

**The Accuracy of Linear and
Nonlinear Estimation in the
Presence of the Zero Lower
Bound**

*Tyler Atkinson, Alexander W. Richter and
Nathaniel A. Throckmorton*

Federal Reserve Bank of Dallas
Research Department
Working Paper 1804
<https://doi.org/10.24149/wp1804>

The Accuracy of Linear and Nonlinear Estimation in the Presence of the Zero Lower Bound*

Tyler Atkinson Alexander W. Richter Nathaniel A. Throckmorton

May 7, 2018

ABSTRACT

This paper evaluates the accuracy of linear and nonlinear estimation methods for dynamic stochastic general equilibrium models. We generate a large sample of artificial datasets using a global solution to a nonlinear New Keynesian model with an occasionally binding zero lower bound (ZLB) constraint on the nominal interest rate. For each dataset, we estimate the nonlinear model—solved globally, accounting for the ZLB—and the linear analogue of the nonlinear model—solved locally, ignoring the ZLB—with a Metropolis-Hastings algorithm where the likelihood function is evaluated with a Kalman filter, unscented Kalman filter, or particle filter. In datasets that resemble the U.S. experience, the nonlinear model estimated with a particle filter is more accurate and has a higher marginal data density than the linear model estimated with a Kalman filter, as long as the measurement error variances in the particle filter are not too big.

Keywords: Bayesian Estimation; Nonlinear Solution; Particle Filter; Unscented Kalman 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 (nathrockmorton@wm.edu). We thank Chris Stackpole and Eric Walter for supporting the supercomputers at our institutions. We also thank Auburn University and the University of Texas at Dallas for providing access to their supercomputers. The views expressed 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 evaluate the likelihood function exactly 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 extended 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)).

Recent empirical work uses a variety of methods to deal with the nonlinearity imposed by the ZLB constraint.² Several papers exclude the post-crisis period from their estimation, skirting the nonlinearity and justifying the continued use of linear methods (e.g., Aruoba et al. (2018); Christiano et al. (2015); Cuba-Borda (2014); Del Negro et al. (2015); Galí et al. (2012)). While that is a reasonable approach, at some point it is necessary to incorporate the more recent data and its effects on the model parameters. Another approach is to ignore the constraint and estimate a linear model with all available data (e.g., Ireland (2011); Suh and Walker (2016)). That approach maintains the simplicity of the estimation procedure but potentially provides inaccurate estimates of the parameters and model predictions. Recognizing this concern, a few recent papers have estimated a fully nonlinear model with an occasionally binding ZLB constraint (e.g., Gust et al. (2017); Plante et al. (2018)).³ This method provides the most comprehensive treatment of the constraint, but it requires numerically intensive nonlinear solution and estimation techniques. Specifically, it uses projection methods to solve the model and a particle filter to evaluate the likelihood function with each draw.

This paper takes an important first step toward comparing the accuracy of linear and nonlinear estimation methods when the ZLB is present in the data. We specify a true parameterization of a nonlinear New Keynesian model with an occasionally binding ZLB constraint, solve the model

¹Schorfheide (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.

²See Fernández-Villaverde et al. (2016) for a detailed overview of the various solution and estimation methods.

³Several earlier papers study the effects of the ZLB constraint in calibrated nonlinear models using solution methods similar to the one for NL-PF (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)).

globally with a projection method, and generate a large sample of artificial datasets. The datasets include the most common observables—real GDP growth, inflation, and the nominal interest rate—as well as the true latent state variables and structural shocks. For each dataset, we estimate the nonlinear model—solved globally, accounting for the ZLB—and the linear analogue of the nonlinear model—solved locally, ignoring the ZLB—with a random-walk Metropolis-Hastings algorithm.

We combine the solution method with three different filters to evaluate the likelihood: the linear model estimated with a Kalman filter (Lin-KF), the nonlinear model estimated with an unscented Kalman filter (NL-UKF), and the nonlinear model estimated with a particle filter (NL-PF). Lin-KF serves as our benchmark since it is the most common method in the literature and the least computationally expensive. Given the posterior distributions, we calculate the marginal data density for each method to measure empirical fit and then determine how well each method is able to recover the true parameters and latent state variables. We also examine how the measurement error (ME) variances in the observation equation of the filter and the number of quarters that the ZLB binds in the data affects the accuracy of each method. We measure accuracy with the root-mean square-error (RMSE) of the parameter draws and filtered state variables relative to the true values. For each method and dataset, we calculate the RMSE statistic and then report quantiles across the datasets.

The federal funds rate was stuck at its ZLB from December 2008 to December 2015, or 28 of the last 100 quarters. In datasets that resemble the U.S. experience, NL-PF has a clear statistical advantage over Lin-KF, both in terms of empirical fit and accuracy. The differences in accuracy are large enough that impulse responses with the Lin-KF parameter estimates are significantly different from the true responses when the ZLB binds. Using the nonlinear solution with the Lin-KF estimates often makes the impulse responses even more inaccurate. Our results, however, require small ME variances in the particle filter (e.g., 5% of the variance in the data). Given high enough variances (e.g., 20% of the variance in the data), the benefits of using NL-PF over Lin-KF with no ME completely disappear. We conclude that the data must be highly nonlinear for NL-PF to overcome positive ME variances and have a statistically significant advantage over Lin-KF without ME.

One of the most important aspects of our results is that they serve as a benchmark for alternative solution and estimation methods. Comparisons of solution methods are typically based on speed tests and Euler equation errors, which are sometimes misleading because they are conditional on specific discretization and numerical integration methods.⁴ Also, while there is a lot of theoretical work on developing new estimation techniques, there is less comparison in terms of accuracy, especially with nonlinear methods. With our datasets, the accuracy and economic implications of any new solution method or aspect of the estimation procedure (e.g., a new filter or an alternative to the random walk Metropolis-Hastings algorithm) is easily compared against the methods examined

⁴Aruoba et al. (2006) compares various nonlinear solution methods in terms of speed and accuracy using a neo-classical growth model. Richter et al. (2014) draw comparisons using a New Keynesian model with a ZLB constraint.

in this paper. With this mode of comparison, a method is tested with the same empirical strategy applied to most DSGE models. If it turns out that a method is computationally more efficient and more accurate than the ones we examine, it should become the standard for future empirical work.

Our paper is similar in spirit to Fernández-Villaverde and Rubio-Ramírez (2005) who find 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 as well as actual data. The primary nonlinearity in their model is high risk aversion, whereas we study the implications of the ZLB constraint. The paper closest to ours is Hirose and Inoue (2016). They generate artificial datasets from a linear model where the ZLB constraint is imposed using anticipated policy shocks and then apply Lin-KF to estimate the model without the constraint. They find the estimated parameters, impulse response functions, and structural shocks become less accurate as the frequency and duration of ZLB events increase in the data. In contrast, we generate data using a global solution to a nonlinear model and consider alternative estimation methods that use nonlinear filters and global solutions.

