUTION AND SIMULATION METHODS

We solve the nonlinear model with the policy function iteration algorithm described in Richter et al. (2014), which is based on the theoretical work on monotone operators in Coleman (1991). We discretize the endogenous state variables and approximate the exogenous states, s_t, z_t, and $\varepsilon_{i,t}$ using the N-state Markov chain in Rouwenhorst (1995). The Rouwenhorst method is attractive because it only requires us to interpolate along the dimensions of the endogenous state variables, which makes the solution more accurate and faster than quadrature methods. To obtain initial conjectures for the nonlinear policy functions, we solve the level-linear analogue of our nonlinear model with Sims's (2002) gensys algorithm. Then we minimize the Euler equation errors on every node in the state space and compute the maximum distance between the updated policy functions and the initial conjectures. Finally, we replace the initial conjectures with the updated policy functions and iterate until the maximum distance is below the tolerance level. See Appendix B for a more detailed description of the solution method.

We generate data for output growth, the inflation rate, and the nominal interest rate by simulating the model using the nonlinear policy functions, so the observables are given by $\mathbf{x}_t = [y_t^g, \pi_t, i_t]$. Each simulation is initialized with a draw from the ergodic distribution and contains 120 quarters, similar to what is often used when estimating models with actual data. We use samples from the DGP with either no ZLB events or a single ZLB event that is 5%, 10%, 15%, 20%, and 25% of the sample. Our sample is 120 quarters, so the ZLB events are either 6, 12, 18, 24, or 30 quarters long. The longest events reflect the recent experiences of some advanced economies, such as the U.S. and Japan. We create 50 datasets for each ZLB duration. Appendix E.6 provides more information.

3 ESTIMATION METHODS

The medium-scale model is costly to estimate with global methods, which causes researchers to work with smaller models. To account for this reality, we simulate data from the fully nonlinear model and test the accuracy of different estimation methods on a small-scale nonlinear model that does not include capital or sticky wages. Therefore, the estimated model contains misspecification. The medium-scale model that generates our data collapses to the small-scale model when $\alpha = \varphi_w = 0$ and $\theta_w \to \infty$. The equilibrium system includes (1), (5), (7), (10), (12), (16), (17), and

$$y_t = a_t n_t, \tag{20}$$

$$w_t = mc_t y_t / n_t, \tag{21}$$

$$w_t = \chi n_t^\eta \lambda_t, \tag{22}$$

$$c_t = y_t^{gdp}, \tag{23}$$

$$y_t^{gdp} = [1 - \varphi_p(\pi_t/\bar{\pi} - 1)^2/2] y_t. \tag{24}$$

Once again, we remove the trend in productivity and provide the detrended equilibrium system in Appendix A. The competitive equilibrium includes sequences of quantities, $\{\tilde{c}_t, \tilde{y}_t, \tilde{y}_t^{gdp}, y_t^g, n_t\}_{t=0}^\infty$, prices, $\{\tilde{w}_t, i_t, i_t^n, \pi_t, \tilde{\lambda}_t, mc_t\}_{t=0}^\infty$, and exogenous variables, $\{s_t, z_t\}_{t=0}^\infty$, that satisfy the detrended system, given the initial conditions, $\{\tilde{c}_{-1}, i_{-1}^n, s_0, z_0, \varepsilon_{i,0}\}$, and shock sequences, $\{\varepsilon_{z,t}, \varepsilon_{s,t}, \varepsilon_{i,t}\}_{t=1}^\infty$.

We estimate the small-scale model with Bayesian methods. For each dataset, we draw parameters from a proposal distribution, solve the model conditional on the draw, and filter the data to evaluate the likelihood function within a random walk Metropolis-Hastings algorithm. Within this framework, we test the accuracy of two promising estimation methods that account for the ZLB.

The first method estimates the fully nonlinear model with a particle filter (NL-PF). We solve the model with the same algorithm we used to generate our datasets. To filter the data, we follow Algorithm 14 in Herbst and Schorfheide (2016) and adapt the basic bootstrap particle filter described in Fernández-Villaverde and Rubio-Ramírez (2007) to include the information contained in the current observation, so the model better matches extreme outliers in the data. NL-PF is well-equipped to handle the nonlinearities in the data, but it is also the most computationally intensive. NL-PF requires solving the fully nonlinear model and performing a large number of simulations to evaluate the likelihood function for each draw in the random walk Metropolis-Hastings algorithm. Appendix C provides a more detailed description of the estimation algorithm and the particle filter.

The second method estimates a piecewise linear version of the nonlinear model with an inversion filter. To solve the model, we use the OccBin toolbox developed by Guerrieri and Iacoviello (2015). The algorithm separates the model into two regimes. In one regime, the ZLB constraint is slack, and the decision rules from the unconstrained linear model are used. In the other regime, the ZLB binds and backwards induction within a guess and verify method solves for the decision rules. For example, if the ZLB binds in the current period, an initial conjecture is made for how many quarters the nominal interest rate will remain at the ZLB. Starting far enough in the future, the algorithm uses the decision rules for when the ZLB does not bind and iterates backward to the current period. The algorithm switches to the decision rules for the ZLB regime when the simulated nominal interest rate indicates that the ZLB binds. The simulation implies a new guess for the ZLB duration. The algorithm iterates until the implied ZLB duration equals the previous guess.

The advantage of using the piecewise linear model is that it solves very quickly. On average, the nonlinear model takes 3.6 seconds to solve (using Fortran with 16 cores), whereas the piecewise linear model takes a fraction of a second. Furthermore, the nonlinear solution time exponentially increases with the size of the model, whereas the model has little effect on the solution time in the piecewise linear model. However, it is numerically too costly to apply a particle filter. For each particle, the piecewise linear solution requires a long enough simulation to return to the regime where the ZLB does not bind, whereas only a 1-period update is needed with the nonlinear solution. To speed up the filter, Guerrieri and Iacoviello (2017) follow Fair and Taylor (1983) and use an inversion filter that requires only one simulation. The inversion filter solves for the shocks that minimize the distance between the observables and the equivalent model predictions each period.

The piecewise linear model estimated with the inversion filter (OB-IF) makes one potentially important simplifying assumption. Households do not account for the possibility that the ZLB may bind in the future when it does not currently bind. That means households ignore the effects of the ZLB in states of the economy where it is likely to bind in the near future because the algorithm uses the unconstrained linear decision rules. The question is whether this simplification creates large enough differences between the two methods to justify the higher estimation time of NL-PF.

As a benchmark, we estimate the linear analogue of the nonlinear model using Sims's (2002) gensys algorithm to solve the model and a Kalman filter to evaluate the likelihood function (Lin-KF). Unlike the other two methods, this method ignores the ZLB constraint, but it is much easier to implement and was the most common method used in the literature before the Great Recession.

For each estimation method, the observation equation is given by $\mathbf{x}_t = H\mathbf{s}_t + \xi_t$, where $\mathbf{s}_t$ is a vector of variables, H is an observable selection matrix, and ξ is a vector of measurement errors (MEs). The inversion filter solves for the shocks that minimize the distance between the observables, $\mathbf{x}_t$, and their model predictions, $H\mathbf{s}_t$, so there is no ME up to a numerical tolerance. With a Kalman filter or particle filter, $\xi \sim \mathbb{N}(0, R)$, where R is a diagonal matrix of ME variances.[6]

[6] Ireland (2004) allows for correlated MEs, but he finds a real business cycle model's out-of-sample forecasts improve when the ME covariance matrix is diagonal. Guerrón-Quintana (2010) finds that introducing *i.i.d.* MEs and fixing the variances to 10% or 20% of the standard deviation of the data improves the empirical fit and forecasting properties of a New Keynesian model. Fernández-Villaverde and Rubio-Ramírez (2007) estimate the ME variances, but Doh (2011) argues that approach can lead to complications because the ME variances are similar to bandwidths in nonparameteric estimation. Given those findings, we use a diagonal ME covariance matrix and fix the ME variances.

We are free to set the ME variances to zero when we use the Kalman filter, since the number of observables is equal to the number of shocks. The particle filter, however, always requires positive ME variances to avoid degeneracy. Unfortunately, there is no consensus on how to set these values, despite their potentially large effect. We consider three values for the ME variances: 2%, 5%, and 10% of the variance in the data. These values encompass the range of values used in the literature.[7]

Parameter	Dist.	Mean (SD)	Parameter	Dist.	Mean (SD)	Parameter	Dist.	Mean (SD)
φ_p	Norm	100 (25)	h	Beta	0.8 (0.1)	σ_z	IGam	0.005 (0.005)
ϕ_π	Norm	2.0 (0.25)	ρ_s	Beta	0.8 (0.1)	σ_s	IGam	0.005 (0.005)
ϕ_y	Norm	0.5 (0.25)	ρ_i	Beta	0.8 (0.1)	σ_i	IGam	0.002 (0.002)

Table 2: Prior distributions, means, and standard deviations of the estimated parameters.

Table 2 displays information about the prior distributions of the estimated parameters. All other parameter values are fixed at their true values. The prior means are set to the true parameter values to isolate the influence of other aspects of the estimation procedure, such as the solution method and filter. Different prior means would most likely affect the accuracy of the estimation and contaminate our results. The prior standard deviations, which are consistent with the values in the literature, are relatively diffuse to give the algorithm flexibility to search the parameter space.

Our estimation procedure has three stages. First, we conduct a mode search to create an initial variance-covariance matrix for the estimated parameters. The covariance matrix is based on the parameters corresponding to the 90th percentile of the likelihoods from $5{,}000$ draws. Second, we perform an initial run of the Metropolis-Hastings algorithm with $25{,}000$ draws from the posterior distribution. We burn off the first $5{,}000$ draws and use the remaining draws to update the variance-covariance matrix from the mode search. Third, we conduct a final run of the Metropolis-Hastings algorithm. We obtain $50{,}000$ draws from the posterior distribution and then record the mean draw.

The algorithm is programmed in Fortran and the datasets are run in parallel across several supercomputers. Each dataset uses one core with OB-IF and Lin-KF, whereas NL-PF uses 16 cores because the solution is parallelized. For example, a supercomputer with 80 cores can simultaneously run 80 datasets with OB-IF but only 5 datasets with NL-PF. To increase the accuracy of the particle filter, we evaluate the likelihood function on each core. Since NL-PF uses 16 cores, we obtain 16 likelihoods and determine whether to accept a draw based on the median likelihood. This key step reduces the variance of the likelihoods from seed effects. The filter uses $40{,}000$ particles.

[7] Some papers set the ME *standard deviations* to 20% or 25% of the sample standard deviations, which is equivalent to setting the ME *variances* to 4% or 6.25% of the sample variances (e.g., An and Schorfheide, 2007; Doh, 2011; Herbst and Schorfheide, 2016; van Binsbergen et al., 2012). Other work directly sets the ME variances to 10% or 25% of the sample variances (e.g., Bocola, 2016; Gust et al., 2017; Plante et al., 2018; Richter and Throckmorton, 2016).

	NL-PF (16 Cores)		OB-IF (1 Core)		Lin-KF (1 Core)	
	0Q	30Q	0Q	30Q	0Q	30Q
Seconds per draw	6.7 $(6.1, 7.9)$	8.4 $(7.5, 9.5)$	0.035 $(0.031, 0.040)$	0.096 $(0.051, 0.135)$	0.002 $(0.002, 0.004)$	0.002 $(0.001, 0.003)$
Hours per dataset	148.8 $(134.9, 176.5)$	186.4 $(167.6, 210.7)$	0.781 $(0.689, 0.889)$	2.137 $(1.133, 3.000)$	0.052 $(0.044, 0.089)$	0.049 $(0.022, 0.067)$

Table 3: Average and $(5, 95)$ percentiles of the estimation times by method and ZLB duration in the data.

Table 3 shows the computing times for each estimation method. We first report the average and $(5, 95)$ percentiles of the combined solution and filter times across our 50 posterior mean estimates. These draws are independent and representative of other draws from the posterior distribution. We then show hours per dataset, which are extrapolated by multiplying seconds per draw by $80{,}000$ draws and dividing by $3{,}600$ seconds per hour. We report times for NL-PF, OB-IF, and Lin-KF in datasets where the ZLB never binds and datasets with one 30 quarter ZLB event. NL-PF is run on 16 cores and the other methods use a single core. The estimation times depend on the hardware, but there are two interesting takeaways. One, OB-IF is slightly slower than Lin-KF, but it only takes a few hours to run on a single core. Two, NL-PF requires significantly more time than OB-IF, but it ran in about a week with 16 cores, so it is possible to estimate the nonlinear model on a workstation.

4 Posterior Estimates and Accuracy

The section begins by showing the accuracy of the parameter estimates for each estimation method. We then compare the filtered estimates of the notional interest rate, expected frequency and duration of the ZLB, responses to a severe recession, and forecasting performance across the methods.

4.1 Parameter Estimates

We measure parameter accuracy by calculating the normalized root-mean square-error (NRMSE) for each estimated parameter. For parameter j and estimation method h, the error is the difference between the parameter estimate for dataset k, $\hat{\theta}_{j,h,k}$, and the true parameter, $\tilde{\theta}_j$. Therefore, the NRMSE for parameter j and estimation method h is given by

$$\mathrm{NRMSE}_h^j = \frac{1}{\tilde{\theta}_j}\sqrt{\frac{1}{N}\sum_{k=1}^{N}(\hat{\theta}_{j,h,k} - \tilde{\theta}_j)^2},$$

where N is the number of datasets. The RMSE is normalized by $\tilde{\theta}_j$ to remove differences in the scales of the parameters and measure the total error. We also compute the coverage ratio given by

$$\mathrm{CR}_h^j = \frac{1}{N}\sum_{k=1}^{N}\mathbb{I}(\hat{\theta}_{j,h,k}^{5\%} < \tilde{\theta}_j) \times \mathbb{I}(\hat{\theta}_{j,h,k}^{95\%} > \tilde{\theta}_j),$$

where $\mathbb{I}$ is an indicator function and $\hat{\theta}^{X\%}$ denodes the Xth percentile of the posterior distribution. This statistic shows how likely it is for the posterior distribution to contain the true parameter value.

Ptr	Truth	NL-PF-5%		OB-IF-0%		Lin-KF-5%	
		0Q	30Q	0Q	30Q	0Q	30Q
φ_p	100	151.1 (134.2, 165.8) {0.52, 0.02}	188.4 (174.7, 202.7) {0.89, 0.00}	142.6 (121.1, 157.3) {0.44, 0.08}	183.4 (169.2, 198.5) {0.84, 0.00}	151.4 (134.0, 165.7) {0.52, 0.00}	191.6 (175.3, 204.1) {0.92, 0.00}
h	0.8	0.66 (0.62, 0.70) {0.18, 0.00}	0.68 (0.64, 0.71) {0.16, 0.00}	0.64 (0.61, 0.67) {0.20, 0.00}	0.63 (0.60, 0.67) {0.21, 0.00}	0.66 (0.62, 0.69) {0.18, 0.00}	0.67 (0.63, 0.70) {0.17, 0.00}
ρ_s	0.8	0.76 (0.72, 0.80) {0.06, 0.70}	0.81 (0.78, 0.84) {0.03, 0.90}	0.76 (0.73, 0.81) {0.05, 0.82}	0.82 (0.79, 0.86) {0.04, 0.78}	0.76 (0.72, 0.80) {0.06, 0.74}	0.82 (0.78, 0.86) {0.04, 0.78}
ρ_i	0.8	0.79 (0.75, 0.82) {0.03, 0.96}	0.80 (0.75, 0.84) {0.03, 0.96}	0.76 (0.71, 0.79) {0.06, 0.52}	0.77 (0.73, 0.81) {0.05, 0.66}	0.79 (0.75, 0.82) {0.03, 0.98}	0.84 (0.80, 0.88) {0.06, 0.56}
σ_z	0.005	0.0032 (0.0023, 0.0039) {0.37, 0.00}	0.0040 (0.0030, 0.0052) {0.23, 0.58}	0.0051 (0.0044, 0.0058) {0.09, 0.92}	0.0059 (0.0050, 0.0069) {0.22, 0.30}	0.0032 (0.0023, 0.0039) {0.36, 0.00}	0.0043 (0.0030, 0.0057) {0.20, 0.68}
σ_s	0.005	0.0052 (0.0040, 0.0066) {0.15, 0.92}	0.0050 (0.0039, 0.0062) {0.13, 0.96}	0.0051 (0.0042, 0.0063) {0.13, 0.92}	0.0046 (0.0036, 0.0056) {0.15, 0.82}	0.0053 (0.0040, 0.0067) {0.15, 0.92}	0.0047 (0.0037, 0.0061) {0.15, 0.92}
σ_i	0.002	0.0017 (0.0014, 0.0020) {0.17, 0.48}	0.0015 (0.0013, 0.0019) {0.24, 0.20}	0.0020 (0.0018, 0.0023) {0.08, 0.90}	0.0020 (0.0019, 0.0024) {0.09, 0.90}	0.0017 (0.0015, 0.0020) {0.16, 0.50}	0.0016 (0.0014, 0.0019) {0.20, 0.28}
ϕ_π	2.0	2.04 (1.88, 2.19) {0.06, 0.98}	2.13 (1.94, 2.31) {0.09, 0.92}	2.01 (1.84, 2.16) {0.06, 0.98}	1.96 (1.77, 2.14) {0.06, 0.98}	2.04 (1.88, 2.20) {0.06, 0.98}	1.73 (1.52, 1.91) {0.15, 0.78}
ϕ_y	0.5	0.35 (0.21, 0.54) {0.36, 0.80}	0.42 (0.27, 0.62) {0.28, 0.98}	0.32 (0.17, 0.48) {0.41, 0.68}	0.44 (0.27, 0.61) {0.25, 0.98}	0.35 (0.22, 0.54) {0.35, 0.80}	0.32 (0.17, 0.47) {0.40, 0.76}
Σ		1.90	2.08	1.53	1.91	1.88	2.28

Table 4: Average, $(5, 95)$ percentiles, and $\{\mathrm{NRMSE}, \mathrm{CR}\}$. Σ is the sum of the NRMSE across the parameters.

Table 4 shows the parameter estimates by specification (first column header) and the duration of the ZLB (second column header). The percentage appended to each specification header corresponds to the size of the ME variances. Each cell includes the average (first row), $(5, 95)$ percentiles (second row), NRMSE (third row, first value), and the coverage ratio (third row, second value).[8]

Across all specifications, the Rotemberg price adjustment cost parameter (φ_p) has the highest NRMSE and it becomes less accurate when the ZLB binds in the data. The upward bias is driven by misspecification, since the small-scale model used for estimation does not include sticky-wages. In the small-scale model, the response of marginal costs to shocks is much larger than in the medium-scale model, so the estimates of φ_p are higher than the true value to flatten the Phillips curve. Another inaccuracy is a downward bias in the estimates of habit persistence (h). The response of output growth to shocks is too small due to the lack of investment in the small-scale model. Lowering h increases the response to shocks, although at the expense of lower persistence. Risk premium persistence (ρ_s) and the monetary response to the output growth gap (ϕ_y) also have a downward bias in the datasets without a ZLB event, but the CR is much higher than the near-zero

[8]For conciseness, we focus on datasets without a ZLB event and those with a 30 quarter event, but the estimates for the datasets with intermediate ZLB durations, as well as the Lin-KF-0% estimates, are provided in Appendix E.2.

values for φ_p and h. Also, the bias of ρ_s and ϕ_y decreases using datasets with a 30 quarter event.

The NL-PF-5% estimates of the productivity growth and monetary policy shock standard deviations (σ_z and σ_i) are biased downward, while the OB-IF-0% estimates are roughly consistent with their true values. In the datasets without a ZLB event, Lin-KF-5% produces identical estimates to NL-PF-5%, suggesting the bias is due to the positive ME variances in the filter. The importance of the ME variances is likely driven by the filter ascribing large shocks to ME rather than the structural shocks, reducing their estimated volatility. However, in datasets with a 30 quarter event, NL-PF-5% is more likely to contain the true risk premium parameters (ρ_s and σ_s) than OB-IF-0%. While the average estimates are similar, the CR is 0.90 for ρ_s with NL-PF-5%, compared to 0.78 with OB-IF-0%. For σ_s the CRs are 0.96 with NL-PF-5% and 0.82 with OB-IF-0%. This is notable because these two parameters have the largest effect on the frequency and duration of ZLB events.