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 recent ZLB period. Our contribution is to determine the accuracy of these nonlinear methods and show under what conditions they outperform other methods.

We find the ME variances in the particle filter play an important role in the estimation procedure. Positive ME variances are necessary to prevent degeneracy—a situation when all but a few particle weights are near zero. 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 Smets and Wouters (2007) model estimated with Lin-KF. To prevent degeneracy and increase the accuracy of the particle filter, most papers set large ME variances.⁵ Similar to our results, Cuba-Borda et al. (2017) show ME reduces the accuracy of the likelihood function using a calibrated model with an occasionally binding borrowing constraint. Herbst and Schorfheide (2017) develop a tempered particle filter that sequentially reduces the ME variances. They assess its accuracy against the Kalman filter on actual U.S. data with a linear DSGE model and find it outperforms the basic, unadapted, bootstrap particle filter. In our analysis, we set the ME variances and then compare the posterior estimates with several alternative values using a particle filter that adapts to the current observation.

The paper proceeds as follows. [Section 2](#) describes our data generating process, including the true model and parameters. [Section 3](#) outlines the estimation and computational procedures. [Section 4](#) defines our accuracy measures and reports the results of our estimation. [Section 5](#) concludes.

⁵Some 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)). Others directly set 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)).

2 DATA GENERATING PROCESS

To test the accuracy of recent linear and nonlinear estimation methods, we generate a large number of artificial datasets from a New Keynesian model with an occasionally binding ZLB constraint, where the frequency and duration of ZLB events controls the degree of nonlinearity in each dataset.

2.1 MODEL A representative household chooses $\{c_t, n_t, b_t\}_{t=0}^{\infty}$ to maximize expected lifetime utility, $E_0 \sum_{t=0}^{\infty} \beta^t a_t [\log(c_t - hc_{t-1}^a) - \chi n_t^{1+\eta} / (1+\eta)]$, where β is the subjective discount factor, χ determines the steady state labor supply, $1/\eta$ is the Frisch elasticity of labor supply, c is consumption, c^a is aggregate consumption, h is the degree of external habit persistence, n is labor hours, b is the real value of a privately-issued 1-period nominal bond, E_0 is the mathematical expectation operator conditional on information available in period 0, and a is a preference shock that follows

$$a_t = 1 - \rho_a + \rho_a a_{t-1} + \sigma_a \varepsilon_{a,t}, \quad 0 \leq \rho_a < 1, \quad \varepsilon_a \sim \mathbb{N}(0, 1). \quad (1)$$

An increase in a_t makes households more impatient, which increases demand in period t . The household's choices are constrained by $c_t + b_t / (i_t(1 + \bar{s})) = w_t n_t + b_{t-1} / \pi_t + d_t$, where π is the gross inflation rate, w is the real wage rate, i is the gross nominal interest rate, \bar{s} is the steady-state risk premium on the nominal bond, and d is a real dividend from ownership of intermediate firms.

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

$$\begin{aligned} a_t \lambda_t &= c_t - hc_{t-1}^a, \\ w_t &= \chi a_t n_t^\eta \lambda_t, \\ 1 &= \beta(1 + \bar{s}) E_t[(\lambda_t / \lambda_{t+1})(i_t / \pi_{t+1})]. \end{aligned}$$

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, $y_t(f)$, according to $y_t(f) = z_t n_t(f)$, where $n(f)$ is the labor hired by firm f and $z_t = g_t z_{t-1}$ is technology, which is common across firms. Deviations from the balanced growth rate, \bar{g} , follow

$$g_t = \bar{g} + \sigma_g \varepsilon_{g,t}, \quad \varepsilon_g \sim \mathbb{N}(0, 1). \quad (2)$$

The final goods firm purchases $y_t(f)$ units from each intermediate firm to produce the final good, $y_t \equiv [\int_0^1 y_t(f)^{(\epsilon-1)/\epsilon} df]^\epsilon / (\epsilon-1)$, where $\epsilon > 1$ is the elasticity of substitution. It then maximizes dividends to determine its demand function for intermediate good f , $y_t(f) = (p_t(f)/p_t)^{-\epsilon} y_t$, where $p_t = [\int_0^1 p_t(f)^{1-\epsilon} df]^{1/(1-\epsilon)}$ is the price level. Following Rotemberg (1982), each intermediate firm pays a price adjustment cost, $adj_t(f) \equiv \varphi(p_t(f)/(\bar{\pi} p_{t-1}(f)) - 1)^2 y_t / 2$, where $\varphi > 0$ scales the cost and $\bar{\pi}$ is the gross inflation rate along the balanced growth path. Given that functional form,

firm f chooses $n_t(f)$ 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 $d_t(f) = p_t(f)y_t(f)/p_t - w_t n_t(f) - adj_t(f)$. In symmetric equilibrium, all firms make identical decisions (i.e., $p_t(f) = p_t$, $n_t(f) = n_t$, and $y_t(f) = y_t$). Therefore, the optimality conditions imply

$$y_t = z_t n_t, \quad (3)$$

$$\varphi(\pi_t/\bar{\pi} - 1)(\pi_t/\bar{\pi}) = 1 - \epsilon + \epsilon w_t/z_t + \beta\varphi E_t[(\lambda_t/\lambda_{t+1})(\pi_{t+1}/\bar{\pi} - 1)(\pi_{t+1}/\bar{\pi})(y_{t+1}/y_t)]. \quad (4)$$

When $\varphi = 0$, $w_t/z_t = (\epsilon - 1)/\epsilon$, which is the inverse of a firm's markup of price over marginal cost.

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

$$i_t = \max\{1, i_t^n\}, \quad (5)$$

$$i_t^n = (i_{t-1}^n)^{\rho_i} (\bar{i}(\pi_t/\bar{\pi})^{\phi_\pi} (y_t^{gdp}/(y_{t-1}^{gdp}\bar{g}))^{\phi_y})^{1-\rho_i} \exp(\sigma_i \varepsilon_{i,t}), \quad 0 \leq \rho_i < 1, \quad \varepsilon_i \sim \mathbb{N}(0, 1), \quad (6)$$

where y^{gdp} is real GDP (the level of output minus the resources lost due to price adjustment costs), i^n is the gross notional interest rate, \bar{i} and $\bar{\pi}$ are the steady-state or target values of the inflation and nominal interest rates, and ϕ_π and ϕ_y determine the central bank's responses to deviations of inflation from the target rate and deviations of real GDP growth from the balanced growth rate. When the net notional rate is positive, $i_t = i_t^n$. When it is negative, the ZLB binds and $i_t = 1$. A more negative net notional rate means the central bank is more constrained and the model is more nonlinear.