The bottom row of table 4 shows the sum of the NRMSE across the parameters. These values provide an aggregate measure of parameter accuracy. In the datasets that are not influenced by the ZLB, OB-IF-0% is more accurate than NL-PF-5%. The results for Lin-KF-5% show the lower accuracy of NL-PF-5% is driven by positive ME variances and that the ZLB is the only important nonlinearity in the model. When the ZLB binds, it reduces the accuracy of every specification, largely due to a single parameter, φ_p.[9] Long ZLB events have the smallest effect on the accuracy of NL-PF-5%. Datasets with a 30 quarter ZLB event reduce accuracy by 0.18 relative to datasets without a ZLB event. For comparison, the accuracy decreases by 0.38 with OB-IF-0% and by 0.30 with Lin-KF-0%. However, NL-PF-5% is less accurate than OB-IF-0% due to the positive ME variances. In other words, NL-PF-5% is the best equipped to handle ZLB events in the data, but the loss in accuracy from the positive ME variances in the particle filter may outweigh those benefits.

Misspecification The absence of sticky wages and other frictions from the data generating process are important drivers of the parameter estimates in the small-scale model. Here we explore the effect of misspecification on only the OB-IF estimates since adding sticky wages substantially increases the computational cost of NL-PF. The first two columns of table 5 repeat the OB-IF-0% estimates of the small-scale model, while the middle columns show the effect of reducing misspecification on the OB-IF-0% estimates by including sticky wages.[10] The right two columns show the OB-IF-0% estimates using the medium-scale model that generates the data, eliminating all misspecification except nonlinearities not captured by the OccBin solution. For the last two cases, we fix the parameters that are not estimated in the small-scale model to their true values.[11]

In datasets with a 30 quarter ZLB event, adding sticky wages reduces the sum of the NRMSE

[9] Appendix E.3 shows there is a small but similar decrease in accuracy due to φ_p when there is no misspecification.

[10] The equilibrium system is the same as the small-scale model, except (43) and (44) are replaced with (28), (32), (40), and a real GDP definition that accounts for sticky wages (i.e., $\tilde{y}_t^{gdp} = [1-\varphi_p(\pi_t/\bar{\pi}-1)^2/2-\varphi_w(w_t^g-1)^2/2]\tilde{y}_t$).

[11] Appendix E.1 further explores the estimated bias by reproducing table 4 with φ_p and h fixed at their true values.

Ptr	Truth	OB-IF-0%		OB-IF-0%-Sticky Wages		OB-IF-0%-DGP	
		0Q	30Q	0Q	30Q	0Q	30Q
φ_p	100	142.6 (121.1, 157.3) {0.44, 0.08}	183.4 (169.2, 198.5) {0.84, 0.00}	100.1 (76.9, 119.6) {0.13, 1.00}	129.8 (105.5, 152.3) {0.33, 0.58}	101.4 (80.1, 120.7) {0.12, 0.98}	128.4 (109.0, 148.1) {0.31, 0.46}
h	0.8	0.64 (0.61, 0.67) {0.20, 0.00}	0.63 (0.60, 0.67) {0.21, 0.00}	0.82 (0.78, 0.86) {0.04, 0.82}	0.80 (0.77, 0.85) {0.03, 0.88}	0.81 (0.75, 0.85) {0.04, 1.00}	0.77 (0.72, 0.84) {0.06, 0.78}
ρ_s	0.8	0.76 (0.73, 0.81) {0.05, 0.82}	0.82 (0.79, 0.86) {0.04, 0.78}	0.82 (0.76, 0.86) {0.04, 0.90}	0.84 (0.80, 0.88) {0.06, 0.58}	0.80 (0.76, 0.85) {0.03, 0.96}	0.82 (0.79, 0.86) {0.04, 0.80}
ρ_i	0.8	0.76 (0.71, 0.79) {0.06, 0.52}	0.77 (0.73, 0.81) {0.05, 0.66}	0.80 (0.77, 0.83) {0.02, 0.98}	0.80 (0.77, 0.84) {0.03, 0.92}	0.79 (0.75, 0.82) {0.03, 0.98}	0.79 (0.75, 0.83) {0.03, 0.92}
σ_z	0.005	0.0051 (0.0044, 0.0058) {0.09, 0.92}	0.0059 (0.0050, 0.0069) {0.22, 0.30}	0.0038 (0.0031, 0.0044) {0.24, 0.16}	0.0047 (0.0039, 0.0055) {0.12, 0.72}	0.0047 (0.0039, 0.0054) {0.11, 0.78}	0.0055 (0.0047, 0.0066) {0.15, 0.70}
σ_s	0.005	0.0051 (0.0042, 0.0063) {0.13, 0.92}	0.0046 (0.0036, 0.0056) {0.15, 0.82}	0.0085 (0.0056, 0.0134) {0.81, 0.44}	0.0074 (0.0050, 0.0107) {0.60, 0.58}	0.0060 (0.0043, 0.0084) {0.30, 0.88}	0.0051 (0.0039, 0.0068) {0.19, 0.92}
σ_i	0.002	0.0020 (0.0018, 0.0023) {0.08, 0.90}	0.0020 (0.0019, 0.0024) {0.09, 0.90}	0.0020 (0.0018, 0.0022) {0.08, 0.84}	0.0020 (0.0018, 0.0023) {0.08, 0.92}	0.0020 (0.0018, 0.0022) {0.08, 0.92}	0.0020 (0.0018, 0.0024) {0.09, 0.88}
ϕ_π	2.0	2.01 (1.84, 2.16) {0.06, 0.98}	1.96 (1.77, 2.14) {0.06, 0.98}	1.91 (1.74, 2.04) {0.07, 1.00}	1.81 (1.63, 1.99) {0.11, 0.72}	1.92 (1.72, 2.08) {0.06, 1.00}	1.81 (1.62, 2.03) {0.11, 0.70}
ϕ_y	0.5	0.32 (0.17, 0.48) {0.41, 0.68}	0.44 (0.27, 0.61) {0.25, 0.98}	0.40 (0.24, 0.58) {0.28, 0.96}	0.50 (0.33, 0.73) {0.23, 0.98}	0.41 (0.24, 0.57) {0.26, 0.96}	0.50 (0.32, 0.74) {0.24, 0.96}
Σ		1.53	1.91	1.71	1.59	1.03	1.23

Table 5: Average, $(5, 95)$ percentiles, and $\{\mathrm{NRMSE}, \mathrm{CR}\}$. Σ is the sum of the NRMSE across the parameters.

from 1.91 to 1.59. That is a clear improvement over NL-PF-5% and is driven by more accurate estimates of φ_p and h that dominate the lower accuracy of σ_s. The CR for φ_p and h also significantly increases. This is consistent with the claim that the bias of φ_p and h in table 4 is primarily due to the lack of sticky wages, which destabilize marginal costs and inflation. The amplification of shocks still remains too low, now for both inflation and output, which leads to an upward bias in σ_s rather than a downward bias in h. The NRMSE for σ_s is much higher and the CR declines.

Making the estimated model consistent with the DGP improves the parameter estimates even further. The sum of the NRMSE declines 1.59 to 1.23 when the ZLB binds for 30 quarters. The primary reason is because σ_s is closer to its true value. The NRMSE in σ_s is significantly lower and the CR is much higher. Also, all of the true parameter values are encompassed by the $(5, 95)$ percentiles of the estimates, except the estimate of φ_p has a large upward bias in the 30 quarter datasets. This indicates the increase in the bias of φ_p as the ZLB duration increases is solely driven by sample selection, not model misspecification. Overall, our results suggest it is more beneficial to reduce misspecification and estimate a richer model with OB-IF than a smaller model with NL-PF. Nonlinear methods more accurately capture the dynamics of the ZLB, but computational limitations often require excluding important features of the model, like sticky wages and capital.

Ptr	Truth	NL-PF-2%		NL-PF-5%		NL-PF-10%	
		0Q	30Q	0Q	30Q	0Q	30Q
φ_p	100	150.2 $(133.5, 165.3)$ $\{0.51, 0.02\}$	192.0 $(176.5, 207.1)$ $\{0.93, 0.00\}$	151.1 $(134.2, 165.8)$ $\{0.52, 0.02\}$	188.4 $(174.7, 202.7)$ $\{0.89, 0.00\}$	149.5 $(132.6, 163.8)$ $\{0.50, 0.02\}$	182.7 $(168.6, 197.3)$ $\{0.83, 0.02\}$
h	0.8	0.66 $(0.62, 0.69)$ $\{0.18, 0.00\}$	0.67 $(0.64, 0.71)$ $\{0.17, 0.00\}$	0.66 $(0.62, 0.70)$ $\{0.18, 0.00\}$	0.68 $(0.64, 0.71)$ $\{0.16, 0.00\}$	0.66 $(0.61, 0.70)$ $\{0.17, 0.00\}$	0.68 $(0.65, 0.72)$ $\{0.15, 0.00\}$
ρ_s	0.8	0.76 $(0.71, 0.79)$ $\{0.06, 0.60\}$	0.81 $(0.78, 0.84)$ $\{0.03, 0.92\}$	0.76 $(0.72, 0.80)$ $\{0.06, 0.70\}$	0.81 $(0.78, 0.84)$ $\{0.03, 0.90\}$	0.76 $(0.72, 0.79)$ $\{0.06, 0.76\}$	0.81 $(0.79, 0.85)$ $\{0.03, 0.88\}$
ρ_i	0.8	0.77 $(0.73, 0.80)$ $\{0.05, 0.76\}$	0.79 $(0.75, 0.83)$ $\{0.03, 0.96\}$	0.79 $(0.75, 0.82)$ $\{0.03, 0.96\}$	0.80 $(0.75, 0.84)$ $\{0.03, 0.96\}$	0.80 $(0.77, 0.84)$ $\{0.03, 0.96\}$	0.81 $(0.76, 0.85)$ $\{0.03, 0.94\}$
σ_z	0.005	0.0038 $(0.0031, 0.0043)$ $\{0.25, 0.16\}$	0.0043 $(0.0035, 0.0052)$ $\{0.18, 0.60\}$	0.0032 $(0.0023, 0.0039)$ $\{0.37, 0.00\}$	0.0040 $(0.0030, 0.0052)$ $\{0.23, 0.58\}$	0.0027 $(0.0020, 0.0035)$ $\{0.46, 0.00\}$	0.0038 $(0.0025, 0.0050)$ $\{0.28, 0.62\}$
σ_s	0.005	0.0052 $(0.0039, 0.0065)$ $\{0.15, 0.88\}$	0.0051 $(0.0040, 0.0061)$ $\{0.13, 0.92\}$	0.0052 $(0.0040, 0.0066)$ $\{0.15, 0.92\}$	0.0050 $(0.0039, 0.0062)$ $\{0.13, 0.96\}$	0.0051 $(0.0041, 0.0065)$ $\{0.14, 0.94\}$	0.0049 $(0.0037, 0.0061)$ $\{0.14, 0.92\}$
σ_i	0.002	0.0019 $(0.0017, 0.0021)$ $\{0.10, 0.70\}$	0.0018 $(0.0016, 0.0021)$ $\{0.14, 0.62\}$	0.0017 $(0.0014, 0.0020)$ $\{0.17, 0.48\}$	0.0015 $(0.0013, 0.0019)$ $\{0.24, 0.20\}$	0.0015 $(0.0012, 0.0018)$ $\{0.25, 0.28\}$	0.0013 $(0.0011, 0.0017)$ $\{0.34, 0.12\}$
ϕ_π	2.0	2.01 $(1.84, 2.16)$ $\{0.06, 0.98\}$	2.14 $(1.96, 2.31)$ $\{0.09, 0.90\}$	2.04 $(1.88, 2.19)$ $\{0.06, 0.98\}$	2.13 $(1.94, 2.31)$ $\{0.09, 0.92\}$	2.06 $(1.89, 2.21)$ $\{0.07, 0.98\}$	2.12 $(1.92, 2.28)$ $\{0.08, 0.96\}$
ϕ_y	0.5	0.31 $(0.18, 0.48)$ $\{0.42, 0.64\}$	0.39 $(0.24, 0.60)$ $\{0.32, 0.92\}$	0.35 $(0.21, 0.54)$ $\{0.36, 0.80\}$	0.42 $(0.27, 0.62)$ $\{0.28, 0.98\}$	0.41 $(0.26, 0.59)$ $\{0.27, 0.98\}$	0.46 $(0.30, 0.66)$ $\{0.24, 1.00\}$
Σ		1.79	2.01	1.90	2.08	1.95	2.13

Table 6: Average, $(5, 95)$ percentiles, and $\{\mathrm{NRMSE}, \mathrm{CR}\}$. Σ is the sum of the NRMSE across the parameters.

ME Variances Table 6 shows the parameter estimates and NRMSEs for NL-PF with three different ME variances: 2%, 5% (baseline), and 10%. Without model misspecification, lowering the ME variances would increase the accuracy of the parameter estimates as long as the effective sample of particles is large enough. In our setup, the presence of misspecification creates a potential tradeoff. On the one hand, lower ME variances force the model to match sharp swings in the data, which could help identify the parameters. On the other hand, higher ME variances give the model a degree of freedom to account for important differences between the estimated model and the DGP.

We find higher ME variances increase the sum of the NRMSE. In datasets with 30 quarter ZLB events, it increases from 2.01 to 2.13 when the ME variances increase from 2% to 10%. For σ_z and σ_i, higher ME variances push the estimates lower, away from their true values. Once again, this result is likely driven by the filter incorrectly ascribing movements in the data to ME rather than the structural shocks. This loss in accuracy as the ME variances increase is partially offset by the increase in the accuracy of most other parameters. Estimates of ϕ_y with all datasets and estimates of φ_p with datasets where the ZLB binds for 30 quarters improve the most. These results show that ME variances are important for accuracy. In some cases, they may compensate for model misspecification. In our setup, however, larger ME variances have a net negative effect on accuracy.

4.2 NOTIONAL INTEREST RATE ESTIMATES We measure the accuracy of the notional rate by calculating the average RMSE across periods when the ZLB binds. For period t and estimation method h, the error is the difference between the filtered notional rate based on the parameter estimates for dataset k, $\hat{i}^n_{t,h,k}$, and the true notional rate, $\tilde{i}^n_t$. The RMSE for method h is given by

$$\mathrm{RMSE}^{i^n}_h = \sqrt{\tfrac{1}{N}\tfrac{1}{\tau}\textstyle\sum_{k=1}^{N}\sum_{j=t}^{t+\tau-1}(\hat{i}^n_{j,h,k}-\tilde{i}^n_j)^2},$$

where t is the first period the ZLB binds and τ is the duration of the ZLB event. There is no reason to normalize the RMSE since the units are the same across periods and we do not sum across states.

Estimates of the notional interest rate are of keen interest to policymakers for two key reasons. One, they summarize the severity of the recession and the nominal interest rate policymakers would like to set in the absence of the ZLB, which help inform decisions about implementing unconventional monetary policy. Two, estimates of the notional rate help determine how long the ZLB is expected to bind, which is necessary to issue forward guidance. The notional rate is also the only latent endogenous state variable in the model that is not directly linked to an observable.

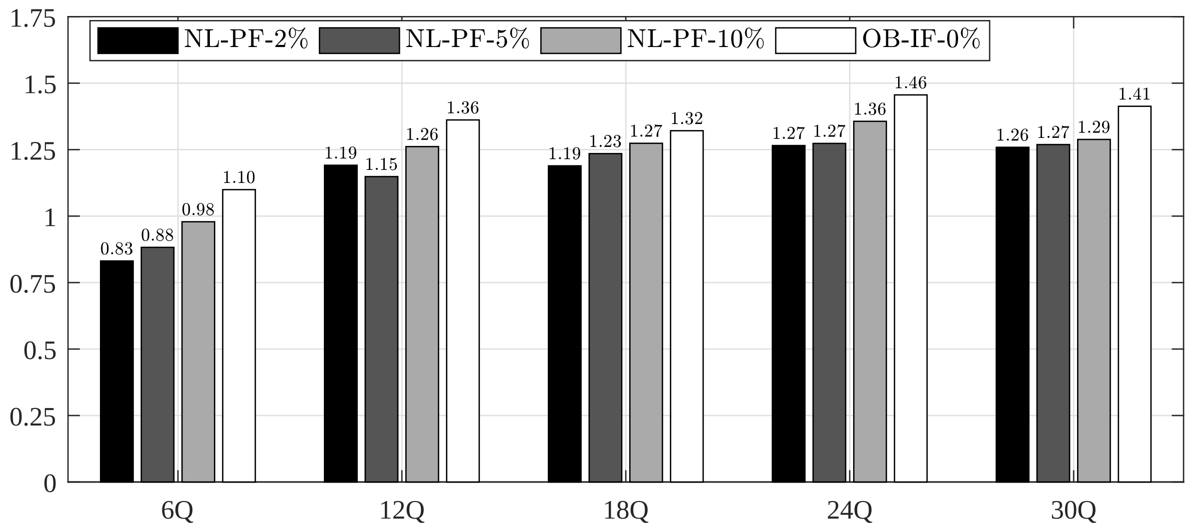


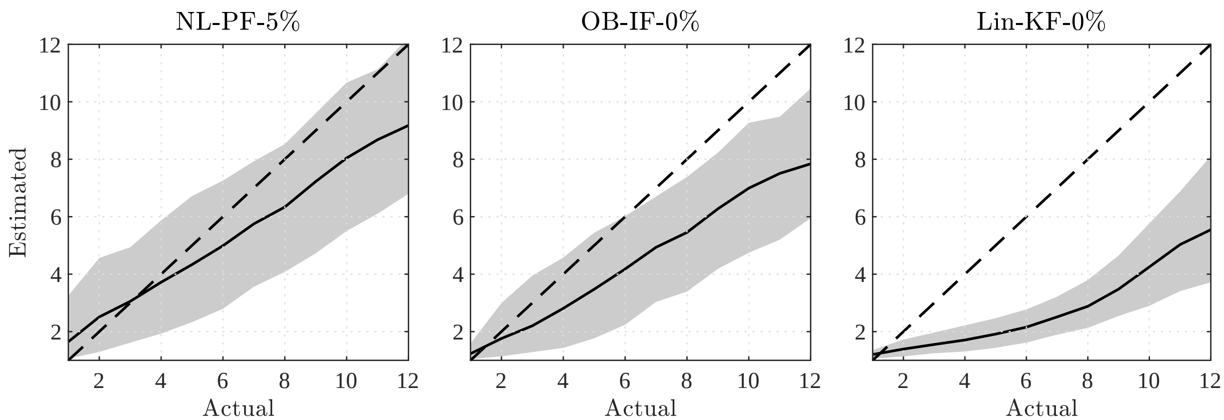
Figure 1: RMSE of the notional interest rate across ZLB durations in the data. Rates are net annualized percentages.

Figure 1 shows the accuracy of the notional rate for our baseline methods, NL-PF-5% and OB-IF-0%. We also show the how different ME variances in the particle filter affect accuracy. We do not present the results for Lin-KF because they are uninformative. Since the linear model does not distinguish between the notional and nominal rates and the nominal rate is an observable, the error in the linear model equals the absolute value of the notional rate when the ZLB binds in the data.

Regardless of the ZLB duration, NL-PF-5% provides more accurate estimates of the notional rate than OB-IF-0%. Depending on the ZLB duration, the average difference between the two methods ranges from 0.1 to 0.25 annualized percentage points. Appendix E.5 shows the RMSE is higher with OB-IF-0% because the estimate of the notional rate is more likely to be above the true value. Nevertheless, the differences in the estimates are not big enough to have a meaningful impact on policy prescriptions. Increasing or decreasing the ME variances also has a modest effect.