The model does not possess a steady-state due to the unit root in technology, z_t . Therefore, we redefine the subset of variables with a trend in terms of technology (i.e., $\tilde{x}_t \equiv x_t/z_t$). The detrended equilibrium system includes the two stochastic processes, (1) and (2), the ZLB constraint, (5), and

$$\tilde{y}_t^g = g_t \tilde{y}_t^{gdp} / \tilde{y}_{t-1}^{gdp}, \quad (7)$$

$$a_t \tilde{\lambda}_t = \tilde{c}_t - h \tilde{c}_{t-1} / g_t, \quad (8)$$

$$\tilde{w}_t = \chi a_t \tilde{y}_t^n \tilde{\lambda}_t, \quad (9)$$

$$1 = \beta(1 + \bar{s}) E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(i_t/(\pi_{t+1} g_{t+1}))], \quad (10)$$

$$\varphi(\pi_t/\bar{\pi} - 1)(\pi_t/\bar{\pi}) = 1 - \epsilon + \epsilon \tilde{w}_t + \beta\varphi E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(\pi_{t+1}/\bar{\pi} - 1)(\pi_{t+1}/\bar{\pi})(\tilde{y}_{t+1}^g/\tilde{y}_t^g)], \quad (11)$$

$$\tilde{c}_t = (1 - \varphi(\pi_t/\bar{\pi} - 1)^2/2) \tilde{y}_t \equiv \tilde{y}_t^{gdp}, \quad (12)$$

$$i_t^n = (i_{t-1}^n)^{\rho_i} (\bar{i}(\pi_t/\bar{\pi})^{\phi_\pi} (\tilde{y}_t^g/\bar{g})^{\phi_y})^{1-\rho_i} \exp(\sigma_i \varepsilon_{i,t}). \quad (13)$$

A competitive equilibrium consists of infinite sequences of quantities, $\{\tilde{c}_t, \tilde{y}_t, \tilde{y}_t^g, \tilde{\lambda}_t\}_{t=0}^{\infty}$, prices, $\{\tilde{w}_t, i_t, i_t^n, \pi_t\}_{t=0}^{\infty}$, and exogenous variables, $\{a_t, g_t\}_{t=0}^{\infty}$, that satisfy the detrended equilibrium system, given the initial conditions, $\{\tilde{c}_{-1}, i_{-1}^n, a_0, g_0, \varepsilon_{i,0}\}$, and sequences of shocks, $\{\varepsilon_{g,t}, \varepsilon_{a,t}, \varepsilon_{i,t}\}_{t=1}^{\infty}$.

Subjective Discount Factor	β	0.9955	Inflation Gap Response	ϕ_π	2.0
Frisch Elasticity of Labor Supply	$1/\eta$	3	Output Gap Response	ϕ_y	0.5
Elasticity of Substitution	ϵ	6	Habit Persistence	h	0.5
Steady-State Labor Hours	\bar{n}	0.33	Preference Shock Persistence	ρ_a	0.8
Steady-State Risk Premium	\bar{s}	0.0060	Notional Rate Persistence	ρ_i	0.8
Steady-State Growth Rate	\bar{g}	1.0034	Growth Rate Shock SD	σ_g	0.0150
Steady-State Inflation Rate	$\bar{\pi}$	1.0048	Preference Shock SD	σ_a	0.0150
Rotemberg Price Adjustment Cost	φ	100	Notional Rate Shock SD	σ_i	0.0025

Table 1: Parameter values for the data generating process.

2.2 PARAMETER VALUES Table 1 shows the true model parameters. In general, the true parameters were chosen so our data generating process is characteristic of actual U.S. data. The steady-state growth rate (\bar{g}), inflation rate ($\bar{\pi}$), and risk-premium (\bar{s}) are equal to the time averages of per capita real GDP growth, the percent change in the GDP implicit price deflator, and the Baa corporate bond yield relative to the yield on the 10-Year Treasury from 1992Q1-2016Q4 (100 quarters).

The subjective discount factor, β , is set to 0.9955, which is the time average of the values implied by the steady-state consumption Euler equation and the federal funds rate. The corresponding annualized steady-state nominal interest rate is 2.7%, 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 elasticity of substitution between intermediate goods, ϵ , is set to 6, which matches the estimate in Christiano et al. (2005) and corresponds to a 20% average markup of price over marginal cost. The Frisch elasticity of labor supply, $1/\eta$, is set to 3 to match the macro estimate in Peterman (2016). The other parameters are set to round numbers that are in line with the posterior parameter estimates from similar models in the literature.

2.3 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 , g_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 log-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 real GDP growth, the inflation rate, and the nominal interest rate by simulating the model using the nonlinear policy functions, so the vector of observables is given by

$\mathbf{x}_t = [\tilde{y}_t^g, \pi_t, i_t]$. We also save the other variables in the equilibrium system to calculate the accuracy of the estimated states. Each simulation is initialized with a draw from the ergodic distribution and contains 104 quarters. We use the first 4 quarters as a training period when filtering the data, so our accuracy measures are based on the last 100 quarters of each data set. We chose this data length because it is similar to what most researchers use when estimating DSGE models with actual data.