4.3 EXPECTED ZLB DURATION AND PROBABILITY In addition to estimates of the notional interest rate, two commonly referenced statistics in the literature are the expected duration and probability of the ZLB constraint. These statistics determine the impact of a ZLB event in the model and are frequently measured against survey data. Figure 2 shows the accuracy of the two statistics.

(a) Estimated vs. Actual Expected ZLB Durations



(b) Estimated vs. Actual Probability of a 4 Quarter or Longer ZLB Event

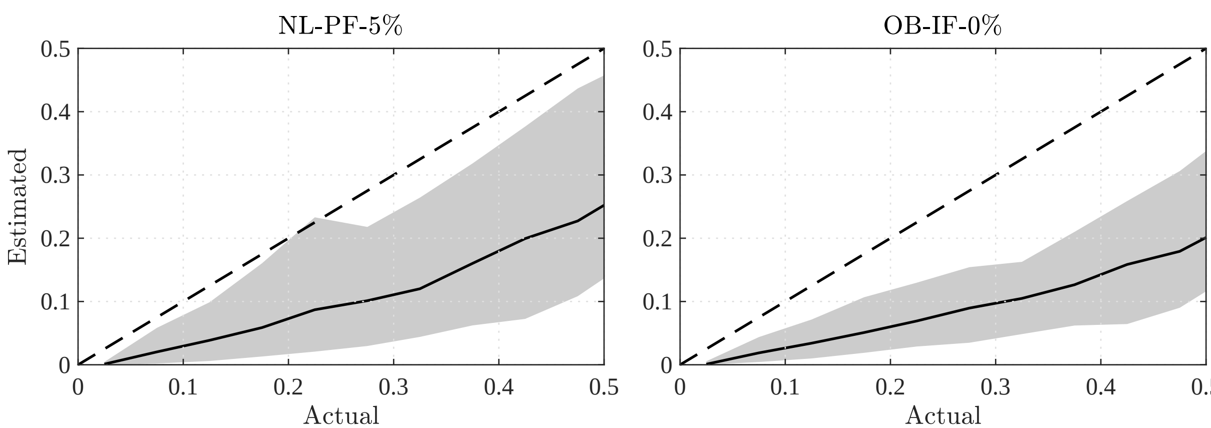


Figure 2: Estimated and actual ZLB statistics. The solid lines are mean estimates and the shaded areas capture the $(5, 95)$ percentiles across the datasets. The dashed line shows where the estimated values would equal the actual values.

The top panel compares the expected ZLB durations given the parameter estimates from the small-scale model to the actual expected ZLB durations from the DGP given the true parameters. The expected ZLB durations are computed as the average across $10{,}000$ simulations of a model initialized at the filtered states (or actual states for the DGP) where the ZLB binds. The solid lines are the mean expected ZLB durations in the small-scale model after pooling across the different ZLB states and datasets. The shaded areas are the $(5, 95)$ percentiles of the durations. The estimated expected ZLB duration equals the actual expected ZLB duration along the dashed 45 degree line.

When the actual expected ZLB duration is relatively short, the NL-PF-5% and OB-IF-0% expected ZLB durations are close to the truth. As the actual expected duration lengthens, both estimates become less accurate. The NL-PF-5% 95th percentile continues to encompass the actual expected durations. However, once the actual value exceeds six quarters, there is a 95% chance or higher of under-estimating the actual expected duration with OB-IF-0%. Furthermore, the OB-IF-0% mean expected duration is typically at least one quarter shorter than the NL-PF-5% mean estimate.[12] These results are likely driven by model misspecification, as the presence of capital and sticky wages in the DGP makes the ZLB more persistent than in the estimated small-scale model.

The Lin-KF-0% estimated ZLB durations are always significantly shorter since that method does not permit a negative notional rate when filtering the data. The only instance when Lin-KF-0% produces an expected ZLB duration beyond one year is when the economy is in a severe downturn and the actual expected duration is extremely long. The Lin-KF-0% estimates are a lower bound on the OB-IF-0% estimates since the solutions are identical when the ZLB does not bind.

The bottom panel is constructed in a similar way as the top panel except the horizontal and vertical axes correspond to the actual and estimated probability of a ZLB event that lasts for at least four quarters. The probability is calculated in all periods where the ZLB does not bind in the data. We do not show the results for Lin-KF-0% because the probability of a four quarter ZLB event is always near zero. NL-PF-5% and OB-IF-0% underestimate the true probability, but the mean NL-PF-5% estimates are slightly closer to the actual probabilities and the 95th percentile almost encompasses the truth. Changing the ME variances in the particle filter has no discernable effect on the estimates. These results illustrate the precautionary savings effects of the ZLB, which are not captured by OB-IF-0%. However, they do not provide overwhelming support for NL-PF-5%.

4.4 RECESSION RESPONSES

To illustrate the economic implications of the differences in accuracy, we compare simulations of the small-scale model given our parameter estimates to simulations of the DGP given the true parameters. The simulations are initialized in steady state and followed by four consecutive 1.5 standard deviation positive risk premium shocks, which generates a 10 quarter ZLB event in the DGP.[13] A risk premium shock is a proxy for a change in demand because it affects households' consumption and saving decisions. Positive shocks cause households to postpone consumption, which reduces current output growth. We focus on this particular shock because it is the primary mechanism for generating ZLB events in the DGP and estimated model.[14]

[12] Prior to instituting date-based forward guidance in 2011, Blue Chip consensus forecasts revealed that people expected the ZLB to bind for three quarters or less. After the forward guidance, the expectation rose to seven quarters.

[13] The simulations are reflective of the Great Recession. The current Congressional Budget Office estimate of the output gap in 2009Q2 is -5.9%, roughly equivalent to the output (level) gap in the true simulation in the fourth period.

[14] Appendix E.4 shows impulse responses to a productivity growth and monetary policy shock in a severe recession.

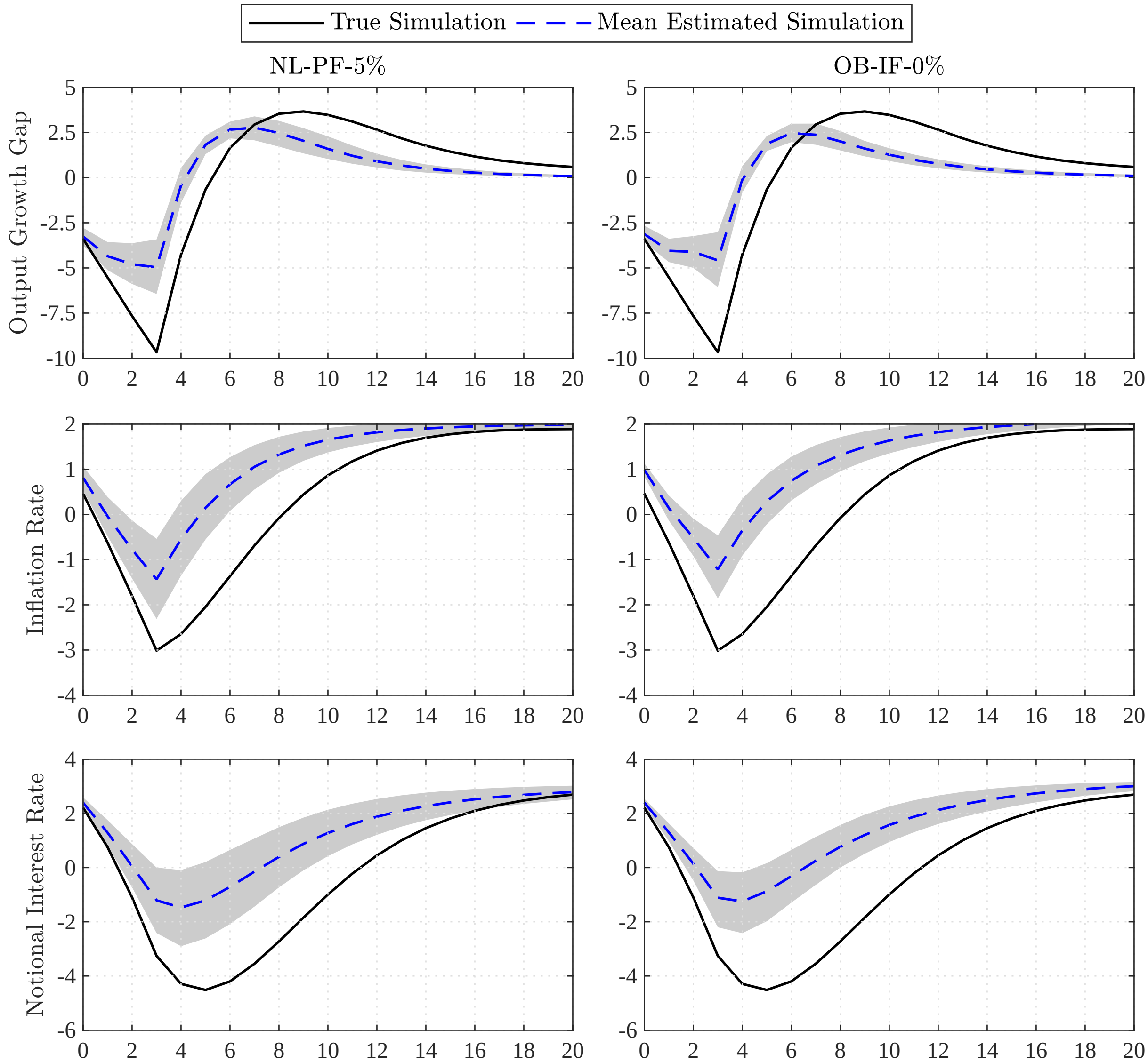


Figure 3: Recession responses. The solid line is the true simulation, the dashed line is the mean estimated simulation, and the shaded area contains the $(5, 95)$ percentiles across the datasets. The simulations are initialized in steady state and followed by four 1.5 standard deviation positive risk premium shocks. All values are net annualized percentages.

Figure 3 shows the simulated paths of the output growth gap, inflation rate, and notional interest rate in annualized net percentages. The NL-PF-5% simulations are shown in the left column and the OB-IF-0% simulations are in the right column. The true simulation of the DGP (solid line) is compared to the mean estimated simulation of the small-scale model (dashed line). The $(5, 95)$ percentiles account for differences in the simulations across the parameter estimates for each dataset.

Model misspecification leads to significantly muted responses relative to the true simulation.[15] None of the estimated simulations for NL-PF-5% or OB-IF-0% can replicate the size of the negative output growth gap, decline in inflation, or policy response at the beginning of the true simu-

[15] Appendix E.3 reproduces the responses without misspecification to confirm it is the source of the muted responses.

lation. Both estimation methods also underestimate the duration of the ZLB event. However, the NL-PF-5% mean simulations of the three variables and the ZLB duration are closer to the truth than the OB-IF-0% simulations. Unlike OccBin, the fully nonlinear solution captures the expectational effects of going to the ZLB, which puts downward pressure on output and inflation and improves accuracy. Although NL-PF-5% is closer to the truth than OB-IF-0%, once again these differences are fairly small and may not justify the significantly longer estimation time.

4.5 FORECAST PERFORMANCE Another important aspect of any model is its ability to forecast. We examine the forecasting performance of each estimation method in the quarter immediately preceding a severe recession that causes the ZLB to bind. The point forecasts are inaccurate since severe recessions are rare. However, there are potentially important differences between the forecast distributions, which assign probabilities to the range of potential outcomes in a given period. The tails of the distribution are particularly important. To measure the accuracy of the forecast distribution of variable j, we compute the continuous rank probability score (CRPS) given by

$$\mathrm{CRPS}^j_{m,k,t,\tau} = \int_{-\infty}^{\tilde{\jmath}_{t+\tau}} [F_{m,k,t}(j_{t+\tau})]^2 dj_{t+\tau} + \int_{\tilde{\jmath}_{t+\tau}}^{\infty} [1 - F_{m,k,t}(j_{t+\tau})]^2 dj_{t+\tau},$$

where m indicates whether the forecast distribution comes from the DGP or an estimated model, k is the dataset, t is the forecast date, $F_{m,k,t}(j_{t+\tau})$ is the cumulative distribution function (CDF) of the τ-quarter ahead forecast, and $\tilde{\jmath}_{t+\tau}$ is the true realization. The CRPS measures the accuracy of the forecast distribution by penalizing probabilities assigned to outcomes that are not realized. It also has the same units as the forecasted variables, which are net percentages, and reduces to the mean absolute error if the forecast is deterministic. A smaller CRPS indicates a more accurate forecast.[16]

For each dataset, we calculate a CRPS for the small-scale model given the parameter estimates and the medium-scale model that generates the data. To approximate the forecast distribution for a given model, we first initialize the forecasts at the filtered state (or actual state for the DGP) one quarter before the ZLB binds in the data. Then we draw random shocks and simulate the model for 8 quarters, $10{,}000$ times. Using the simulations, we approximate the CDF of the forecast distribution 8-quarters ahead.[17] Finally, we average the CRPS for a given model across the datasets.

Figure 4 shows the mean CRPS across the datasets for the DGP and each estimation method. The horizontal axis denotes the ZLB duration in the data. Due to model misspecification, none of the estimation methods perform as well as the DGP. The DGP has at least a 0.5 percentage point advantage over the estimated models, regardless of the forecasted variable or ZLB duration in the data. Interestingly, the CRPS is similar across the estimation methods. The differences are most pronounced for the nominal interest rate forecasts in datasets where the ZLB binds for 30 quarters.

[16] Appendix D shows the CDF for a specific dataset to illustrate what each term represents in the CRPS calculation.
[17] We obtain similar results with a four quarter forecast horizon, as well as with the RMSE of the point forecast.

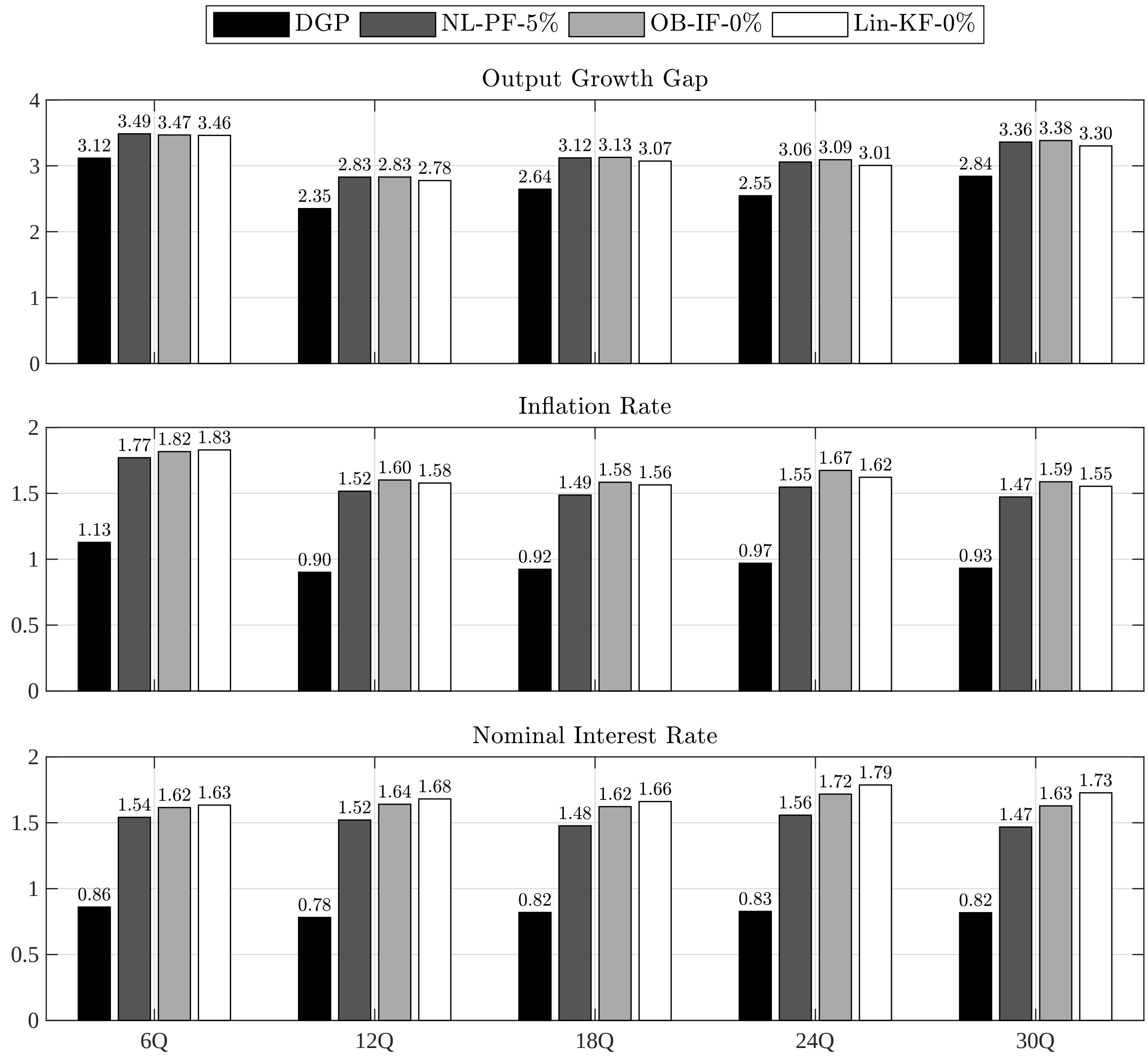


Figure 4: Mean CRPS of 8-quarter ahead forecasts. Forecasts are made one quarter before the ZLB binds in the data.

The NL-PF-5% CRPS is only 179% of the DGP CRPS, compared to 199% for OB-IF-0% and 211% for Lin-KF-0%. The NL-PF-5% forecasts of the inflation rate are also consistently more accurate than the other estimation methods. However, in all cases the differences in accuracy are small relative to the DGP. These findings are consistent with our previous results. NL-PF-5% has an advantage over OB-IF-0%, but it is small and may not be worth the added computational costs.

5 CONCLUSION

During the Great Recession, many central banks lowered their policy rate to its ZLB, creating a kink in the policy rule and calling into question linear estimation methods. There are two promising alternatives: estimate a fully nonlinear model that accounts for the expectational effects of going to the ZLB or a piecewise linear model that is faster but ignores the expectational effects. This paper compares the accuracy of the two methods. We find the predictions of the nonlinear model are typically more accurate than the piecewise linear model, but the differences are often small. There are far larger gains in accuracy from estimating a richer, less misspecified piecewise linear model.

Our results suggest that researchers are better off using piecewise linear models rather than a simpler but properly solved nonlinear model when examining the empirical implications of the ZLB constraint. However, it is important to caution that further research is needed to examine whether our findings in the ZLB context are generalizable to other settings. It is also important to emphasize that the nonlinear model is considerably more versatile. While the piecewise linear and nonlinear models can handle any combination of occasionally binding constraints, only the nonlinear model can account for other nonlinear features emphasized in the literature (e.g., stochastic volatility, asymmetric adjustment costs, non-Gaussian shocks, search frictions, time-varying policy rules, changes in steady states). Our results will serve as an important starting point for future research that explores these nonlinear features or makes advances in nonlinear estimation methods.

REFERENCES

AN, S. AND F. SCHORFHEIDE (2007): “Bayesian Analysis of DSGE Models,” *Econometric Reviews*, 26, 113–172, https://doi.org/10.1080/07474930701220071.

ARUOBA, S., P. CUBA-BORDA, AND F. SCHORFHEIDE (2018): “Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries,” *The Review of Economic Studies*, 85, 87–118, https://doi.org/10.1093/restud/rdx027.

BOCOLA, L. (2016): “The Pass-Through of Sovereign Risk,” *Journal of Political Economy*, 124, 879–926, https://doi.org/10.1086/686734.

CANOVA, F., F. FERRONI, AND C. MATTHES (2014): “Choosing The Variables To Estimate Singular Dsge Models,” *Journal of Applied Econometrics*, 29, 1099–1117, https://doi.org/10.1002/jae.2414.