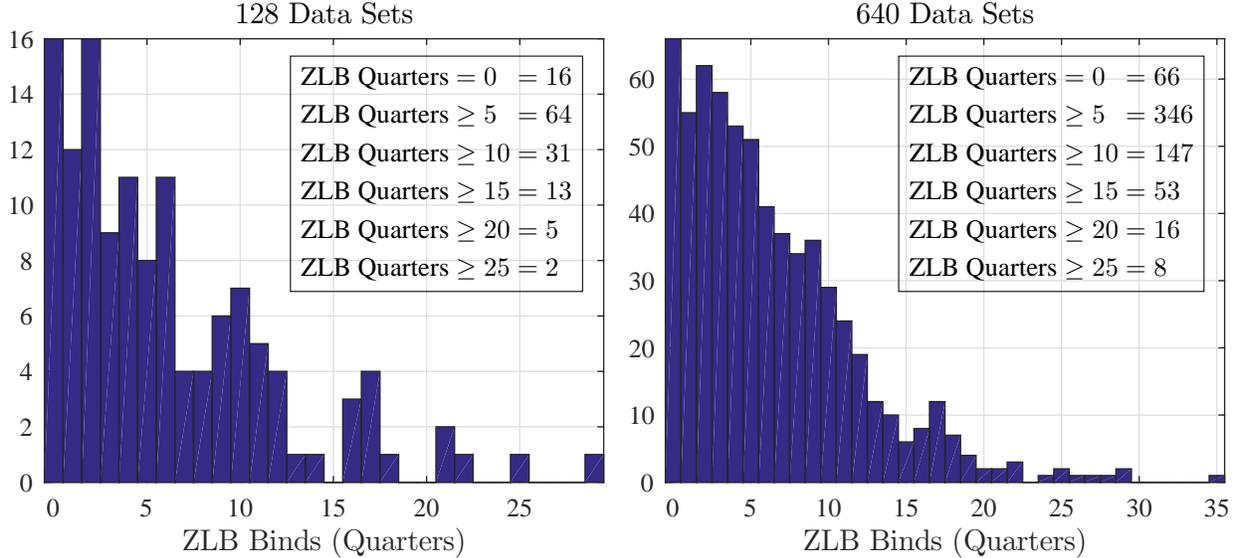


Figure 1: Probability mass functions of the number quarters the ZLB binds in our datasets.

We create 640 datasets to ensure large differences in their degree of nonlinearity. Figure 1 plots histograms of the number of quarters the ZLB binds in each dataset. Some of our results are based on a subset of the datasets, so we provide a histogram for the first 128 datasets and all 640 datasets. The fraction of quarters where the ZLB binds largely determines the degree of nonlinearity in the data, which will affect the accuracy of each estimation method. The ZLB binds in at least 1 quarter in 574 of the 640 datasets. On average, the ZLB binds in about 6 of the 100 quarters in each dataset. When we condition on datasets where the ZLB binds for at least 5 (10, 15, 25) quarters, the average increases to 10 (14.3, 19.2, 28) quarters. The average when the ZLB binds for at least 25 quarters equals the number of quarters that the federal funds rate was stuck at zero over the last 100 quarters.

Another important statistic is the duration of the longest ZLB event in a given dataset because it is a key indicator of the severity of the ZLB constraint. Across all datasets, the average duration of the longest ZLB event is about 3.3 quarters. Conditional on datasets with at least 5 (10, 15, 25) ZLB quarters, that value increases to about 5 (6.5, 8.5, 10.1) quarters. These values fall well short of the 28 quarter ZLB event in the U.S. from December 2008 to December 2015. However, recent estimates of the notional rate indicate it was close to zero for large portions of that period, suggesting the Fed was not constrained for all 28 quarters (Gust et al. (2017); Plante et al. (2018)). Datasets with longer durations than we consider would increase the importance of using nonlinear methods.

3 ESTIMATION METHODS

We estimate the model described in [section 2.1](#) with Bayesian methods. For each data set, we draw parameters from a proposal distribution, solve the model conditional on that draw, filter the data to evaluate the likelihood function, and use a random walk Metropolis-Hastings algorithm to determine whether to accept or reject the draw. We examine the accuracy of several estimation methods. First, we estimate the log-linear analogue of the nonlinear model, solved using Sims’s (2002) gensys algorithm, with a Kalman filter (Lin-KF). This method is by far the least computationally expensive and the most common approach in the literature for estimating DSGE models.

Second, we estimate the nonlinear model with a particle filter (NL-PF). We follow Algorithm 12 in Herbst and Schorfheide (2016) and adapt the 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 far better equipped to handle the nonlinearities in the data, but it also imposes a much larger computational burden than Lin-KF because it requires us to solve the nonlinear model for each draw from the proposal distribution.

Finally, as a potential middle-ground between Lin-KF and NL-PF, we estimate the nonlinear model with an unscented Kalman filter (NL-UKF) following Julier and Uhlmann (1997). The unscented Kalman filter modifies the basic Kalman filter for nonlinear models. It propagates a deterministic set of points through the nonlinear state equation to approximate the mean and variance of the conditional state distribution, whereas the particle filter approximates the entire distribution using a much larger number of points. Therefore, the unscented Kalman filter is faster than the particle filter but potentially less accurate. [Appendix C](#) describes each estimation procedure in detail.

The measurement equation is independent of the solution method and given by $\mathbf{x}_t = H\mathbf{s}_t + \xi_t$, where $\mathbf{s}_t = [\tilde{c}_t, \tilde{y}_t, \tilde{y}_t^g, \tilde{\lambda}_t, \tilde{w}_t, i_t, i_t^n, \pi_t, a_t, g_t]'$ is the state of the system, H is an observable selection matrix, and $\xi \sim \mathbb{N}(0, R)$ is a vector of measurement errors (MEs) with covariance matrix R . The primary difference between estimating the linear and nonlinear models is the state equation. In the former case, it is given by $\mathbf{s}_t = T(\vartheta)\mathbf{s}_{t-1} + M(\vartheta)\varepsilon_t$ and in the latter case it is given by $\mathbf{s}_t = \Psi(\vartheta, \mathbf{s}_{t-1}, \varepsilon_t)$, where $\vartheta = [\beta, \eta, \epsilon, \bar{n}, \bar{s}, \bar{g}, \bar{\pi}, \varphi, \phi_\pi, \phi_y, h, \rho_a, \rho_i, \sigma_g, \sigma_a, \sigma_i]'$ is a vector of parameters, T and M are the transition and impact matrices from the linear solution, Ψ is a vector-valued function of the nonlinear equilibrium system of equations, and $\varepsilon_t = [\varepsilon_{g,t}, \varepsilon_{a,t}, \varepsilon_{i,t}]'$ is a vector of shocks.

[Table 2](#) displays information about the priors. 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 are relatively diffuse to give the algorithm the flexibility to search the parameter space and consistent with the values used in the literature.

We are free to set the ME variances to zero when we use the Kalman filter, since the number of

Parameter	Dist.	Mean (SD)	Parameter	Dist.	Mean (SD)	Parameter	Dist.	Mean (SD)
φ	Norm	100.00 (25.000)	h	Beta	0.5000 (0.2000)	σ_g	IGam	0.0150 (0.0150)
ϕ_π	Norm	2.0000 (0.2500)	ρ_a	Beta	0.8000 (0.2000)	σ_a	IGam	0.0150 (0.0150)
ϕ_y	Norm	0.5000 (0.2500)	ρ_i	Beta	0.8000 (0.2000)	σ_i	IGam	0.0025 (0.0025)

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

observables is equal to the number of shocks. The particle filter, however, always requires positive ME variances to avoid degeneracy. Unfortunately, there is very little consensus in the literature on how to set these values, despite their potentially large effect on the posterior estimates. For example, Ireland (2004) allows for cross- and auto-correlated MEs. However, he finds the 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 their variance to 10% or 20% of the standard deviation of the data improves the empirical fit and forecasting properties of a medium-scale New Keynesian models. Fernández-Villaverde and Rubio-Ramírez (2007) estimate the ME variances instead of fixing them, 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 decided to use a diagonal ME covariance matrix and fix the variances, consistent with most other papers that use a particle filter (e.g., An and Schorfheide (2007); Bocola (2016); Doh (2011); Gust et al. (2017); Plante et al. (2018); van Binsbergen et al. (2012)). We consider three values for the ME variance of real GDP growth, inflation, and the nominal interest rate: 5%, 10%, and 20% of their variance in the data. These percentages capture the wide range of values in the literature. As a benchmark, we also use 0% ME variances when we estimate Lin-KF.

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 100,000 draws from the posterior distribution and then thin by 100 to limit the effects of serial correlation. Therefore, each posterior distribution contains a sample of 1,000 draws.

The algorithm is programmed in Fortran using Open MPI and executed on a cluster. We run the chains in parallel across several supercomputers. When we estimate Lin-KF, each chain uses a single core. To estimate NL-UKF and NL-PF, each chain uses 20 cores because we parallelize the non-linear solution across the nodes in the state space. For example, if a cluster has 100 cores available, we can simultaneously run 100 Lin-KF chains but only 5 NL-PF chains. To increase the accuracy

of the particle filter, we evaluate the likelihood function on each core. Since each NL-PF chain uses 20 cores, we obtain 20 likelihoods and then determine whether to accept or reject a candidate draw based on the median likelihood. This step reduces the variance of the likelihoods from seed effects at no additional computational cost. The filter uses 40,000 particles for our main results, but we also examine the accuracy of NL-PF when the filter uses 20,000 and 80,000 particles in [Appendix A](#).

	Lin-KF (1 core)	NL-UKF (20 cores)	NL-PF (20 cores)
Seconds per Draw	0.0020 (0.0019, 0.0023)	3.1767 (2.3042, 5.3735)	8.6813 (6.4974, 13.9510)
Hours per Dataset	0.0737 (0.0670, 0.0836)	114.7 (83.2, 194.0)	313.5 (234.6, 503.8)

Table 3: Medians and (5%, 95%) credible sets of the estimation times.

[Table 3](#) reports the medians and credible sets of the computing times for each estimation method across 128 datasets. We approximate the times by calculating the average seconds per draw across the 1,000 draws from the posterior distribution. We also provide hours per dataset, which we extrapolated by multiplying seconds per draw by 130,000 draws and dividing by 3,600 seconds per hour. We report the times for Lin-KF, NL-UKF, and NL-PF on 1, 20, and 20 cores, respectively. For each method, we set the ME variances to 5%. The estimation times depend on the hardware, but there are a couple interesting points. One, with 20 cores, NL-UKF used only about 3.2 seconds per draw, which corresponds to 115 hours or less than 5 days per dataset. Therefore, nonlinear estimation is possible in a reasonable amount of time on a single workstation. Two, even NL-PF finished in less than 2 weeks, which made it feasible for us to estimate a large number of datasets.

4 POSTERIOR ESTIMATES AND ACCURACY

4.1 ACCURACY MEASURES We measure the accuracy of each estimation method by calculating the root-mean square-error in the estimated parameters, $RMSE^\theta$, and filtered state variables, $RMSE^s$. In the former case, the error is the difference between a draw from the posterior distribution, $\hat{\theta}_d$, and the true parameter, $\tilde{\theta}$. In the latter case, the error is the difference between the median filtered state variables based on a draw from the posterior distribution, \hat{s}_d , and the true state variables, \tilde{s} . For parameter i , state variable j , dataset k , and method h , the statistics are given by

$$RMSE_{i,k,h}^\theta = \frac{1}{\tilde{\theta}_i} \sqrt{\frac{1}{N_d} \sum_{d=1}^{N_d} (\hat{\theta}_{d,i,k,h} - \tilde{\theta}_i)^2}, \quad (14)$$

$$RMSE_{j,k,h}^s = \sum_{t=1}^T \sqrt{\frac{1}{N_d} \sum_{d=1}^{N_d} (\hat{s}_{d,j,k,h,t} - \tilde{s}_{j,t})^2}, \quad (15)$$

where $N_d = 1000$ is the number of posterior draws after thinning and $T = 100$ is the number of quarters in each data sample. $\hat{\theta}_{d,i,k,h} \sim p(\theta_i | k, h)$ is the d th draw of parameter i from its posterior

distribution, conditional on dataset k and estimation method h . $\hat{s}_{d,j,k,h,t}$ is j th median filtered state variable based on parameter draw d , conditional on dataset k , estimation method h , and quarter t .

There are two primary benefits of the $RMSE$ statistics. One, they assess the accuracy of the entire distribution, rather than just the median draw. Specifically, the $RMSE$ statistics naturally weight the errors according to their posterior distribution, so the errors near the center of the distribution have a larger weight than the errors in the tails. Two, we are able to obtain an overall measure of accuracy across the estimated parameters because $RMSE^\theta$ is normalized by the true parameter, $\tilde{\theta}_i$, to remove scale differences. Similarly, we report the accuracy of the state variables by summing the errors across all periods in our sample. [Appendix A](#) shows how well each method is able to recover the true means, standard deviations, and covariances in each dataset. For each statistic, we report quantiles across the datasets to identify differences in the accuracy of each estimation method.

4.2 PARAMETER ESTIMATES AND MARGINAL DATA DENSITIES We begin by comparing the posterior parameter estimates to the true values in [table 4](#). There are three estimation methods: the linear model estimated with a Kalman filter and 0% ME variances (Lin-KF-0%), the linear model estimated with a Kalman filter and 5% ME variances (Lin-KF-5%), and the nonlinear model estimated with a particle filter and 5% ME variances (NL-PF-5%). We report the medians and 90% credible sets for each parameter and method after pooling the draws from all datasets and the subsample of datasets where the ZLB binds for at least 15 quarters. Differentiating the results based on the frequency of ZLB events shows how the amount of nonlinearity in the data affects the estimates.