CHUGH, S. K. (2006): “Optimal Fiscal and Monetary Policy with Sticky Wages and Sticky Prices,” *Review of Economic Dynamics*, 9, 683–714, https://doi.org/10.1016/j.red.2006.07.001.

COLEMAN, II, W. J. (1991): “Equilibrium in a Production Economy with an Income Tax,” *Econometrica*, 59, 1091–1104, https://doi.org/10.2307/2938175.

CUBA-BORDA, P., L. GUERRIERI, M. IACOVIELLO, AND M. ZHONG (2017): “Likelihood Evaluation of Models with Occasionally Binding Constraints,” Federal Reserve Board.

DOH, T. (2011): “Yield Curve in an Estimated Nonlinear Macro Model,” *Journal of Economic Dynamics and Control*, 35, 1229 – 1244, https://doi.org/10.1016/j.jedc.2011.03.003.

Fair, R. C. and J. B. Taylor (1983): "Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models," *Econometrica*, 51, 1169–1185, https://doi.org/10.2307/1912057.

Fernald, J. G. (2012): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," Federal Reserve Bank of San Francisco Working Paper 2012-19.

Fernández-Villaverde, J., G. Gordon, P. Guerrón-Quintana, and J. F. Rubio-Ramírez (2015): "Nonlinear Adventures at the Zero Lower Bound," *Journal of Economic Dynamics and Control*, 57, 182–204, https://doi.org/10.1016/j.jedc.2015.05.014.

Fernández-Villaverde, J. and J. F. Rubio-Ramírez (2005): "Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood," *Journal of Applied Econometrics*, 20, 891–910, https://doi.org/10.1002/jae.814.

——— (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *The Review of Economic Studies*, 74, 1059–1087, https://doi.org/10.1111/j.1467-937X.2007.00437.x.

Gavin, W. T., B. D. Keen, A. W. Richter, and N. A. Throckmorton (2015): "The Zero Lower Bound, the Dual Mandate, and Unconventional Dynamics," *Journal of Economic Dynamics and Control*, 55, 14–38, https://doi.org/10.1016/j.jedc.2015.03.007.

Gordon, N. J., D. J. Salmond, and A. F. M. Smith (1993): "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," *IEE Proceedings F - Radar and Signal Processing*, 140, 107–113, https://doi.org/10.1049/ip-f-2.1993.0015.

Guerrieri, L. and M. Iacoviello (2015): "OccBin: A Toolkit for Solving Dynamic Models with Occasionally Binding Constraints Easily," *Journal of Monetary Economics*, 70, 22–38, https://doi.org/10.1016/j.jmoneco.2014.08.005.

——— (2017): "Collateral Constraints and Macroeconomic Asymmetries," *Journal of Monetary Economics*, 90, 28–49, https://doi.org/10.1016/j.jmoneco.2017.06.004.

Guerrón-Quintana, P. A. (2010): "What You Match Does Matter: The Effects of Data on DSGE Estimation," *Journal of Applied Econometrics*, 25, 774–804, https://doi.org/10.1002/jae.1106.

Gust, C., E. Herbst, D. López-Salido, and M. E. Smith (2017): "The Empirical Implications of the Interest-Rate Lower Bound," *American Economic Review*, 107, 1971–2006, https://doi.org/10.1257/aer.20121437.

Herbst, E. and F. Schorfheide (2018): "Tempered Particle Filtering," *Journal of Econometrics*, forthcoming, https://doi.org/10.1016/j.jeconom.2018.11.003.

Herbst, E. P. and F. Schorfheide (2016): *Bayesian Estimation of DSGE Models*, Princeton, NJ: Princeton University Press.

HIROSE, Y. AND A. INOUE (2016): "The Zero Lower Bound and Parameter Bias in an Estimated DSGE Model," *Journal of Applied Econometrics*, 31, 630–651, https://doi.org/10.1002/jae.2447.

HIROSE, Y. AND T. SUNAKAWA (2015): "Parameter Bias in an Estimated DSGE Model: Does Nonlinearity Mattter?" Centre for Applied Macroceonomic Analysis Working Paper 46/2015.

IIBOSHI, H., M. SHINTANI, AND K. UEDA (2018): "Estimating a Nonlinear New Keynesian Model with a Zero Lower Bound for Japan," Tokyo Center for Economic Research Working Paper E-120.

IRELAND, P. N. (2004): "A Method for Taking Models to the Data," *Journal of Economic Dynamics and Control*, 28, 1205–1226, https://doi.org/10.1016/S0165-1889(03)00080-0.

KEEN, B. D., A. W. RICHTER, AND N. A. THROCKMORTON (2017): "Forward Guidance and the State of the Economy," *Economic Inquiry*, 55, 1593–1624, https://doi.org/10.1111/ecin.12466.

KITAGAWA, G. (1996): "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models," *Journal of Computational and Graphical Statistics*, 5, pp. 1–25, https://doi.org/10.2307/1390750.

KOOP, G., M. H. PESARAN, AND S. M. POTTER (1996): "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74, 119–147, https://doi.org/10.1016/0304-4076(95)01753-4.

KOPECKY, K. AND R. SUEN (2010): "Finite State Markov-chain Approximations to Highly Persistent Processes," *Review of Economic Dynamics*, 13, 701–714, https://doi.org/10.1016/j.red.2010.02.002.

LAUBACH, T. AND J. C. WILLIAMS (2016): "Measuring the Natural Rate of Interest Redux," Finance and Economics Discussion Series 2016-011.

MERTENS, K. AND M. O. RAVN (2014): "Fiscal Policy in an Expectations Driven Liquidity Trap," *The Review of Economic Studies*, 81, 1637–1667, https://doi.org/10.1093/restud/rdu016.

NAKATA, T. (2017): "Uncertainty at the Zero Lower Bound," *American Economic Journal: Macroeconomics*, 9, 186–221, https://doi.org/10.1257/mac.20140253.

NAKOV, A. (2008): "Optimal and Simple Monetary Policy Rules with Zero Floor on the Nominal Interest Rate," *International Journal of Central Banking*, 4, 73–127.

NGO, P. V. (2014): "Optimal Discretionary Monetary Policy in a Micro-Founded Model with a Zero Lower Bound on Nominal Interest Rate," *Journal of Economic Dynamics and Control*, 45, 44–65, https://doi.org/10.1016/j.jedc.2014.05.010.

OTROK, C. (2001): "On Measuring the Welfare Cost of Business Cycles," *Journal of Monetary Economics*, 47, 61–92, https://doi.org/10.1016/S0304-3932(00)00052-0.

PETERMAN, W. B. (2016): "Reconciling Micro and Macro Estimates of the Frisch Labor Supply Elasticity," *Economic Inquiry*, 54, 100–120, https://doi.org/10.1111/ecin.12252.

PLANTE, M., A. W. RICHTER, AND N. A. THROCKMORTON (2018): "The Zero Lower Bound and Endogenous Uncertainty," *Economic Journal*, 128, 1730–1757, https://doi.org/10.1111/ecoj.12445.

RICHTER, A. W. AND N. A. THROCKMORTON (2015): "The Zero Lower Bound: Frequency, Duration, and Numerical Convergence," *B.E. Journal of Macroeconomics*, 15, 157–182, https://doi.org/10.1515/bejm-2013-0185.

——— (2016): "Is Rotemberg Pricing Justified by Macro Data?" *Economics Letters*, 149, 44–48, https://doi.org/10.1016/j.econlet.2016.10.011.

RICHTER, A. W., N. A. THROCKMORTON, AND T. B. WALKER (2014): "Accuracy, Speed and Robustness of Policy Function Iteration," *Computational Economics*, 44, 445–476, https://doi.org/10.1007/s10614-013-9399-2.

ROTEMBERG, J. J. (1982): "Sticky Prices in the United States," *Journal of Political Economy*, 90, 1187–1211, https://doi.org/10.1086/261117.

ROUWENHORST, K. G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, Princeton, NJ: Princeton University Press, 294–330.

SCHORFHEIDE, F. (2000): "Loss Function-Based Evaluation of DSGE Models," *Journal of Applied Econometrics*, 15, 645–670, https://doi.org/10.1002/jae.582.

SIMS, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20, 1–20, https://doi.org/10.1023/A:1020517101123.

STEWART, L. AND P. MCCARTY, JR (1992): "Use of Bayesian Belief Networks to Fuse Continuous and Discrete Information for Target Recognition, Tracking, and Situation Assessment," *Proc. SPIE*, 1699, 177–185, https://doi.org/10.1117/12.138224.

VAN BINSBERGEN, J. H., J. FERNÁNDEZ-VILLAVERDE, R. S. KOIJEN, AND J. RUBIO-RAMÍREZ (2012): "The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences," *Journal of Monetary Economics*, 59, 634–648, https://doi.org/10.1016/j.jmoneco.2012.09.002.

WOLMAN, A. L. (2005): "Real Implications of the Zero Bound on Nominal Interest Rates," *Journal of Money, Credit and Banking*, 37, 273–96, https://doi.org/10.1353/mcb.2005.0026.

A Detrended Equilibrium System

Medium-Scale Model The detrended system includes (1), (6), (7), (9), (16), (17) and

$$\tilde{y}_t = (\upsilon_t \tilde{k}_{t-1}/z_t)^\alpha n_t^{1-\alpha}, \tag{25}$$

$$r_t^k = \alpha mc_t z_t \tilde{y}_t/(\upsilon_t \tilde{k}_{t-1}), \tag{26}$$

$$\tilde{w}_t = (1-\alpha) mc_t \tilde{y}_t/n_t, \tag{27}$$

$$w_t^g = \pi_t z_t \tilde{w}_t/(\bar{\pi}\bar{z}\tilde{w}_{t-1}), \tag{28}$$

$$\tilde{y}_t^{gdp} = [1 - \varphi_p(\pi_t/\bar{\pi}-1)^2/2 - \varphi_w(w_t^g-1)^2/2]\tilde{y}_t - u_t\tilde{k}_{t-1}/z_t, \tag{29}$$

$$y_t^g = z_t \tilde{y}_t^{gdp}/(\bar{z}\tilde{y}_{t-1}^{gdp}), \tag{30}$$

$$\tilde{\lambda}_t = \tilde{c}_t - h\tilde{c}_{t-1}/z_t, \tag{31}$$

$$\tilde{w}_t^f = \chi n_t^\eta \tilde{\lambda}_t, \tag{32}$$

$$\tilde{c}_t + \tilde{x}_t = \tilde{y}_t, \tag{33}$$

$$x_t^g = z_t \tilde{x}_t/(\bar{z}\tilde{x}_{t-1}), \tag{34}$$

$$\tilde{k}_t = (1-\delta)(\tilde{k}_{t-1}/z_t) + \tilde{x}_t(1 - \nu(x_t^g-1)^2/2), \tag{35}$$

$$1 = \beta E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(s_t i_t/(z_{t+1}\pi_{t+1}))], \tag{36}$$

$$q_t = \beta E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(r_{t+1}^k \upsilon_{t+1} - u_{t+1} + (1-\delta)q_{t+1})/z_{t+1}], \tag{37}$$

$$1 = q_t[1 - \nu(x_t^g-1)^2/2 - \nu(x_t^g-1)x_t^g] + \beta\nu\bar{z}E_t[q_{t+1}(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(x_{t+1}^g)^2(x_{t+1}^g-1)/z_{t+1}], \tag{38}$$

$$\varphi_p(\pi_t/\bar{\pi}-1)(\pi_t/\bar{\pi}) = 1 - \theta_p + \theta_p mc_t + \beta\varphi_p E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(\pi_{t+1}/\bar{\pi}-1)(\pi_{t+1}/\bar{\pi})(\tilde{y}_{t+1}/\tilde{y}_t)], \tag{39}$$

$$\varphi_w(w_t^g-1)w_t^g = [(1-\theta_w)\tilde{w}_t + \theta_w\tilde{w}_t^f]n_t/\tilde{y}_t + \beta\varphi_w E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(w_{t+1}^g-1)w_{t+1}^g(\tilde{y}_{t+1}/\tilde{y}_t)]. \tag{40}$$

The variables are $\tilde{c}, n, \tilde{x}, \tilde{k}, \tilde{y}, \tilde{y}^{gdp}, u, \upsilon, w^g, x^g, y^g, \tilde{w}^f, \tilde{w}, r^k, \pi, i, i^n, q, mc, \tilde{\lambda}, z$, and s.

Small-Scale Model The detrended system includes (1), (7), (16), (17), (30), (31), (36), (39), and

$$\tilde{y}_t = n_t, \tag{41}$$

$$\tilde{w}_t = mc_t \tilde{y}_t/n_t, \tag{42}$$

$$\tilde{y}_t^{gdp} = [1 - \varphi_p(\pi_t/\bar{\pi}-1)^2/2]\tilde{y}_t, \tag{43}$$

$$\tilde{w}_t = \chi n_t^\eta \tilde{\lambda}_t, \tag{44}$$

$$\tilde{c}_t = \tilde{y}_t^{gdp}. \tag{45}$$

The variables are $\tilde{c}, n, \tilde{y}, \tilde{y}^{gdp}, y^g, \tilde{w}, \pi, i, i^n, mc, \tilde{\lambda}, z$, and s.

B Nonlinear Solution Method

We begin by compactly writing the detrended nonlinear equilibrium system as

$$E[f(\mathbf{s}_{t+1}, \mathbf{s}_t, \varepsilon_{t+1})|\mathbf{z}_t, \vartheta] = 0,$$

where f is a vector-valued function, $\mathbf{s}_t$ is a vector of variables, $\varepsilon_t \equiv [\varepsilon_{s,t}, \varepsilon_{z,t}, \varepsilon_{i,t}]'$ is a vector of shocks, $\mathbf{z}_t$ is a vector of states ($\mathbf{z}_t \equiv [\tilde{c}_{t-1}, i^n_{t-1}, \tilde{k}_{t-1}, \tilde{x}_{t-1}, \tilde{w}_{t-1}, s_t, z_t, \varepsilon_{i,t}]'$ for the model with capital and $\mathbf{z}_t \equiv [\tilde{c}_{t-1}, i^n_{t-1}, s_t, z_t, \varepsilon_{i,t}]'$ for the model without capital), and ϑ is a vector of parameters.

There are many ways to discretize the exogenous state variables, s_t, z_t, and $\varepsilon_{i,t}$. We use the Markov chain in Rouwenhorst (1995), which Kopecky and Suen (2010) show outperforms other methods for approximating autoregressive processes. The bounds on $\tilde{c}_{t-1}$, i^n_{t-1}, $\tilde{k}_{t-1}$, $\tilde{x}_{t-1}$, and $\tilde{w}_{t-1}$ are respectively set to $\pm 2.5\%$, $\pm 6\%$, $\pm 8\%$, $\pm 15\%$, $\pm 4\%$ of their deterministic steady state. These bounds were chosen so the grids contain 99.9% of the simulated values for each state variable and ZLB duration. We discretize the states into 7 evenly-spaced points, except for capital and the risk premium which use 11 and 13 points, respectively. The product of the points in each dimension, D, represents the total nodes in the state space ($D = 16{,}823{,}807$ for the model with capital and $D = 31{,}213$ for the model without capital). The realization of $\mathbf{z}_t$ on node d is denoted $\mathbf{z}_t(d)$. The Rouwenhorst method provides integration nodes, $[s_{t+1}(m), z_{t+1}(m), \varepsilon_{i,t+1}(m)]$, with weights, $\phi(m)$, for $m \in \{1, \ldots, M\}$. Since the exogenous variables evolve according to a Markov chain, the number of future realizations is the same as the state variables, $(13, 7, 7)$ or $M = 637$.

The vector of policy functions is denoted $\mathbf{pf}_t$ and the realization on node d is denoted $\mathbf{pf}_t(d)$ ($\mathbf{pf}_t \equiv [\tilde{c}_t(\mathbf{z}_t), \pi^{gap}_t(\mathbf{z}_t), n_t(\mathbf{z}_t), q_t(\mathbf{z}_t), \upsilon_t(\mathbf{z}_t)]$ for the capital model and $\mathbf{pf}_t \equiv [\tilde{c}_t(\mathbf{z}_t), \pi^{gap}_t(\mathbf{z}_t)]$ for the model without capital, where $\pi^{gap}_t(\mathbf{z}_t) \equiv \pi_t(\mathbf{z}_t)/\bar{\pi}$). Our choice of policy functions, while not unique, simplifies solving for the other variables in the nonlinear system of equations given $\mathbf{z}_t$.

The following steps outline our global policy function iteration algorithm:

1. Use Sims's (2002) `gensys` algorithm to solve the level-linear model without the ZLB constraint. Then map the solution to the discretized state space to initialize the policy functions.

2. On iteration $j \in \{1, 2, \ldots\}$ and each node $d \in \{1, \ldots, D\}$, use Chris Sims's `csolve` to find $\mathbf{pf}_t(d)$ to satisfy $E[f(\cdot)|\mathbf{z}_t(d), \vartheta] \approx 0$. Guess $\mathbf{pf}_t(d) = \mathbf{pf}_{j-1}(d)$. Then apply the following:

   (a) Solve for all variables dated at time t, given $\mathbf{pf}_t(d)$ and $\mathbf{z}_t(d)$.

   (b) Linearly interpolate the policy functions, $\mathbf{pf}_{j-1}$, at the updated state variables, $\mathbf{z}_{t+1}(m)$, to obtain $\mathbf{pf}_{t+1}(m)$ on every integration node, $m \in \{1, \ldots, M\}$.

   (c) Given $\{\mathbf{pf}_{t+1}(m)\}_{m=1}^M$, solve for the other elements of $\mathbf{s}_{t+1}(m)$ and compute

$$\mathbb{E}[f(\mathbf{s}_{t+1}, \mathbf{s}_t(d), \varepsilon_{t+1})|\mathbf{z}_t(d), \vartheta] \approx \textstyle\sum_{m=1}^M \phi(m) f(\mathbf{s}_{t+1}(m), \mathbf{s}_t(d), \varepsilon_{t+1}(m)).$$

   When `csolve` converges, set $\mathbf{pf}_j(d) = \mathbf{pf}_t(d)$.

3. Repeat step 2 until $\text{maxdist}_j < 10^{-6}$, where $\text{maxdist}_j \equiv \max\{|\mathbf{pf}_j - \mathbf{pf}_{j-1}|\}$. When that criterion is satisfied, the algorithm has converged to an approximate nonlinear solution.

C ESTIMATION ALGORITHM

We use a random walk Metropolis-Hastings algorithm to estimate the model in section 3 with artificial data of 120 quarters. To measure how well the model fits the data, we use either the adapted particle filter described in Algorithm 14 in Herbst and Schorfheide (2016), which modifies the basic bootstrap filter in Stewart and McCarty (1992) and Gordon et al. (1993) to better account for the outliers in the data, or the inversion filter recently used by Guerrieri and Iacoviello (2017).

C.1 METROPOLIS-HASTINGS ALGORITHM

The following steps outline the algorithm:

1. Generate artificial data consisting of the output growth gap, the inflation rate, and the nominal interest rate, $\mathbf{x}_t \equiv [y_t^g, \pi_t, i_t]'$, where $N_x = 3$ is the number of observable variables.

2. Specify the prior distributions, means, variances, and bounds of each element of the vector of N_e estimated parameters, $\theta \equiv [\varphi_p, \phi_\pi, \phi_y, h, \rho_s, \rho_i, \sigma_z, \sigma_s, \sigma_i]'$.

3. Find the posterior mode to initialize the preliminary Metropolis-Hastings step.

   (a) For all $i \in \{1, \ldots, N_m\}$, where $N_m = 5{,}000$, apply the following steps:

      i. Draw $\hat{\theta}_i$ from the joint prior distribution and calculate its density value:

      $$\log \ell_i^{prior} = \textstyle\sum_{j=1}^{N_e} \log p(\hat{\theta}_{i,j} | \mu_j, \sigma_j^2),$$

      where p is the prior density function of parameter j with mean μ_j and variance σ_j^2.

      ii. Solve the model given $\hat{\theta}_i$. Follow Appendix B for the nonlinear model and use OccBin for the PW linear model. Repeat 3(a)i if the algorithm does not converge.

      iii. Obtain the model log-likelihood, $\log \ell_i^{model}$. Apply the particle filter described in section C.2 to the nonlinear model and the inversion filter to the PW linear model.

      iv. The posterior log-likelihood is $\log \ell_i^{post} = \log \ell_i^{prior} + \log \ell_i^{model}$

   (b) Calculate $\max(\log \ell_1^{post}, \ldots, \log \ell_{N_m}^{post})$ and find the corresponding parameter vector, $\hat{\theta}_0$.

4. Approximate the covariance matrix for the joint posterior distribution of the parameters, Σ, which is used to obtain candidate draws during the preliminary Metropolis-Hastings step.

   (a) Locate the draws with a likelihood in the top decile. Stack the $N_{m,sub} = (1-p)N_m$ draws in a $N_{m,sub} \times N_e$ matrix, $\hat{\Theta}$, and define $\tilde{\Theta} = \hat{\Theta} - \sum_{i=1}^{N_{m,sub}} \hat{\theta}_{i,j}/N_{m,sub}$.

   (b) Calculate $\Sigma = \tilde{\Theta}'\tilde{\Theta}/N_{m,sub}$ and verify it is positive definite, otherwise repeat step 3.

5. Perform an initial run of the random walk Metropolis-Hastings algorithm.

(a) For all $i \in \{0, \ldots, N_d\}$, where $N_d = 25{,}000$, perform the following steps:

i. Draw a candidate vector of parameters, $\hat{\theta}_i^{cand}$, where

$$\hat{\theta}_i^{cand} \sim \begin{cases} \mathbb{N}(\hat{\theta}_0, c_0\Sigma) & \text{for } i = 0, \\ \mathbb{N}(\hat{\theta}_{i-1}, c\Sigma) & \text{for } i > 0. \end{cases}$$

We set $c_0 = 0$ and tune c to target an overall acceptance rate of roughly 30%.

ii. Calculate the prior density value, $\log \ell_i^{prior}$, of the candidate draw, $\hat{\theta}_i^{cand}$, as in 3(a)i.

iii. Solve the model given $\hat{\theta}_i^{cand}$. If the algorithm does not converge repeat 5(a)i.

iv. Obtain the model log-likelihood value, $\log \ell_i^{model}$, using the methods in 3(a)iii.

v. Accept or reject the candidate draw according to

$$(\hat{\theta}_i, \log \ell_i) = \begin{cases} (\hat{\theta}_i^{cand}, \log \ell_i^{cand}) & \text{if } i = 0, \\ (\hat{\theta}_i^{cand}, \log \ell_i^{cand}) & \text{if } \min(1, \ell_i^{cand}/\ell_{i-1}) > \hat{u}, \\ (\hat{\theta}_{i-1}, \log \ell_{i-1}) & \text{otherwise,} \end{cases}$$

where $\hat{u}$ is a draw from a uniform distribution, $\mathbb{U}[0, 1]$, and the posterior log-likelihood associated with the candidate draw is $\log \ell_i^{cand} = \log \ell_i^{prior} + \log \ell_i^{model}$.

(b) Burn the first $N_b = 5{,}000$ draws and use the remaining sample to calculate the mean draw, $\hat{\theta}^{5(b)} = \sum_{i=N_b+1}^{N_d} \hat{\theta}_i/(N_d - N_b)$, and the covariance matrix, $\Sigma^{5(b)}$. We follow step 4 to calculate $\Sigma^{5(b)}$ but use all $N_d - N_b$ draws instead of just the upper pth percentile.

6. Conduct a final run of the Metropolis-Hastings algorithm by repeating step 5, where $N_d = 50{,}000$, $\hat{\theta}_0 = \hat{\theta}^{5(b)}$, and $\Sigma = \Sigma^{5(b)}$. The final posterior mean estimates are $\hat{\theta} = \sum_{i=1}^{N_d} \hat{\theta}_i/N_d$.

C.2 ADAPTED PARTICLE FILTER Henceforth, our definition of $\mathbf{s}_t$ from Appendix B is referred to as the state vector, which should not be confused with the state variables for the nonlinear model.

1. Initialize the filter by drawing $\{\varepsilon_{t,p}\}_{t=-24}^{0}$ for all $p \in \{0, \ldots, N_p\}$ and simulating the model, where N_p is the number of particles. We initialize the filter with the final state vector, $\mathbf{s}_{0,p}$, which is approximately a draw from the model's ergodic distribution. We set $N_p = 40{,}000$.

2. For $t \in \{1, \ldots, T\}$, sequentially filter the nonlinear model as follows:

(a) For $p \in \{1, \ldots, N_p\}$, draw shocks from an adapted distribution, $\varepsilon_{t,p} \sim \mathbb{N}(\bar{\varepsilon}_t, I)$, where $\bar{\varepsilon}_t$ maximizes $p(\xi_t|\mathbf{s}_t)p(\mathbf{s}_t|\bar{\mathbf{s}}_{t-1})$ and $\bar{\mathbf{s}}_{t-1} = \sum_{p=1}^{N_p} \mathbf{s}_{t-1,p}/N_p$ is the mean state vector.

i. Use the model solution to update the state vector, $\mathbf{s}_t$, given $\bar{\mathbf{s}}_{t-1}$ and a guess for $\bar{\varepsilon}_t$. Define $\mathbf{s}_t^h \equiv H\mathbf{s}_t$, where H selects the observable variables from the state vector.

ii. Calculate the measurement error, $\xi_t = \mathbf{s}_t^h - \mathbf{x}_t$, which is assumed to be multivariate normally distributed, $p(\xi_t|\mathbf{s}_t) = (2\pi)^{-3/2}|R|^{-1/2}\exp(-\xi_t' R^{-1}\xi_t/2)$, where $R \equiv \mathrm{diag}(\sigma^2_{me,y^g}, \sigma^2_{me,\pi}, \sigma^2_{me,i})$ is a diagonal matrix of measurement error variances.

iii. The probability of observing the current state, $\mathbf{s}_t$, conditional on $\bar{\mathbf{s}}_{t-1}$, is given by

$$p(\mathbf{s}_t|\bar{\mathbf{s}}_{t-1}) = (2\pi)^{-3/2}\exp(-\bar{\varepsilon}_t'\bar{\varepsilon}_t/2).$$

iv. Maximize $p(\xi_t|\mathbf{s}_t)p(\mathbf{s}_t|\bar{\mathbf{s}}_{t-1}) \propto \exp(-\xi_t' R^{-1}\xi_t/2)\exp(-\bar{\varepsilon}_t'\bar{\varepsilon}_t/2)$ by solving for the optimal $\bar{\varepsilon}_t$. We use MATLAB's `fminsearch` routine converted to Fortran.

(b) Use the model solution to predict the state vector, $\mathbf{s}_{t,p}$, given $\mathbf{s}_{t-1,p}$ and $\varepsilon_{t,p}$.

(c) Calculate $\xi_{t,p} = \mathbf{s}_{t,p}^h - \mathbf{x}_t$. The unnormalized weight on particle p is given by

$$\omega_{t,p} = \frac{p(\xi_t|\mathbf{s}_{t,p})p(\mathbf{s}_{t,p}|\mathbf{s}_{t-1,p})}{g(\mathbf{s}_{t,p}|\mathbf{s}_{t-1,p},\mathbf{x}_t)} \propto \frac{\exp(-\xi_{t,p}' R^{-1}\xi_{t,p}/2)\exp(-\varepsilon_{t,p}'\varepsilon_{t,p}/2)}{\exp(-(\varepsilon_{t,p}-\bar{\varepsilon}_t)'(\varepsilon_{t,p}-\bar{\varepsilon}_t)/2)}.$$

Without adaptation, $\bar{\varepsilon}_t = 0$ and $\omega_{t,p} = p(\xi_t|\mathbf{s}_{t,p})$, as in a basic bootstrap particle filter. The time-t contribution to the model log-likelihood is $\ell_t^{model} = \sum_{p=1}^{N_p}\omega_{t,p}/N_p$.

(d) Normalize the weights, $W_{t,p} = \omega_{t,p}/\sum_{p=1}^{N_p}\omega_{t,p}$. Then use systematic resampling with replacement from the swarm of particles as described in Kitagawa (1996) to get a set of particles that represents the filter distribution and reshuffle $\{\mathbf{s}_{t,p}\}_{p=1}^{N_p}$ accordingly.

3. The model log-likelihood is $\log \ell^{model} = \sum_{t=1}^{T}\log \ell_t^{model}$.

Aruoba et al. (2018) apply the same methodology to a New Keynesian model with sunspot shocks. See Herbst and Schorfheide (2016) for a comprehensive discussion of the different particle filters.

D CONTINUOUS RANK PROBABILITY SCORE (CRPS) EXAMPLE

Figure 5 shows an example of the 8-quarter ahead forecast distribution of the nominal interest rate given the parameter estimates from NL-PF-5%. We picked a dataset where the ZLB binds for six quarters, from period 90 to 95 in the sample. The forecasts are initialized at the filtered state in period 89, immediately before the ZLB first binds, and the forecast distribution is approximated based on $10{,}000$ simulations. Due to a strong tendency for the forecasts to revert to the stochastic steady state, the mean forecast for the nominal interest rate is 2.32%. However, the probability density function (PDF) in the left panel shows a significant number of forecasts remain near or at the ZLB, even after 8 quarters. The true realization equals 1.94%, which means there is significant probability mass under the PDF above and below the true value. The right panel shows the cumulative distribution function (CDF) of the forecasts. The CRPS for this dataset and estimation method is closely related to the shaded area, which has the same units as the forecasted variable.

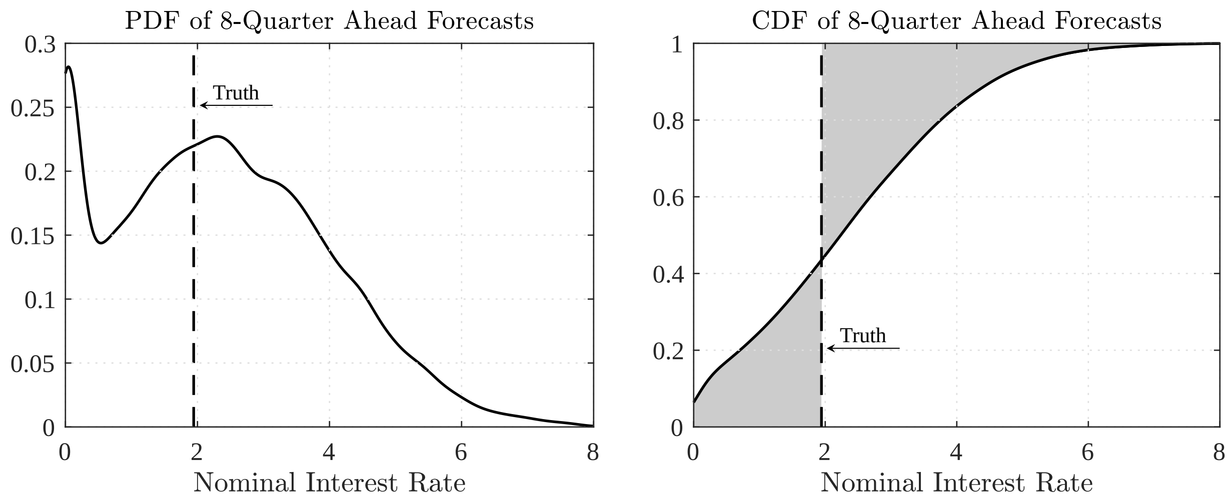


Figure 5: Example forecast distribution in the period before the ZLB binds in the data.

E ADDITIONAL RESULTS

First, we examine the sources of the bias in the estimates of the habit persistence and price adjustment cost parameters. Second, we report the parameter estimates for datasets with ZLB events between 0 and 30 quarters long. Third, we show how misspecification affects the parameter estimates and impulse responses using generated data from our small-scale model. Fourth, we plot impulse responses to a productivity growth and monetary policy shock when the ZLB binds. Fifth, we compare the filtered paths of the notional interest rate. Sixth, we provide additional statistics about the ZLB events in our datasets. Finally, we examine how government spending affects our results.

E.1 PRICE ADJUSTMENT COST AND HABIT PERSISTENCE In table 4, estimates of the price adjustment cost (φ_p) and habit persistence (h) parameters have some of the largest NRMSEs, even in datasets without a ZLB event. These parameters are critical for output and inflation dynamics, so understanding the source of the bias is important for interpreting our results. The small-scale model lacks important shock amplifiers for output, such as sticky wages and variable capital utilization. Therefore, the response of output growth is too small when the model is parameterized with the true values. Conversely, the lack of sticky wages means marginal costs are overly volatile and inflation is too sensitive to shocks. If misspecification impacted the responses of output growth and inflation in the same direction, the estimated shock size would have been affected. Instead, estimates of h are lower than the true value, amplifying the response of output. Estimates of φ_p are biased upward, flattening the price Phillips curve and stabilizing inflation despite overly volatile marginal costs.

Another potentially important source of the bias is the misspecification in the aggregate resource constraint. Movements in wage adjustment costs, capital utilization costs and other terms could be interpreted as price adjustment costs through a larger estimate of φ_p. However, that is unlikely to drive the bias in estimates of φ_p and h. The NL-PF-5% and Lin-KF-5% estimates of φ_p and h are very similar, despite the absence of price adjustment costs in the aggregate resource constraint in the linear model (i.e., $\hat{y}_t = \hat{y}_t^{gdp} = \hat{c}_t$). Therefore, the upward bias in φ_p is not the result of price adjustment costs absorbing the gap between consumption and output in the DGP.

The middle columns of table 5, where only sticky wages are added to the small-scale model, support these conclusions. In particular, in datasets without a ZLB event, there is virtually no bias in the OB-IF-0% estimates of φ_p and h, but there is a large upward bias in σ_s. When sticky wages are added, the volatility of output growth is still too small due to the absence of investment and capital utilization, but the volatility of inflation is now proportionally too small as well. σ_s increases to match the dynamics of the output and inflation data, while h and φ_p remain close to their true values. In the right two columns of table 5, the full model is estimated and σ_s is close to the truth.

Ptr	Truth	NL-PF-5%		OB-IF-0%		Lin-KF-5%	
		0Q	30Q	0Q	30Q	0Q	30Q
ρ_s	0.8	0.43 (0.37, 0.50) {0.47, 0.00}	0.52 (0.40, 0.68) {0.37, 0.04}	0.39 (0.31, 0.47) {0.52, 0.00}	0.44 (0.26, 0.71) {0.48, 0.04}	0.43 (0.35, 0.50) {0.47, 0.00}	0.55 (0.40, 0.77) {0.33, 0.12}
ρ_i	0.8	0.74 (0.69, 0.78) {0.09, 0.26}	0.75 (0.71, 0.81) {0.07, 0.52}	0.69 (0.65, 0.73) {0.15, 0.00}	0.68 (0.62, 0.73) {0.16, 0.00}	0.74 (0.70, 0.78) {0.09, 0.30}	0.79 (0.73, 0.84) {0.05, 0.86}
σ_z	0.005	0.0052 (0.0041, 0.0067) {0.17, 0.88}	0.0062 (0.0037, 0.0134) {0.54, 0.82}	0.0086 (0.0069, 0.0099) {0.73, 0.00}	0.0107 (0.0071, 0.0163) {1.28, 0.00}	0.0053 (0.0041, 0.0067) {0.17, 0.86}	0.0078 (0.0042, 0.0138) {0.83, 0.44}
σ_s	0.005	0.0166 (0.0139, 0.0212) {2.37, 0.00}	0.0196 (0.0113, 0.0261) {3.04, 0.12}	0.0183 (0.0143, 0.0230) {2.71, 0.00}	0.0239 (0.0085, 0.0355) {4.15, 0.04}	0.0169 (0.0141, 0.0216) {2.42, 0.00}	0.0169 (0.0065, 0.0257) {2.59, 0.12}
σ_i	0.002	0.0018 (0.0015, 0.0022) {0.13, 0.64}	0.0016 (0.0014, 0.0021) {0.21, 0.38}	0.0021 (0.0019, 0.0023) {0.09, 0.78}	0.0021 (0.0019, 0.0025) {0.11, 0.78}	0.0018 (0.0015, 0.0022) {0.13, 0.64}	0.0017 (0.0015, 0.0020) {0.16, 0.44}
ϕ_π	2.0	2.04 (1.81, 2.23) {0.07, 0.96}	2.03 (1.84, 2.33) {0.07, 0.90}	1.96 (1.70, 2.21) {0.08, 0.96}	1.84 (1.53, 2.24) {0.14, 0.80}	2.01 (1.78, 2.22) {0.07, 0.98}	1.64 (1.41, 1.89) {0.19, 0.44}
ϕ_y	0.5	0.23 (0.11, 0.40) {0.56, 0.32}	0.29 (0.14, 0.50) {0.49, 0.54}	0.13 (0.05, 0.22) {0.75, 0.02}	0.20 (0.05, 0.35) {0.65, 0.10}	0.24 (0.11, 0.41) {0.56, 0.30}	0.19 (0.08, 0.36) {0.64, 0.18}
Σ		3.86	4.80	5.02	6.96	3.91	4.81

Table 7: Average, $(5, 95)$ percentiles, and $\{\text{NRMSE}, \text{CR}\}$. Σ is the sum of the NRMSE across the parameters.

Lastly, we fixed φ_p and h to their true values and re-estimated each specification. Table 7 reports the results, which show how other parameters adjust. In particular, σ_s is now 3 to 4 times higher than its true value and ρ_s drops to roughly half of its true value. The NRMSEs for σ_s are by far the largest of any parameter and the CRs are all near 0. In this exercise, h cannot fall to compensate for the missing frictions, so the size of the risk premium shocks must increase. This effect, in addition to not allowing φ_p to increase to compensate for the lack of sticky wages, induces too much inflation volatility. Therefore, the estimate of risk premium persistence, ρ_s, falls. Unlike its shock size, its persistence affects the inflation response more than the output growth response.