In the unconditional sample, it appears there is little benefit to using NL-PF-5%, even though the ZLB binds on average for 6 quarters. For all but two parameters, the median NL-PF-5% estimates are further from the truth than the Lin-KF-0% estimates and the values are similar to the Lin-KF-5% estimates. The benefits of NL-PF-5% are clearer when the data spends at least 15 quarters at the ZLB. While most of the NL-PF-5% estimates are just as accurate as they were in the unconditional sample, the Lin-KF-0% and Lin-KF-5% estimates are typically farther from the truth. NL-PF-5% has the biggest advantage in estimating the persistence and standard deviation of the preference shock (ρ_a and σ_a) because those parameters affect the frequency and duration of ZLB events. However, the median NL-PF-5% estimates are less accurate than the linear estimates along some dimensions, in spite of the nonlinearity in the data. For example, the true price adjustment cost parameter is 100, but it increases from 105.8 in the unconditional sample to 111.5 in the ZLB sample.

The second to last row in [table 4](#) shows the median and 90% credible sets of the marginal data densities, ℓ , across our datasets. The density is based on Geweke's (1999) harmonic mean estimator, which indicates how well each method fits the data. We adjust the values for the number of mode search draws where the nonlinear solution method did not converge, so the prior density integrates to one. In the unconditional sample, the differences between the median data densities are

Ptr	Truth	Unconditional (640 datasets)			ZLB quarters ≥ 15 (53 datasets)		
		Lin-KF-0%	Lin-KF-5%	NL-PF-5%	Lin-KF-0%	Lin-KF-5%	NL-PF-5%
φ	100	98.5 (63.7, 136.7)	105.9 (70.6, 144.0)	105.8 (70.7, 143.8)	103.3 (69.7, 139.7)	110.8 (76.8, 147.4)	111.5 (78.4, 148.0)
ϕ_π	2.0	1.982 (1.584, 2.385)	2.006 (1.610, 2.408)	2.034 (1.647, 2.432)	1.981 (1.592, 2.380)	2.009 (1.621, 2.407)	2.081 (1.699, 2.473)
ϕ_y	0.5	0.497 (0.227, 0.805)	0.496 (0.219, 0.806)	0.506 (0.230, 0.815)	0.473 (0.215, 0.773)	0.467 (0.202, 0.774)	0.487 (0.217, 0.786)
h	0.5	0.483 (0.322, 0.625)	0.481 (0.317, 0.625)	0.492 (0.335, 0.630)	0.474 (0.326, 0.604)	0.471 (0.321, 0.602)	0.498 (0.357, 0.624)
ρ_a	0.8	0.808 (0.649, 0.910)	0.831 (0.677, 0.926)	0.811 (0.668, 0.883)	0.857 (0.752, 0.931)	0.877 (0.778, 0.944)	0.838 (0.748, 0.894)
ρ_i	0.8	0.801 (0.730, 0.852)	0.821 (0.752, 0.872)	0.816 (0.748, 0.866)	0.818 (0.755, 0.862)	0.840 (0.778, 0.885)	0.819 (0.755, 0.867)
σ_g	0.015	0.0146 (0.0112, 0.0190)	0.0136 (0.0104, 0.0178)	0.0137 (0.0105, 0.0180)	0.0145 (0.0115, 0.0186)	0.0134 (0.0106, 0.0174)	0.0139 (0.0109, 0.0180)
σ_a	0.015	0.0160 (0.0116, 0.0223)	0.0158 (0.0114, 0.0224)	0.0145 (0.0108, 0.0191)	0.0186 (0.0139, 0.0262)	0.0185 (0.0138, 0.0265)	0.0157 (0.0120, 0.0202)
σ_i	0.0025	0.0024 (0.0020, 0.0030)	0.0021 (0.0016, 0.0027)	0.0022 (0.0017, 0.0028)	0.0022 (0.0018, 0.0028)	0.0019 (0.0014, 0.0024)	0.0022 (0.0016, 0.0028)
ℓ		1280.4 (1260.5, 1301.0)	1280.3 (1260.3, 1300.4)	1282.2 (1263.1, 1303.7)	1278.9 (1257.3, 1296.1)	1278.5 (1257.8, 1296.9)	1290.8 (1266.1, 1307.9)
$\Delta\ell$		—	-0.085 (-3.360, 2.213)	1.882 (-3.709, 11.111)	—	-0.365 (-3.532, 1.872)	10.857* (3.825, 18.631)

Table 4: Posterior medians and (5%, 95%) credible sets of the estimated parameters and marginal data densities. An asterisk indicates the differences in the marginal data densities (values different from 0) are significant at a 10% level.

small across the three methods. When we condition on datasets where the ZLB binds for at least 15 quarters, the nonlinear model fits the data better, while the linear models fit the data slightly worse.

To get an indication of whether the differences in empirical fit are statistically significant, we compute the distribution of the differences between an alternative method and Lin-KF-0% with each dataset. The difference between any two median density values, $\Delta\ell$, is known as the Bayes factor (in logs). A value greater than 4.6 ($\log 10^2$) is viewed as decisive evidence in favor of the alternative method. The median and 90% credible sets of the differences are shown in the last row. Positive values indicate an alternative method fits the data better than Lin-KF-0%. If the credible set does not include zero, it indicates that the differences in the marginal data densities are statistically significant at the 10% level. The results confirm our previous findings. NL-PF-5% does not provide a statistically significant advantage across all datasets. However, when restricting attention to datasets with at least 15 ZLB quarters, the differences are significant and decisive. These results demonstrate that the data must contain a reasonable amount of nonlinearity for the nonlinear model to have a statistically significant advantage over the linear model in fitting the data.

4.3 PARAMETER ACCURACY Table 5 shows how well each estimation method is able to recover the true parameters. For each parameter, dataset, and method, we first calculate $RMSE^\theta$ as well as the sum of the errors across the parameters (bottom row, denoted by Σ). We then calculate the ratio of the errors with each alternative method relative to the error with Lin-KF-0% and report

the median and 90% credible sets of the ratios across the datasets. A value less (greater) than one indicates an alternative method is more (less) accurate than Lin-KF-0%. If the credible set does not include one, it indicates that the differences in accuracy are statistically significant at the 10% level.

Ptr	Unconditional (640 datasets)		ZLB quarters ≥ 15 (53 datasets)		ZLB quarters ≥ 25 (8 datasets)	
	Lin-KF-5%	NL-PF-5%	Lin-KF-5%	NL-PF-5%	Lin-KF-5%	NL-PF-5%
φ	1.058 (0.834, 1.246)	1.042 (0.830, 1.261)	1.137 (0.861, 1.257)	1.111 (0.825, 1.342)	1.134* (1.022, 1.357)	1.176* (1.016, 1.387)
ϕ_π	0.993 (0.923, 1.069)	0.991 (0.856, 1.127)	0.999 (0.926, 1.080)	1.032 (0.774, 1.322)	1.014 (0.972, 1.060)	1.107 (0.857, 1.379)
ϕ_y	1.014 (0.947, 1.096)	1.013 (0.882, 1.166)	1.018 (0.952, 1.126)	1.021 (0.763, 1.252)	1.022 (0.969, 1.127)	0.965 (0.731, 1.237)
h	1.021 (0.941, 1.108)	0.971 (0.817, 1.179)	1.027 (0.935, 1.127)	0.953 (0.669, 1.240)	1.003 (0.941, 1.158)	0.893 (0.664, 1.263)
ρ_a	1.046 (0.777, 1.300)	0.827 (0.626, 1.075)	1.158* (1.030, 1.355)	0.774* (0.504, 0.996)	1.135* (1.058, 1.291)	0.659* (0.502, 0.884)
ρ_i	1.162 (0.708, 1.612)	1.061 (0.709, 1.471)	1.353 (0.910, 1.945)	1.026 (0.745, 1.490)	1.339* (1.173, 1.566)	0.922 (0.633, 1.155)
σ_g	1.147 (0.748, 1.403)	1.104 (0.712, 1.394)	1.203 (0.823, 1.382)	1.097 (0.858, 1.347)	1.120 (0.891, 1.373)	1.048 (0.955, 1.332)
σ_a	1.010 (0.845, 1.268)	0.771 (0.399, 1.160)	0.999 (0.865, 1.267)	0.508* (0.271, 0.681)	0.953 (0.856, 1.196)	0.399* (0.260, 0.598)
σ_i	1.645 (0.696, 2.226)	1.392 (0.726, 1.907)	1.797* (1.022, 2.721)	1.223 (0.824, 1.854)	1.537* (1.325, 2.103)	0.925 (0.639, 1.223)
Σ	1.069 (0.958, 1.189)	0.987 (0.864, 1.101)	1.123* (1.042, 1.205)	0.923 (0.734, 1.071)	1.102* (1.042, 1.167)	0.825* (0.720, 0.954)

Table 5: Median and (5%, 95%) credible sets of the root-mean square-error in the estimated parameters relative to Lin-KF-0%. An asterisk indicates the differences in the errors (values different from 1) are significant at a 10% level.

Across all datasets, the nonlinear model has a slight advantage over Lin-KF-0% along a few dimensions, including the sum of the errors across all parameters. However, none of the differences are statistically significant, even though the ZLB binds on average for 6 quarters. When we condition on datasets where the ZLB binds for at least 15 quarters, the results change. Positive ME variances in the Kalman filter lead to even less accurate estimates than in the unconditional sample, including two parameters (σ_i and ρ_a) that are significantly different from Lin-KF-0%. In contrast, the nonlinear model becomes relatively more accurate along most dimensions. The differences are statistically significant for the two parameters (ρ_a and σ_a) with the largest impact on whether the ZLB binds each quarter. Interestingly, the sum of the errors is not statistically different from Lin-KF-0%.

When turning to the small subset of datasets where the ZLB binds for at least 25 quarters, the results are even more stark. The total error as well as four of the nine parameters with Lin-KF-5% are statistically less accurate than Lin-KF-0%. The opposite is true of the nonlinear model. The total error, as well as the *RMSE* in two of the nine parameters, is statistically less than Lin-KF-0%. Two key takeaways emerge from these results: (1) Positive ME variances in the observation equation of the filter hurt the accuracy of the estimation; (2) the data must be highly nonlinear for the particle filter to overcome positive ME variances and become statistically more accurate than Lin-KF-0%.

$h \backslash \text{ME}$	5%	10%	20%	5%	10%	20%
	Unconditional (128 datasets)			ZLB quarters ≥ 5 (64 datasets)		
Lin-KF	1.070 (0.965, 1.189)	1.161* (1.023, 1.319)	1.304* (1.133, 1.490)	1.099 (0.996, 1.191)	1.195* (1.049, 1.344)	1.318* (1.187, 1.544)
NL-UKF	0.995 (0.870, 1.132)	1.077 (0.937, 1.252)	1.227* (1.051, 1.450)	0.980 (0.867, 1.148)	1.062 (0.937, 1.272)	1.203* (1.038, 1.480)
NL-PF	0.994 (0.857, 1.091)	1.071 (0.942, 1.233)	1.208* (1.056, 1.432)	0.971 (0.842, 1.097)	1.045 (0.946, 1.241)	1.185* (1.038, 1.480)
	ZLB quarters ≥ 10 (31 datasets)			ZLB quarters ≥ 15 (13 datasets)		
Lin-KF	1.090 (0.992, 1.187)	1.180* (1.077, 1.342)	1.318* (1.125, 1.428)	1.090* (1.042, 1.194)	1.199* (1.083, 1.338)	1.335* (1.132, 1.405)
NL-UKF	0.948 (0.800, 1.144)	1.033 (0.904, 1.235)	1.167* (1.035, 1.408)	0.939 (0.796, 1.174)	1.026 (0.862, 1.203)	1.132* (1.036, 1.285)
NL-PF	0.941 (0.785, 1.063)	1.013 (0.882, 1.198)	1.145 (0.970, 1.358)	0.905 (0.735, 1.031)	0.974 (0.792, 1.126)	1.106 (0.921, 1.265)

(a) Median and (5%, 95%) credible sets of the root-mean square-error in the estimated parameters relative to Lin-KF-0%.

$h \backslash \text{ME}$	5%	10%	20%	5%	10%	20%
	Unconditional (128 datasets)			ZLB quarters ≥ 5 (64 datasets)		
Lin-KF	0.082 (-4.97, 3.83)	-2.486 (-11.20, 4.50)	-10.376 (-26.11, 3.32)	-0.099 (-5.04, 2.37)	-2.840 (-11.24, 3.39)	-11.756 (-26.25, 3.62)
NL-UKF	-0.098 (-9.14, 14.33)	-1.998 (-15.05, 7.01)	-10.077 (-27.08, 3.41)	1.829 (-9.43, 14.54)	-1.801 (-15.52, 7.08)	-10.903 (-27.39, 3.58)
NL-PF	1.763 (-6.58, 22.52)	-1.077 (-12.89, 11.98)	-9.404 (-25.75, 4.95)	4.930 (-6.31, 23.16)	-0.260 (-13.25, 12.52)	-9.920 (-27.39, 3.58)
	ZLB quarters ≥ 10 (31 datasets)			ZLB quarters ≥ 15 (13 datasets)		
Lin-PF	-0.304 (-4.10, 2.37)	-3.096 (-11.00, 3.39)	-11.852 (-26.25, 1.49)	-0.283 (-2.92, 1.71)	-3.011 (-9.03, 1.04)	-11.763* (-23.57, -3.39)
NL-UKF	3.084 (-3.59, 14.54)	-1.301 (-9.32, 7.08)	-10.954 (-24.11, 2.36)	4.427 (-0.64, 14.54)	-0.928 (-5.41, 7.08)	-13.669* (-21.68, -2.04)
NL-PF	7.855 (-0.46, 23.16)	1.502 (-8.43, 12.52)	-9.810 (-23.62, 2.83)	9.638* (6.59, 23.16)	1.855 (-2.25, 12.52)	-10.257* (-20.44, -0.56)

(b) Median and (5%, 95%) credible sets of the marginal data densities relative to Lin-KF-0%.