Ptr	Truth	0Q	6Q	12Q	18Q	24Q	30Q
				NL-PF-5%			
φ_p	100	151.1 $(134.2, 165.8)$ $\{0.52, 0.02\}$	161.0 $(143.2, 179.3)$ $\{0.62, 0.00\}$	172.1 $(153.8, 193.4)$ $\{0.73, 0.00\}$	180.6 $(161.3, 201.4)$ $\{0.81, 0.18\}$	187.2 $(167.0, 204.5)$ $\{0.88, 0.00\}$	188.4 $(174.7, 202.7)$ $\{0.89, 0.00\}$
h	0.8	0.66 $(0.62, 0.70)$ $\{0.18, 0.00\}$	0.66 $(0.61, 0.71)$ $\{0.17, 0.00\}$	0.67 $(0.62, 0.71)$ $\{0.17, 0.00\}$	0.67 $(0.63, 0.71)$ $\{0.16, 0.00\}$	0.68 $(0.64, 0.72)$ $\{0.15, 0.00\}$	0.68 $(0.64, 0.71)$ $\{0.16, 0.00\}$
ρ_s	0.8	0.76 $(0.72, 0.80)$ $\{0.06, 0.70\}$	0.77 $(0.74, 0.81)$ $\{0.04, 0.86\}$	0.79 $(0.75, 0.82)$ $\{0.03, 0.98\}$	0.80 $(0.77, 0.84)$ $\{0.03, 0.92\}$	0.81 $(0.78, 0.83)$ $\{0.02, 0.96\}$	0.81 $(0.78, 0.84)$ $\{0.03, 0.90\}$
ρ_i	0.8	0.79 $(0.75, 0.82)$ $\{0.03, 0.96\}$	0.79 $(0.75, 0.82)$ $\{0.04, 0.90\}$	0.79 $(0.77, 0.82)$ $\{0.02, 1.00\}$	0.80 $(0.76, 0.83)$ $\{0.03, 0.94\}$	0.80 $(0.76, 0.84)$ $\{0.03, 0.94\}$	0.80 $(0.75, 0.84)$ $\{0.03, 0.96\}$
σ_z	0.0050	0.0032 $(0.0023, 0.0039)$ $\{0.37, 0.00\}$	0.0032 $(0.0023, 0.0041)$ $\{0.38, 0.08\}$	0.0034 $(0.0024, 0.0044)$ $\{0.34, 0.18\}$	0.0037 $(0.0027, 0.0049)$ $\{0.29, 0.38\}$	0.0038 $(0.0027, 0.0047)$ $\{0.28, 0.46\}$	0.0040 $(0.0030, 0.0052)$ $\{0.23, 0.58\}$
σ_s	0.0050	0.0052 $(0.0040, 0.0066)$ $\{0.15, 0.92\}$	0.0052 $(0.0042, 0.0068)$ $\{0.15, 0.92\}$	0.0051 $(0.0040, 0.0060)$ $\{0.13, 0.98\}$	0.0051 $(0.0034, 0.0064)$ $\{0.18, 0.86\}$	0.0050 $(0.0041, 0.0064)$ $\{0.12, 1.00\}$	0.0050 $(0.0039, 0.0062)$ $\{0.13, 0.96\}$
σ_i	0.0020	0.0017 $(0.0014, 0.0020)$ $\{0.17, 0.48\}$	0.0017 $(0.0014, 0.0019)$ $\{0.18, 0.40\}$	0.0016 $(0.0014, 0.0019)$ $\{0.21, 0.30\}$	0.0016 $(0.0013, 0.0019)$ $\{0.24, 0.26\}$	0.0015 $(0.0013, 0.0018)$ $\{0.25, 0.20\}$	0.0015 $(0.0013, 0.0019)$ $\{0.24, 0.20\}$
ϕ_π	2.0	2.04 $(1.88, 2.19)$ $\{0.06, 0.98\}$	2.06 $(1.87, 2.24)$ $\{0.07, 0.96\}$	2.12 $(1.94, 2.33)$ $\{0.08, 0.92\}$	2.13 $(1.90, 2.41)$ $\{0.10, 0.94\}$	2.10 $(1.84, 2.33)$ $\{0.09, 0.90\}$	2.13 $(1.94, 2.31)$ $\{0.09, 0.92\}$
ϕ_y	0.5	0.35 $(0.21, 0.54)$ $\{0.36, 0.80\}$	0.39 $(0.22, 0.61)$ $\{0.31, 0.92\}$	0.41 $(0.27, 0.60)$ $\{0.27, 1.00\}$	0.40 $(0.26, 0.54)$ $\{0.27, 0.92\}$	0.41 $(0.26, 0.61)$ $\{0.27, 0.98\}$	0.42 $(0.27, 0.62)$ $\{0.28, 0.98\}$
Σ		1.90	1.96	1.99	2.12	2.09	2.08
				OB-IF-0%			
φ_p	100	142.6 $(121.1, 157.3)$ $\{0.44, 0.08\}$	152.5 $(131.3, 170.7)$ $\{0.54, 0.02\}$	164.5 $(140.8, 185.5)$ $\{0.66, 0.00\}$	174.7 $(153.9, 202.0)$ $\{0.76, 0.00\}$	183.1 $(165.3, 204.1)$ $\{0.84, 0.00\}$	183.4 $(169.2, 198.5)$ $\{0.84, 0.00\}$
h	0.8	0.64 $(0.61, 0.67)$ $\{0.20, 0.00\}$	0.64 $(0.61, 0.68)$ $\{0.20, 0.00\}$	0.63 $(0.60, 0.67)$ $\{0.21, 0.00\}$	0.63 $(0.61, 0.67)$ $\{0.21, 0.00\}$	0.63 $(0.59, 0.67)$ $\{0.21, 0.00\}$	0.63 $(0.60, 0.67)$ $\{0.21, 0.00\}$
ρ_s	0.8	0.76 $(0.73, 0.81)$ $\{0.05, 0.82\}$	0.77 $(0.73, 0.81)$ $\{0.04, 0.92\}$	0.80 $(0.76, 0.83)$ $\{0.03, 0.96\}$	0.81 $(0.78, 0.85)$ $\{0.03, 0.86\}$	0.82 $(0.80, 0.85)$ $\{0.03, 0.76\}$	0.82 $(0.79, 0.86)$ $\{0.04, 0.78\}$
ρ_i	0.8	0.76 $(0.71, 0.79)$ $\{0.06, 0.52\}$	0.75 $(0.71, 0.80)$ $\{0.07, 0.50\}$	0.76 $(0.73, 0.79)$ $\{0.06, 0.54\}$	0.76 $(0.68, 0.80)$ $\{0.06, 0.58\}$	0.76 $(0.72, 0.81)$ $\{0.06, 0.58\}$	0.77 $(0.73, 0.81)$ $\{0.05, 0.66\}$
σ_z	0.0050	0.0051 $(0.0044, 0.0058)$ $\{0.09, 0.92\}$	0.0053 $(0.0048, 0.0068)$ $\{0.13, 0.82\}$	0.0056 $(0.0047, 0.0066)$ $\{0.19, 0.60\}$	0.0059 $(0.0051, 0.0079)$ $\{0.24, 0.54\}$	0.0060 $(0.0051, 0.0074)$ $\{0.25, 0.46\}$	0.0059 $(0.0050, 0.0069)$ $\{0.22, 0.30\}$
σ_s	0.0050	0.0051 $(0.0042, 0.0063)$ $\{0.13, 0.92\}$	0.0051 $(0.0041, 0.0063)$ $\{0.14, 0.96\}$	0.0048 $(0.0039, 0.0058)$ $\{0.13, 0.90\}$	0.0047 $(0.0031, 0.0058)$ $\{0.18, 0.76\}$	0.0045 $(0.0037, 0.0053)$ $\{0.15, 0.80\}$	0.0046 $(0.0036, 0.0056)$ $\{0.15, 0.82\}$
σ_i	0.0020	0.0020 $(0.0018, 0.0023)$ $\{0.08, 0.90\}$	0.0020 $(0.0018, 0.0023)$ $\{0.07, 0.90\}$	0.0020 $(0.0018, 0.0022)$ $\{0.07, 0.98\}$	0.0020 $(0.0018, 0.0024)$ $\{0.09, 0.82\}$	0.0020 $(0.0018, 0.0023)$ $\{0.08, 0.88\}$	0.0020 $(0.0019, 0.0024)$ $\{0.09, 0.90\}$
ϕ_π	2.0	2.01 $(1.84, 2.16)$ $\{0.06, 0.98\}$	1.96 $(1.77, 2.16)$ $\{0.07, 0.98\}$	1.99 $(1.78, 2.16)$ $\{0.06, 0.98\}$	1.97 $(1.73, 2.23)$ $\{0.08, 0.96\}$	1.94 $(1.69, 2.19)$ $\{0.08, 0.90\}$	1.96 $(1.77, 2.14)$ $\{0.06, 0.98\}$
ϕ_y	0.5	0.32 $(0.17, 0.48)$ $\{0.41, 0.68\}$	0.35 $(0.18, 0.53)$ $\{0.37, 0.76\}$	0.39 $(0.24, 0.56)$ $\{0.30, 0.90\}$	0.36 $(0.20, 0.52)$ $\{0.35, 0.80\}$	0.41 $(0.21, 0.62)$ $\{0.29, 0.90\}$	0.44 $(0.27, 0.61)$ $\{0.25, 0.98\}$
Σ		1.53	1.63	1.71	2.01	1.99	1.91

Table 8: Average, $(5, 95)$ percentiles, and $\{\mathrm{NRMSE}, \mathrm{CR}\}$. Σ is the sum of the NRMSE across the parameters.

Ptr	Truth	0Q	6Q	12Q	18Q	24Q	30Q
				Lin-KF-0%			
φ_p	100	143.0 $(125.9, 157.7)$ $\{0.44, 0.04\}$	153.3 $(134.2, 168.4)$ $\{0.54, 0.00\}$	167.2 $(147.0, 196.6)$ $\{0.69, 0.00\}$	177.5 $(157.1, 204.9)$ $\{0.79, 0.00\}$	186.3 $(165.6, 204.5)$ $\{0.87, 0.00\}$	186.9 $(168.5, 201.1)$ $\{0.88, 0.00\}$
h	0.8	0.64 $(0.61, 0.68)$ $\{0.20, 0.00\}$	0.64 $(0.60, 0.68)$ $\{0.20, 0.00\}$	0.64 $(0.60, 0.67)$ $\{0.20, 0.00\}$	0.64 $(0.62, 0.67)$ $\{0.20, 0.00\}$	0.64 $(0.60, 0.67)$ $\{0.20, 0.00\}$	0.63 $(0.60, 0.67)$ $\{0.21, 0.00\}$
ρ_s	0.8	0.76 $(0.72, 0.80)$ $\{0.06, 0.74\}$	0.77 $(0.74, 0.80)$ $\{0.04, 0.88\}$	0.80 $(0.76, 0.83)$ $\{0.03, 1.00\}$	0.81 $(0.76, 0.84)$ $\{0.03, 0.92\}$	0.82 $(0.80, 0.85)$ $\{0.03, 0.82\}$	0.82 $(0.80, 0.85)$ $\{0.04, 0.78\}$
ρ_i	0.8	0.76 $(0.73, 0.79)$ $\{0.06, 0.62\}$	0.77 $(0.72, 0.80)$ $\{0.05, 0.70\}$	0.78 $(0.75, 0.81)$ $\{0.04, 0.92\}$	0.79 $(0.74, 0.84)$ $\{0.03, 0.88\}$	0.80 $(0.77, 0.85)$ $\{0.03, 0.90\}$	0.81 $(0.77, 0.85)$ $\{0.03, 0.90\}$
σ_z	0.0050	0.0049 $(0.0043, 0.0054)$ $\{0.07, 0.90\}$	0.0051 $(0.0045, 0.0058)$ $\{0.08, 0.88\}$	0.0055 $(0.0048, 0.0066)$ $\{0.16, 0.56\}$	0.0057 $(0.0051, 0.0067)$ $\{0.17, 0.50\}$	0.0060 $(0.0049, 0.0071)$ $\{0.23, 0.32\}$	0.0059 $(0.0051, 0.0068)$ $\{0.21, 0.28\}$
σ_s	0.0050	0.0052 $(0.0043, 0.0064)$ $\{0.14, 0.86\}$	0.0051 $(0.0042, 0.0062)$ $\{0.14, 0.96\}$	0.0048 $(0.0040, 0.0058)$ $\{0.12, 0.96\}$	0.0048 $(0.0035, 0.0059)$ $\{0.15, 0.86\}$	0.0045 $(0.0038, 0.0053)$ $\{0.15, 0.78\}$	0.0045 $(0.0036, 0.0052)$ $\{0.15, 0.86\}$
σ_i	0.0020	0.0020 $(0.0018, 0.0022)$ $\{0.07, 0.96\}$	0.0020 $(0.0018, 0.0022)$ $\{0.07, 0.88\}$	0.0020 $(0.0018, 0.0023)$ $\{0.08, 0.88\}$	0.0020 $(0.0016, 0.0022)$ $\{0.08, 0.82\}$	0.0020 $(0.0017, 0.0022)$ $\{0.08, 0.88\}$	0.0019 $(0.0017, 0.0022)$ $\{0.08, 0.88\}$
ϕ_π	2.0	2.01 $(1.85, 2.15)$ $\{0.06, 0.98\}$	1.96 $(1.71, 2.17)$ $\{0.07, 1.00\}$	1.85 $(1.60, 2.07)$ $\{0.10, 0.94\}$	1.78 $(1.51, 2.04)$ $\{0.14, 0.76\}$	1.65 $(1.42, 1.92)$ $\{0.19, 0.44\}$	1.69 $(1.46, 1.89)$ $\{0.17, 0.64\}$
ϕ_y	0.5	0.32 $(0.18, 0.48)$ $\{0.40, 0.72\}$	0.32 $(0.20, 0.52)$ $\{0.41, 0.60\}$	0.28 $(0.11, 0.48)$ $\{0.48, 0.50\}$	0.26 $(0.14, 0.43)$ $\{0.51, 0.32\}$	0.25 $(0.15, 0.37)$ $\{0.51, 0.32\}$	0.28 $(0.17, 0.44)$ $\{0.47, 0.44\}$
Σ		1.49	1.62	1.89	2.10	2.30	2.24
				Lin-KF-5%			
φ_p	100	151.4 $(134.0, 165.7)$ $\{0.52, 0.00\}$	161.1 $(142.0, 179.5)$ $\{0.62, 0.00\}$	174.8 $(153.7, 198.6)$ $\{0.76, 0.00\}$	183.1 $(163.0, 208.5)$ $\{0.84, 0.00\}$	191.1 $(172.1, 210.9)$ $\{0.92, 0.00\}$	191.6 $(175.3, 204.1)$ $\{0.92, 0.00\}$
h	0.8	0.66 $(0.62, 0.69)$ $\{0.18, 0.00\}$	0.66 $(0.61, 0.71)$ $\{0.18, 0.00\}$	0.67 $(0.62, 0.71)$ $\{0.17, 0.00\}$	0.67 $(0.63, 0.70)$ $\{0.17, 0.00\}$	0.67 $(0.64, 0.71)$ $\{0.16, 0.00\}$	0.67 $(0.63, 0.70)$ $\{0.17, 0.00\}$
ρ_s	0.8	0.76 $(0.72, 0.80)$ $\{0.06, 0.74\}$	0.78 $(0.74, 0.81)$ $\{0.04, 0.92\}$	0.80 $(0.75, 0.83)$ $\{0.03, 1.00\}$	0.81 $(0.78, 0.85)$ $\{0.03, 0.00\}$	0.82 $(0.79, 0.85)$ $\{0.03, 0.88\}$	0.82 $(0.78, 0.86)$ $\{0.04, 0.78\}$
ρ_i	0.8	0.79 $(0.75, 0.82)$ $\{0.03, 0.98\}$	0.80 $(0.75, 0.83)$ $\{0.04, 0.96\}$	0.81 $(0.78, 0.84)$ $\{0.03, 0.94\}$	0.83 $(0.78, 0.86)$ $\{0.04, 0.00\}$	0.83 $(0.80, 0.88)$ $\{0.05, 0.70\}$	0.84 $(0.80, 0.88)$ $\{0.06, 0.56\}$
σ_z	0.0050	0.0032 $(0.0023, 0.0039)$ $\{0.36, 0.00\}$	0.0033 $(0.0025, 0.0041)$ $\{0.36, 0.12\}$	0.0036 $(0.0027, 0.0045)$ $\{0.31, 0.32\}$	0.0040 $(0.0029, 0.0052)$ $\{0.24, 0.00\}$	0.0042 $(0.0029, 0.0054)$ $\{0.22, 0.66\}$	0.0043 $(0.0030, 0.0057)$ $\{0.20, 0.68\}$
σ_s	0.0050	0.0053 $(0.0040, 0.0067)$ $\{0.15, 0.92\}$	0.0052 $(0.0042, 0.0068)$ $\{0.15, 0.90\}$	0.0051 $(0.0041, 0.0062)$ $\{0.14, 0.94\}$	0.0050 $(0.0033, 0.0063)$ $\{0.18, 0.00\}$	0.0048 $(0.0039, 0.0059)$ $\{0.12, 0.96\}$	0.0047 $(0.0037, 0.0061)$ $\{0.15, 0.92\}$
σ_i	0.0020	0.0017 $(0.0015, 0.0020)$ $\{0.16, 0.50\}$	0.0016 $(0.0014, 0.0019)$ $\{0.20, 0.20\}$	0.0017 $(0.0014, 0.0020)$ $\{0.17, 0.44\}$	0.0016 $(0.0012, 0.0019)$ $\{0.22, 0.00\}$	0.0016 $(0.0014, 0.0020)$ $\{0.19, 0.32\}$	0.0016 $(0.0014, 0.0019)$ $\{0.20, 0.28\}$
ϕ_π	2.0	2.04 $(1.88, 2.20)$ $\{0.06, 0.98\}$	2.00 $(1.72, 2.21)$ $\{0.07, 1.00\}$	1.89 $(1.67, 2.09)$ $\{0.08, 1.00\}$	1.83 $(1.62, 2.09)$ $\{0.11, 0.00\}$	1.72 $(1.52, 1.93)$ $\{0.16, 0.78\}$	1.73 $(1.52, 1.91)$ $\{0.15, 0.78\}$
ϕ_y	0.5	0.35 $(0.22, 0.54)$ $\{0.35, 0.80\}$	0.36 $(0.21, 0.56)$ $\{0.36, 0.84\}$	0.33 $(0.14, 0.54)$ $\{0.42, 0.70\}$	0.31 $(0.18, 0.50)$ $\{0.43, 0.00\}$	0.31 $(0.19, 0.45)$ $\{0.42, 0.66\}$	0.32 $(0.17, 0.47)$ $\{0.40, 0.76\}$
Σ		1.88	2.01	2.11	2.27	2.28	2.28

Table 9: Average, $(5, 95)$ percentiles, and $\{\mathrm{NRMSE}, \mathrm{CR}\}$. Σ is the sum of the NRMSE across the parameters.

E.2 SHORTER ZLB DURATIONS The paper focuses on the accuracy of NL-PF and OB-IF in datasets with either no ZLB events or a single 30 quarter event. This section shows the results when the ZLB binds for durations that are shorter than 30 quarters. We show the NRMSE for each estimated parameter as well as the sum of the NRMSE to measure overall accuracy. Table 8 shows the results with NL-PF-5% and OB-IF-0%, while table 9 focuses on Lin-KF-0% and Lin-KF-5%.

Ptr	Truth	No Misspecification: DGP and Estimation Use Small-Scale Model NL-PF-5%		OB-IF-0%		Lin-KF-5%	
		0Q	30Q	0Q	30Q	0Q	30Q
φ_p	100	96.8 $(81.6, 109.9)$ $\{0.09, 0.96\}$	109.8 $(89.5, 130.3)$ $\{0.15, 0.90\}$	94.3 $(81.8, 108.3)$ $\{0.11, 0.96\}$	110.6 $(95.3, 125.1)$ $\{0.15, 0.96\}$	103.7 $(92.6, 118.4)$ $\{0.09, 0.98\}$	128.5 $(111.2, 145.3)$ $\{0.30, 0.46\}$
h	0.8	0.79 $(0.76, 0.82)$ $\{0.02, 0.94\}$	0.79 $(0.77, 0.82)$ $\{0.02, 0.94\}$	0.79 $(0.75, 0.82)$ $\{0.02, 0.92\}$	0.79 $(0.77, 0.82)$ $\{0.02, 0.96\}$	0.80 $(0.76, 0.83)$ $\{0.02, 0.96\}$	0.79 $(0.76, 0.82)$ $\{0.03, 0.92\}$
ρ_s	0.8	0.80 $(0.76, 0.83)$ $\{0.03, 0.96\}$	0.83 $(0.78, 0.86)$ $\{0.04, 0.60\}$	0.81 $(0.76, 0.85)$ $\{0.04, 0.98\}$	0.84 $(0.80, 0.87)$ $\{0.06, 0.58\}$	0.82 $(0.77, 0.86)$ $\{0.05, 0.90\}$	0.87 $(0.83, 0.91)$ $\{0.10, 0.10\}$
ρ_i	0.8	0.82 $(0.79, 0.84)$ $\{0.03, 0.88\}$	0.82 $(0.78, 0.85)$ $\{0.03, 0.80\}$	0.79 $(0.77, 0.82)$ $\{0.02, 0.98\}$	0.79 $(0.74, 0.82)$ $\{0.03, 0.90\}$	0.82 $(0.79, 0.84)$ $\{0.03, 0.94\}$	0.86 $(0.83, 0.88)$ $\{0.08, 0.26\}$
σ_z	0.005	0.0037 $(0.0029, 0.0046)$ $\{0.27, 0.24\}$	0.0035 $(0.0025, 0.0045)$ $\{0.33, 0.18\}$	0.0051 $(0.0044, 0.0056)$ $\{0.08, 0.98\}$	0.0052 $(0.0043, 0.0061)$ $\{0.11, 0.86\}$	0.0038 $(0.0029, 0.0046)$ $\{0.26, 0.28\}$	0.0034 $(0.0026, 0.0044)$ $\{0.33, 0.16\}$
σ_s	0.005	0.0047 $(0.0035, 0.0058)$ $\{0.19, 0.90\}$	0.0043 $(0.0032, 0.0058)$ $\{0.22, 0.72\}$	0.0049 $(0.0039, 0.0060)$ $\{0.16, 0.86\}$	0.0046 $(0.0034, 0.0057)$ $\{0.17, 0.80\}$	0.0047 $(0.0034, 0.0059)$ $\{0.21, 0.90\}$	0.0036 $(0.0027, 0.0046)$ $\{0.32, 0.38\}$
σ_i	0.002	0.0016 $(0.0013, 0.0020)$ $\{0.20, 0.24\}$	0.0014 $(0.0010, 0.0018)$ $\{0.31, 0.18\}$	0.0020 $(0.0017, 0.0022)$ $\{0.07, 0.90\}$	0.0019 $(0.0016, 0.0022)$ $\{0.10, 0.78\}$	0.0016 $(0.0013, 0.0019)$ $\{0.20, 0.24\}$	0.0015 $(0.0012, 0.0017)$ $\{0.27, 0.10\}$
ϕ_π	2.0	2.00 $(1.81, 2.21)$ $\{0.06, 0.96\}$	2.01 $(1.82, 2.20)$ $\{0.06, 1.00\}$	1.95 $(1.74, 2.14)$ $\{0.06, 1.00\}$	1.80 $(1.58, 2.06)$ $\{0.12, 0.76\}$	1.97 $(1.76, 2.18)$ $\{0.07, 0.96\}$	1.62 $(1.42, 1.86)$ $\{0.20, 0.38\}$
ϕ_y	0.5	0.45 $(0.29, 0.61)$ $\{0.22, 1.00\}$	0.48 $(0.28, 0.61)$ $\{0.18, 1.00\}$	0.46 $(0.30, 0.63)$ $\{0.21, 1.00\}$	0.52 $(0.32, 0.73)$ $\{0.23, 1.00\}$	0.46 $(0.31, 0.63)$ $\{0.22, 1.00\}$	0.50 $(0.34, 0.66)$ $\{0.19, 1.00\}$
Σ		1.12	1.35	0.78	0.99	1.14	1.82

Table 10: Average, $(5, 95)$ percentiles, and $\{\mathrm{NRMSE}, \mathrm{CR}\}$. Σ is the sum of the NRMSE across the parameters.

E.3 NO MISSPECIFICATION Table 10 compares the parameter estimates after removing model misspecification. Since it is numerically very expensive to estimate the medium-scale model used to generate the data with NL-PF, we created new datasets from the small-scale model. The sum of the NRMSE shows about 40% of the error is due to model misspecification. For example, in datasets without any ZLB events, the error with NL-PF-5% increases from 1.12 to 1.90 when misspecification is added to the estimated model. Removing misspecification has the largest impact on the accuracy of φ_p, h, and ϕ_y because the estimates no longer have to compensate for the lack of sticky wages and investment, which creates large differences in the model's sensitivity to shocks. Notably, the NL-PF-5% estimate of φ_p declines from 151.1 to 96.8 and the estimate of h rises from 0.66 to 0.79 in datasets without ZLB events. The CR rises from near 0 to consistently above 0.9.

The other results emphasized in the paper are unchanged. The shock standard deviations are biased downward with NL-PF-5% because the filter incorrectly assigns some of the fluctuations to ME, reducing the estimated variances. When the ZLB binds in the data, it biases the estimates of φ_p and ρ_s upward, though NL-PF-5% and OB-IF-0% are both far more accurate than Lin-KF-5%.

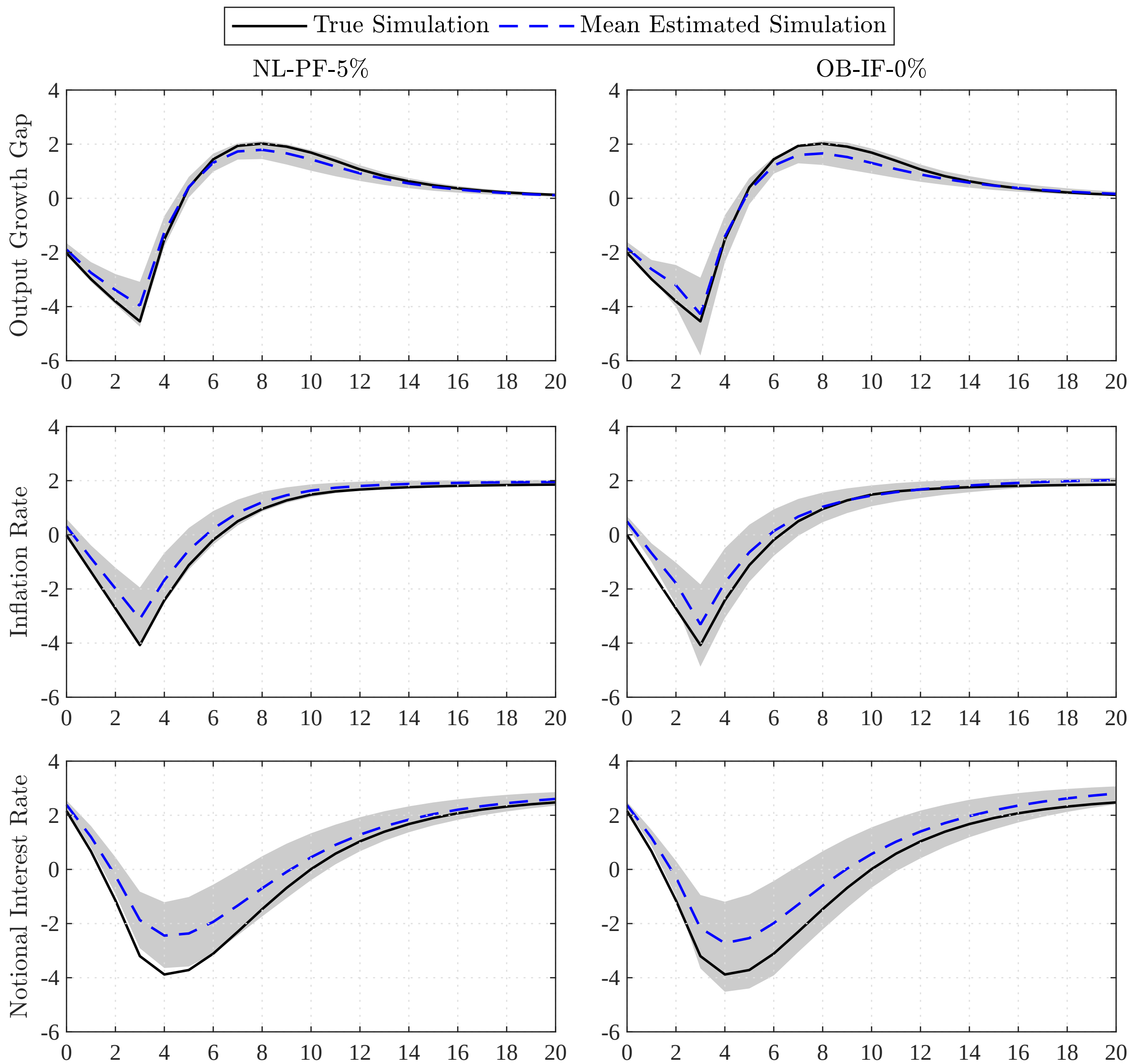


Figure 6: Recession responses without model misspecification. The solid line is the true simulation, the dashed line is the mean estimated simulation, and the shaded area contains the $(5, 95)$ percentiles across the datasets. The simulations are initialized in steady state and followed by four consecutive 1.5 standard deviation positive risk premium shocks.

Figure 6 plots the recession responses in figure 3 without misspecification. The solid line shows the responses based on the true parameterization of the small-scale model, rather than the medium-scale model that generates our original datasets. The dashed line shows the mean responses, given the parameter estimates with our alternative datasets. Consistent with the previous results, the responses based on the NL-PF-5% and OB-IF-0% parameter estimates are very similar. The key difference is that the mean estimated simulations are much closer to the true simulation and the $(5, 95)$ percentiles almost always encompass the truth. This result shows the muted responses in figure 3 are primarily driven by model misspecification, rather than inaccuracies in the estimation methods.

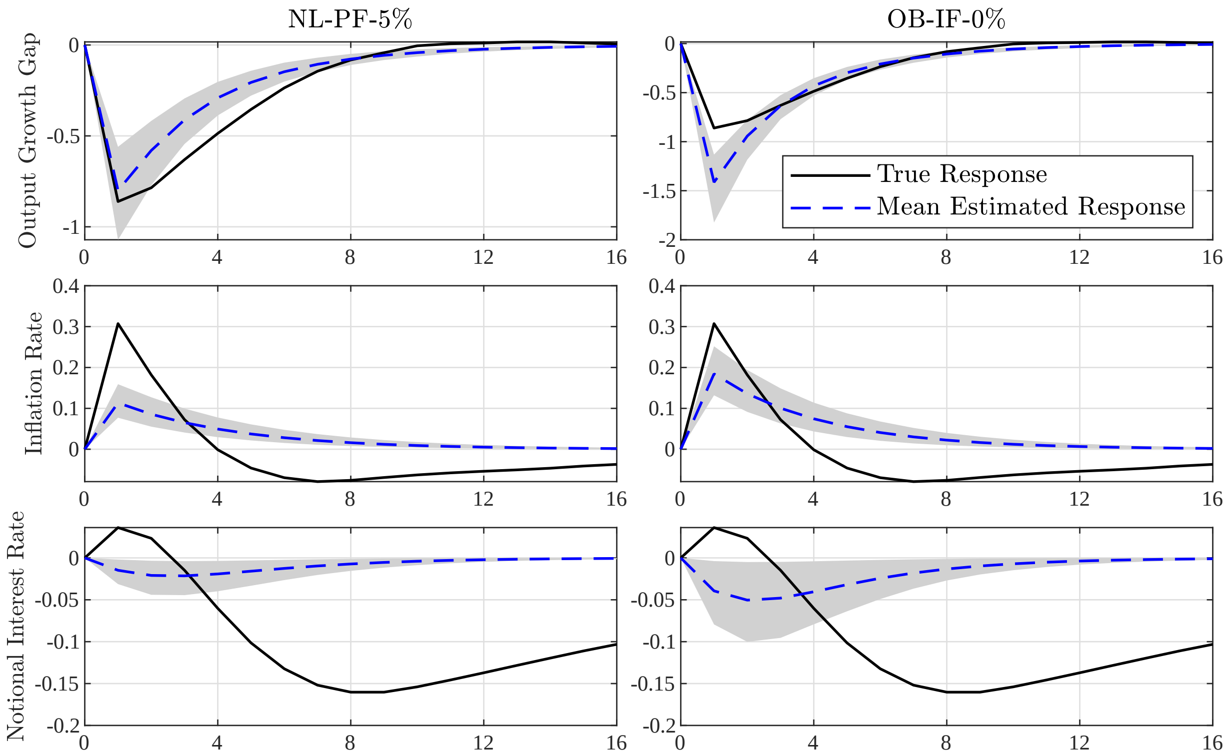
E.4 IMPULSE RESPONSES This section shows generalized impulse response functions (GIRFs) of a productivity growth and monetary policy shock when the economy is in a severe recession and the ZLB binds. To compute the GIRFs, we follow Koop et al. (1996). We first calculate the mean of $10{,}000$ simulations, conditional on random shocks in every quarter (i.e., the baseline path). We then calculate a second mean from another set of $10{,}000$ simulations, but this time the shock in the first quarter is replaced with a two standard deviation negative productivity growth or monetary policy shock (i.e., the impulse path). Finally, we plot the differences between the two mean paths.

The benefit of a GIRF over a more traditional impulse response function is that it allows us to calculate the responses in any state of the economy without the influence of mean reversion. For the true model, we initialize at the state following four consecutive 1.5 standard deviation positive risk premium shocks, consistent with figure 3. We then find a sequence of four equally sized risk premium shocks that produce the same notional rate in our estimated model as the true model, so the simulations begin at the same point. The NL-PF-5% simulations are shown in the left column and the OB-IF-0% simulations are in the right column. The true simulation of the DGP (solid line) is compared to the mean estimated simulation of the small-scale model (dashed line). The $(5, 95)$ percentiles account for differences in the simulations across the parameter estimates for each dataset.

Figure 7a shows the responses to a productivity growth shock. Qualitatively the responses of output growth and inflation are similar across the specifications. Higher productivity growth increases the output growth gap and decreases the inflation rate like a typical supply shock. Since the Fed faces a tradeoff between stabilizing the inflation and output gaps, the notional interest rate response depends on the parameterization. The notional rate rises with the DGP, but falls with both of the estimated models. Quantitatively, there are important differences between all of the responses. Consistent with figure 3, model misspecification leads to muted responses of the output growth gap and the inflation rate. There are also differences in the magnitudes of the estimated responses, but most of that is driven by the downward bias in the shock standard deviation with NL-PF-5%.

Figure 7b shows the responses to a monetary policy shock. Although the ZLB binds in the true and estimated models, the shock is expansionary because it lowers the expected nominal interest rate in future periods. Therefore, the output growth gap and the inflation rate both increase in all three models. Unlike with the other two shocks, model misspecification has a relatively small effect on the responses, as the $(5, 95)$ percentiles of the estimated responses encompass the true responses in most periods. There are some differences in the NL-PF-5% and OB-IF-0% responses, but they are smaller than in figure 7a and are never large enough to have meaningful policy implications.

(a) Productivity Growth Shock



(b) Monetary Policy Shock

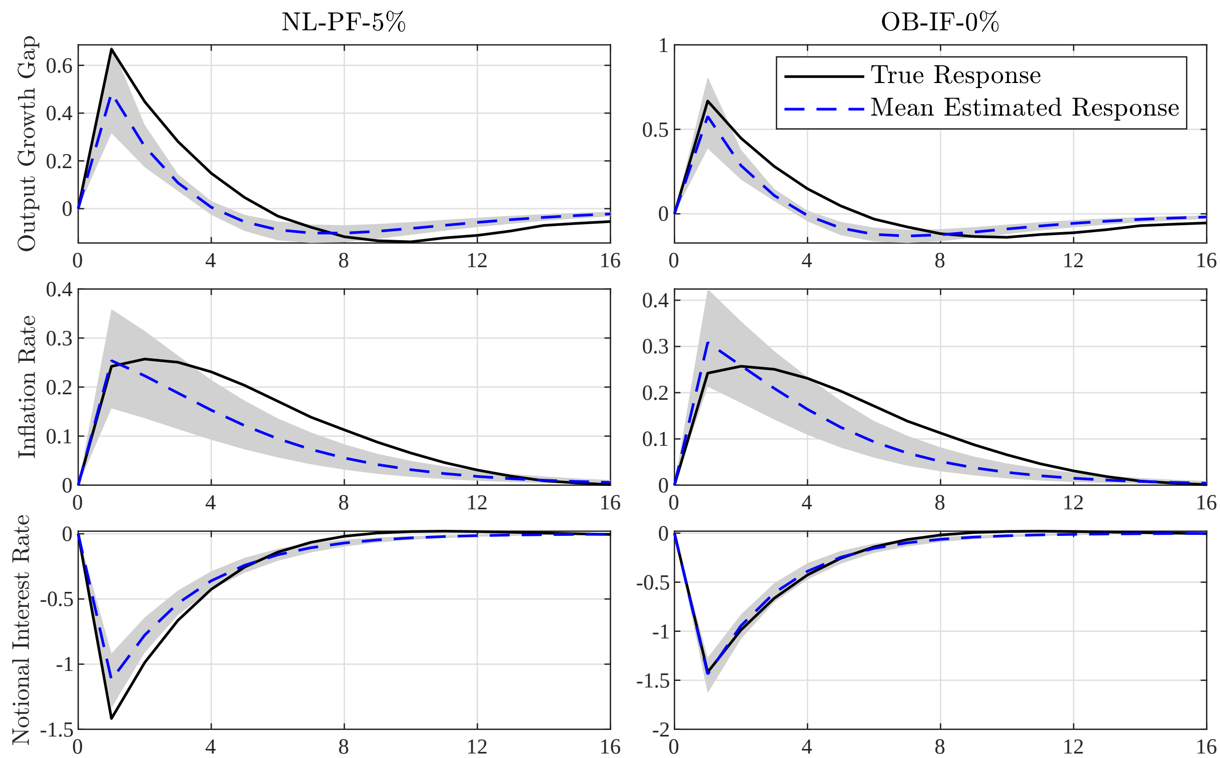


Figure 7: Impulse responses to a -2 standard deviation shock in a severe recession. The solid line is the true response, the dashed line is the mean estimated response, and the shaded area contains the $(5, 95)$ percentiles of the responses.

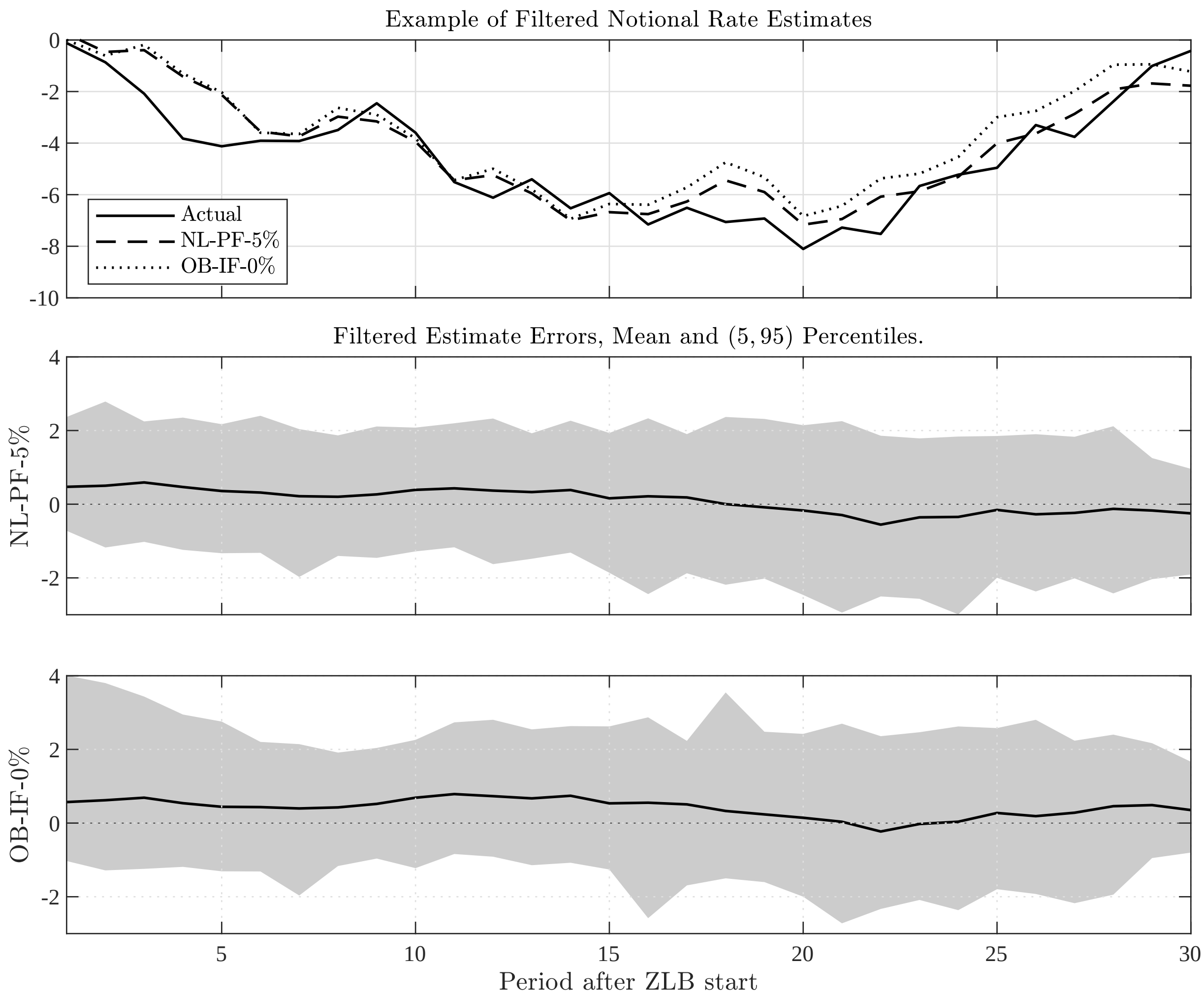


Figure 8: Estimates of the notional rate in datasets with a 30 quarter ZLB event. Rates are net annualized percentages.

E.5 NOTIONAL RATE ESTIMATES Figure 8 provides more intuition about what is driving the relative accuracy of the filtered estimates of the notional rate in figure 1. The top panel plots the actual notional rate from an example dataset with a 30 quarter ZLB event, as well as the filtered estimates from NL-PF-5% and OB-IF-0%. Over time, the OB-IF-0% estimate increases towards zero faster than NL-PF-5%. This may be driven by the lower estimate of ρ_i (0.77) with OB-IF-0%, which is slightly below the NL-PF-5% estimate and the true value (0.80). The bottom two panels plot the error in the average filtered notional rate estimates during the 30 quarter ZLB event across the 50 datasets (solid line). The shaded region shows the $(5, 95)$ percentiles. This suggests the example dataset in the top panel is fairly representative. The distribution of errors for OB-IF-0% is slightly shifted up from the NL-PF-5% error distribution, and increasingly so over time. This may seem somewhat at odds with the results in figure 1, as OB-IF-0% is even less accurate relative to NL-PF-5% in the datasets with shorter ZLB events. However the OB-IF-0% estimates of ρ_i and ϕ_y have an even larger downward bias in datasets with shorter ZLB duration, as shown in table 8.

	6Q	12Q	18Q	24Q	30Q
CDF of ZLB Event Durations	0.678	0.885	0.966	0.992	0.998
Number of Simulations to Reach 50 Datasets	150,300	154,950	256,950	391,950	1,030,300

Table 11: Probability of ZLB event durations in a long simulation of the medium-scale model.

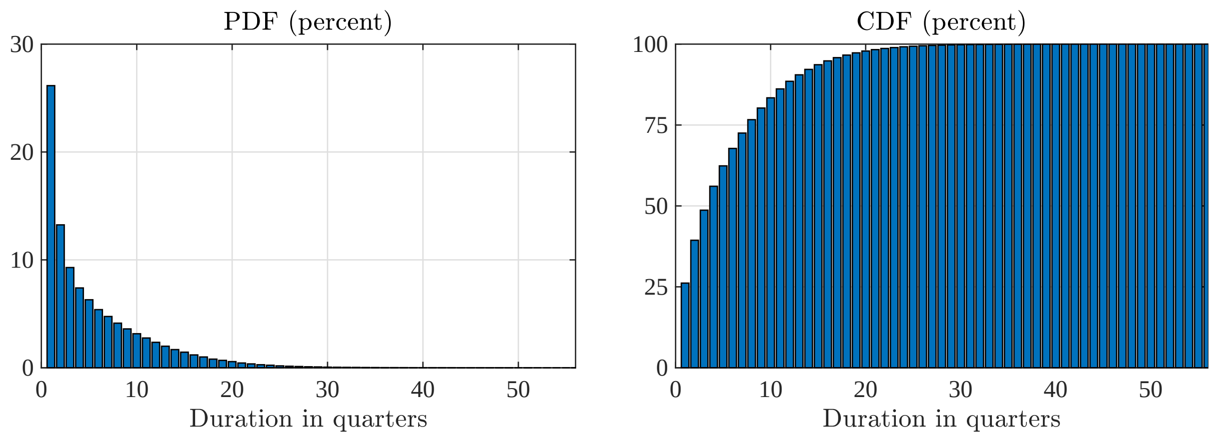


Figure 9: Duration of ZLB events in a long simulation of the medium-scale model.

E.6 ADDITIONAL DATASET STATISTICS ZLB events are frequent in the medium-scale model that generates the datasets, which allows us to find simulations with up to 30 consecutive quarters at the ZLB without imposing restrictions on the shocks. In a long simulation of the model, the unconditional probability of being at the ZLB is 24 percent. This is roughly equivalent to the U.S. experience of 7 years, since our sample is 30 years. Most of the ZLB events in the simulation are short, with the policy rate rising above zero within one year or less, as shown in table 11 and figure 9. However, long ZLB events are not incredibly uncommon, as 0.25 percent of ZLB events have a duration of at least 30 quarters. When generating our datasets, we impose an additional requirement that the ZLB event in our sample is unique so it reflects actual data. The number of 120 quarter simulations required to find 50 simulations that meet that criterion is shown in the last row of table 11.

E.7 GOVERNMENT SPENDING This section shows how government spending affects our results. Government spending is a potentially important feature because it adds a shock that directly enters the aggregate resource constraint. Without government spending, any shock in the DGP that affects the resource constraint is absorbed by consumption or price adjustment costs in the small-scale model, since output and inflation are observed. Without a wedge between consumption and output, it could cause significant bias in the habit persistence and price adjustment cost parameters.

We assume the share of government spending devoted to output follows

$$g_t^s = (1-\rho_g)\bar{g}^s + \rho_g g_{t-1}^s + \sigma_g \varepsilon_{g,t}, 0 \leq \rho_g < 1, \varepsilon_g \sim \mathbb{N}(0,1), \tag{46}$$

where the steady-state share, $\bar{g}^s$, is set to 0.2129 to match the time average from 1988Q1-2017Q4.

With the addition of government spending, the aggregate resource constraint is given by

$$c_t + x_t = (1 - g_t^s) y_t^{gdp}. \tag{47}$$

All other equations in the equilibrium system are unchanged. We add government spending to the medium-scale model that generates our datasets and our small-scale model for estimation. We estimate the small-scale model with (g^s-4obs) and without (g^s-3obs) including real per capita consumption growth as an additional observable. In this second specification, the government spending shock is less constrained, potentially absorbing the adjustment costs left out of the small-scale model and reducing inaccuracy driven by misspecification in the aggregate resource constraint. The specification without government spending (no-g^s) excludes g^s from the DGP and the estimated model, just like in the main paper. In each case, the true parameterization is unchanged, except the shock standard deviations were reduced from 0.005 to 0.004. This change is necessary because the additional volatility in the model with government spending causes the model to spend too much time at the ZLB and not converge at the previous parameterization. Table 12 shows the parameter estimates using datasets where the ZLB binds for 30 quarters and table 13 is based on datasets where the ZLB never binds in the data. OB-IF-0% is not used to estimate these specifications, since it is not possible to have more shocks than observables in the inversion filter.

Interestingly, the differences in the parameter estimates between g^s-4obs and no-g^s are fairly small, especially in datasets where the ZLB binds for 30 quarters. The g^s-4obs estimates of φ_p and h are more accurate than the no-g^s estimates, but they are still significantly biased. Furthermore, the improvement in those estimates is not as significant as what occurs when we add sticky wages to the model estimated with OB-IF-0%. This implies that the presence of government spending helps increase the volatility of output growth, but not enough to compensate for the lack of sticky wages, which we see as the most important misspecification driving the bias in φ_p and h. It is also important to note that the estimates of the productivity growth and risk premium shock standard deviations (σ_z and σ_s) are biased downward to a greater extent than in the model without government spending. As a consequence, the sum of the NRMSE with government spending is higher than without government spending, regardless of the estimation method or the duration of the ZLB. This result occurs even though the g^s-4obs estimates included an additional observable.

Excluding the additional observable (g^s-3obs) also does not improve the overall accuracy of the parameter estimates. The productivity growth and risk premium shock standard deviations become more accurate than no-g^s, but the estimates of φ_p are largely unchanged and the downward bias in h becomes even larger. As a result, the NRMSE of g^s-3obs is higher than the g^s-4obs or no-g^s estimates. Once again, this is consistent with the lack of sticky wages as the most important misspecification, while misspecification in the resource constraint appears to play a smaller role.

		NL-PF-5% (30Q)			Lin-KF-5% (30Q)		
Ptr	Truth	no-g^s	g^s-4obs	g^s-3obs	no-g^s	g^s-4obs	g^s-3obs
φ_p	100	180.8 $(167.2, 193.5)$ $\{0.81, 0.00\}$	164.2 $(145.1, 188.9)$ $\{0.65, 0.06\}$	183.3 $(165.2, 203.5)$ $\{0.84, 0.00\}$	182.8 $(168.0, 194.5)$ $\{0.83, 0.00\}$	170.0 $(150.3, 196.3)$ $\{0.71, 0.00\}$	188.6 $(167.6, 210.5)$ $\{0.90, 0.00\}$
h	0.8	0.66 $(0.63, 0.71)$ $\{0.17, 0.00\}$	0.71 $(0.67, 0.74)$ $\{0.11, 0.00\}$	0.56 $(0.47, 0.62)$ $\{0.31, 0.00\}$	0.65 $(0.62, 0.70)$ $\{0.18, 0.00\}$	0.71 $(0.66, 0.74)$ $\{0.12, 0.00\}$	0.54 $(0.43, 0.61)$ $\{0.33, 0.00\}$
ρ_s	0.8	0.84 $(0.81, 0.86)$ $\{0.05, 0.48\}$	0.86 $(0.84, 0.88)$ $\{0.08, 0.10\}$	0.84 $(0.80, 0.87)$ $\{0.05, 0.62\}$	0.85 $(0.82, 0.87)$ $\{0.06, 0.36\}$	0.87 $(0.85, 0.90)$ $\{0.09, 0.12\}$	0.84 $(0.81, 0.88)$ $\{0.06, 0.58\}$
ρ_i	0.8	0.81 $(0.78, 0.84)$ $\{0.03, 0.94\}$	0.81 $(0.77, 0.84)$ $\{0.03, 0.96\}$	0.81 $(0.77, 0.85)$ $\{0.03, 0.92\}$	0.83 $(0.80, 0.86)$ $\{0.04, 0.80\}$	0.83 $(0.80, 0.88)$ $\{0.05, 0.70\}$	0.85 $(0.81, 0.89)$ $\{0.07, 0.28\}$
ρ_{gs}	0.8	–	0.89 $(0.85, 0.93)$ $\{0.12, 0.28\}$	0.82 $(0.80, 0.84)$ $\{0.03, 1.00\}$	–	0.89 $(0.85, 0.93)$ $\{0.12, 0.20\}$	0.83 $(0.82, 0.86)$ $\{0.04, 1.00\}$
σ_z	0.004	0.0030 $(0.0023, 0.0037)$ $\{0.26, 0.40\}$	0.0028 $(0.0019, 0.0037)$ $\{0.33, 0.20\}$	0.0034 $(0.0026, 0.0047)$ $\{0.21, 0.94\}$	0.0031 $(0.0024, 0.0038)$ $\{0.25, 0.40\}$	0.0029 $(0.0021, 0.0041)$ $\{0.30, 0.28\}$	0.0036 $(0.0025, 0.0052)$ $\{0.22, 0.88\}$
σ_s	0.004	0.0031 $(0.0025, 0.0039)$ $\{0.25, 0.50\}$	0.0024 $(0.0020, 0.0030)$ $\{0.40, 0.04\}$	0.0036 $(0.0026, 0.0049)$ $\{0.20, 0.82\}$	0.0029 $(0.0023, 0.0036)$ $\{0.30, 0.26\}$	0.0023 $(0.0018, 0.0029)$ $\{0.44, 0.00\}$	0.0034 $(0.0025, 0.0047)$ $\{0.22, 0.70\}$
σ_i	0.002	0.0015 $(0.0013, 0.0018)$ $\{0.24, 0.22\}$	0.0015 $(0.0011, 0.0018)$ $\{0.26, 0.28\}$	0.0015 $(0.0011, 0.0017)$ $\{0.29, 0.22\}$	0.0014 $(0.0011, 0.0016)$ $\{0.33, 0.00\}$	0.0015 $(0.0012, 0.0017)$ $\{0.26, 0.10\}$	0.0015 $(0.0012, 0.0017)$ $\{0.27, 0.10\}$
σ_g	0.004	–	0.0044 $(0.0039, 0.0049)$ $\{0.13, 0.74\}$	0.0025 $(0.0018, 0.0032)$ $\{0.39, 0.16\}$	–	0.0044 $(0.0039, 0.0049)$ $\{0.13, 0.70\}$	0.0025 $(0.0018, 0.0033)$ $\{0.40, 0.20\}$
ϕ_π	2.0	2.27 $(2.13, 2.47)$ $\{0.14, 0.64\}$	2.09 $(1.85, 2.34)$ $\{0.08, 0.90\}$	2.23 $(2.00, 2.45)$ $\{0.13, 0.68\}$	2.10 $(1.91, 2.32)$ $\{0.08, 0.92\}$	1.73 $(1.31, 2.04)$ $\{0.17, 0.72\}$	1.90 $(1.62, 2.13)$ $\{0.09, 0.96\}$
ϕ_y	0.5	0.38 $(0.26, 0.55)$ $\{0.29, 0.98\}$	0.50 $(0.34, 0.63)$ $\{0.18, 0.98\}$	0.47 $(0.24, 0.64)$ $\{0.21, 0.96\}$	0.36 $(0.22, 0.51)$ $\{0.33, 0.94\}$	0.41 $(0.30, 0.58)$ $\{0.25, 0.98\}$	0.44 $(0.31, 0.64)$ $\{0.23, 0.98\}$
Σ		2.26	2.38	2.70	2.39	2.64	2.83

Table 12: Average, $(5, 95)$ percentiles, and $\{\text{NRMSE}, \text{CR}\}$. Σ is the sum of the NRMSE across the parameters.

		NL-PF-5% (0Q)			Lin-KF-5% (0Q)		
Ptr	Truth	no-g^s	g^s-4obs	g^s-3obs	no-g^s	g^s-4obs	g^s-3obs
φ_p	100	157.9 $(130.0, 175.8)$ $\{0.59, 0.00\}$	128.8 $(109.2, 143.7)$ $\{0.31, 0.34\}$	148.8 $(128.8, 163.8)$ $\{0.50, 0.00\}$	157.7 $(130.1, 175.3)$ $\{0.59, 0.02\}$	128.8 $(109.5, 142.8)$ $\{0.31, 0.38\}$	149.2 $(129.4, 164.3)$ $\{0.50, 0.00\}$
h	0.8	0.64 $(0.60, 0.69)$ $\{0.20, 0.00\}$	0.68 $(0.65, 0.72)$ $\{0.15, 0.00\}$	0.57 $(0.47, 0.66)$ $\{0.30, 0.00\}$	0.64 $(0.60, 0.69)$ $\{0.20, 0.00\}$	0.68 $(0.65, 0.72)$ $\{0.15, 0.00\}$	0.57 $(0.48, 0.66)$ $\{0.29, 0.00\}$
ρ_s	0.8	0.79 $(0.74, 0.82)$ $\{0.03, 0.94\}$	0.81 $(0.76, 0.85)$ $\{0.03, 0.90\}$	0.78 $(0.72, 0.83)$ $\{0.05, 0.86\}$	0.79 $(0.74, 0.83)$ $\{0.03, 0.96\}$	0.81 $(0.77, 0.85)$ $\{0.04, 0.90\}$	0.78 $(0.72, 0.83)$ $\{0.05, 0.86\}$
ρ_i	0.8	0.79 $(0.74, 0.82)$ $\{0.04, 0.86\}$	0.78 $(0.74, 0.82)$ $\{0.04, 0.84\}$	0.80 $(0.76, 0.83)$ $\{0.03, 0.98\}$	0.79 $(0.74, 0.82)$ $\{0.04, 0.88\}$	0.78 $(0.75, 0.82)$ $\{0.03, 0.92\}$	0.80 $(0.76, 0.83)$ $\{0.03, 0.98\}$
ρ_{gs}	0.8	–	0.82 $(0.76, 0.87)$ $\{0.05, 0.94\}$	0.81 $(0.76, 0.84)$ $\{0.03, 1.00\}$	–	0.82 $(0.77, 0.86)$ $\{0.04, 0.94\}$	0.80 $(0.75, 0.83)$ $\{0.03, 1.00\}$
σ_z	0.004	0.0029 $(0.0022, 0.0037)$ $\{0.29, 0.22\}$	0.0023 $(0.0018, 0.0029)$ $\{0.43, 0.00\}$	0.0027 $(0.0019, 0.0036)$ $\{0.36, 0.54\}$	0.0029 $(0.0022, 0.0037)$ $\{0.29, 0.28\}$	0.0023 $(0.0018, 0.0029)$ $\{0.43, 0.00\}$	0.0027 $(0.0019, 0.0036)$ $\{0.36, 0.50\}$
σ_s	0.004	0.0032 $(0.0025, 0.0038)$ $\{0.23, 0.52\}$	0.0025 $(0.0021, 0.0030)$ $\{0.38, 0.02\}$	0.0036 $(0.0026, 0.0049)$ $\{0.19, 0.84\}$	0.0032 $(0.0025, 0.0039)$ $\{0.23, 0.54\}$	0.0025 $(0.0020, 0.0030)$ $\{0.38, 0.02\}$	0.0037 $(0.0027, 0.0049)$ $\{0.19, 0.84\}$
σ_i	0.002	0.0018 $(0.0015, 0.0021)$ $\{0.15, 0.60\}$	0.0018 $(0.0015, 0.0021)$ $\{0.15, 0.60\}$	0.0017 $(0.0014, 0.0020)$ $\{0.17, 0.48\}$	0.0018 $(0.0015, 0.0021)$ $\{0.15, 0.62\}$	0.0018 $(0.0015, 0.0020)$ $\{0.15, 0.56\}$	0.0017 $(0.0014, 0.0020)$ $\{0.16, 0.50\}$
σ_g	0.004	–	0.0041 $(0.0037, 0.0046)$ $\{0.08, 0.84\}$	0.0033 $(0.0025, 0.0039)$ $\{0.20, 0.52\}$	–	0.0041 $(0.0036, 0.0046)$ $\{0.08, 0.84\}$	0.0033 $(0.0025, 0.0038)$ $\{0.20, 0.56\}$
ϕ_π	2.0	2.11 $(1.97, 2.24)$ $\{0.07, 1.00\}$	1.92 $(1.67, 2.25)$ $\{0.09, 1.00\}$	2.08 $(1.87, 2.34)$ $\{0.08, 0.94\}$	2.10 $(1.97, 2.24)$ $\{0.07, 0.98\}$	1.92 $(1.66, 2.27)$ $\{0.09, 0.98\}$	2.08 $(1.86, 2.32)$ $\{0.08, 0.96\}$
ϕ_y	0.5	0.39 $(0.26, 0.53)$ $\{0.26, 1.00\}$	0.53 $(0.34, 0.70)$ $\{0.22, 0.98\}$	0.52 $(0.30, 0.69)$ $\{0.23, 1.00\}$	0.39 $(0.27, 0.52)$ $\{0.27, 1.00\}$	0.53 $(0.34, 0.70)$ $\{0.22, 0.98\}$	0.52 $(0.30, 0.68)$ $\{0.23, 1.00\}$
Σ		1.87	1.92	2.12	1.86	1.91	2.12

Table 13: Average, $(5, 95)$ percentiles, and $\{\mathrm{NRMSE}, \mathrm{CR}\}$. Σ is the sum of the NRMSE across the parameters.