Table 6: Comparison of the root-mean square-error (top panel) and marginal data densities (bottom panel) with an expanded set of specifications. An asterisk indicates the differences from Lin-KF-0% are significant at a 10% level.

4.4 ADDITIONAL SPECIFICATIONS Table 6 examines a much broader set of specifications. The table shows results for the nonlinear model estimated with the unscented Kalman filter (NL-UKF) and for each specification with 10% and 20% ME variances. We also condition on datasets with at least 5 quarters and at least 10 quarters at the ZLB to provide more variation in the degree of nonlinearity in the data. To accommodate the additional specifications, we lowered the number of datasets from 640 to 128, though most of the previous results are robust to the smaller sample size.

The top panel shows the parameter accuracy for each method relative to Lin-KF-0%. The values are based on the total error across the parameters so they are analogous to the values in the bottom row of table 5. There are several key findings. One, the ME variances have a large impact on accuracy. Regardless of the estimation method and dataset, lower ME variances lead to more ac-

curate estimates. As the ME variances increase, the filter is more likely to treat extreme realizations in the data as ME, even though outliers are the most informative about the true parameter values. Setting the ME variances to 0% treats all tail realizations as informative variations in the data.

Two, for a given value of the ME variances, NL-PF is more accurate than the other methods and the differences in accuracy become larger as the fraction of time spent at the ZLB increases.

Three, NL-UKF is consistently more accurate than Lin-KF and the accuracy increases when the data is more nonlinear, similar to our results with NL-PF. NL-UKF is always less accurate than NL-PF, but takes less than half as much time to estimate NL-PF and it is nearly as accurate. Therefore, it may be reasonable to prefer NL-UKF for some models, even though it is less accurate.

Four, if the ME variances are too large, Lin-KF-0% is more accurate than both NL-UKF and NL-PF, even if the data is highly nonlinear. For example, when the ZLB binds for at least 15 quarters and the ME variances are set to 5%, NL-PF is considerably more accurate than Lin-KF-0%, whereas the opposite result occurs with 20% ME variances. When using all of the datasets, NL-PF-20% is statistically less accurate than Lin-KF-0%. These results are important, since it is common in the literature to specify large ME variances to increase the accuracy of the particle filter.

The values in the bottom panel are the marginal data densities for the same specifications in the top panel, relative to Lin-KF-0%. These values are analogous to those in the bottom row of [table 4](#). There are several similarities with the accuracy results. One, the data density always increases as the ME variances decline. Two, NL-PF typically has the highest data density and the advantage over the Lin-KF-0% density increases with the number of ZLB quarters in the data. The one counterexample is when the ME variances are equal to 20%. Given a sufficiently high ME variance, any improvement in the data density from using a nonlinear model instead of a linear model is completely washed away, even when the data spends many quarters at the ZLB. This result demonstrates that the presence of ME variance significantly biases the estimates. Any research that compares the estimates from linear and nonlinear models could erroneously conclude there is either no or a very small benefit to using the nonlinear model simply because the ME variances are too large.

4.5 ACCURACY OF THE LATENT STATES [Table 7](#) shows how well each method is able to recover the latent preference shock state across time. For each dataset, we calculate $RMSE^s$, which sums the errors across time based on the median filtered states. Just like in previous tables, we calculate the ratio of the errors with each alternative method relative to the error with Lin-KF-0% and report the median and 90% credible sets of the ratios across the datasets. We focus on the preference shock state because it has the largest impact on whether the ZLB binds in each period. It is also the only state variable that is persistent across time and not linked to one of the observables.

Similar to the results for the parameters, the accuracy of the filtered states improves as the ME variances decrease, regardless of the estimation method. As long as the ME variances are suffi-

$h \backslash \text{ME}$	5%	10%	20%	5%	10%	20%
	Unconditional (128 datasets)			ZLB quarters ≥ 5 (64 datasets)		
Lin-KF	1.139 (0.99, 1.64)	1.268* (1.04, 1.84)	1.402* (1.15, 2.25)	1.190* (1.02, 1.59)	1.310* (1.10, 1.96)	1.408* (1.16, 2.29)
NL-UKF	0.720* (0.53, 0.96)	0.840 (0.62, 1.10)	1.016 (0.71, 1.47)	0.702* (0.46, 0.91)	0.828 (0.58, 1.04)	0.963 (0.66, 1.37)
NL-PF	0.705* (0.52, 0.88)	0.826 (0.59, 1.09)	1.008 (0.68, 1.39)	0.674* (0.44, 0.87)	0.789 (0.55, 1.04)	0.928 (0.66, 1.37)
	ZLB quarters ≥ 10 (31 datasets)			ZLB quarters ≥ 15 (13 datasets)		
Lin-KF	1.261* (1.04, 1.63)	1.411* (1.12, 2.02)	1.445* (1.19, 2.31)	1.296* (1.05, 1.68)	1.536* (1.06, 2.07)	1.585* (1.08, 2.30)
NL-UKF	0.692* (0.43, 0.88)	0.789 (0.46, 1.02)	0.895 (0.63, 1.32)	0.685* (0.34, 0.96)	0.689 (0.41, 1.03)	0.878 (0.52, 1.28)
NL-PF	0.615* (0.34, 0.79)	0.714* (0.41, 0.95)	0.883 (0.55, 1.29)	0.568* (0.31, 0.78)	0.681* (0.37, 0.98)	0.833 (0.43, 1.24)

Table 7: Median and (5%, 95%) credible sets of the root-mean square-error in the preference shock state (a) relative to Lin-KF-0%. An asterisk indicates the differences in the errors (values different from 1) are significant at a 10% level.

ciently low (e.g., 5% of the variance in the data), both NL-UKF and NL-PF provide statistically more accurate estimates of the preference shock state across all datasets. When we condition on datasets that spend a larger number of quarters at the ZLB, the differences in accuracy between the linear and nonlinear estimation methods become even larger, particularly with NL-PF. The nonlinear model provides significantly more accurate estimates than the linear model for two reasons. One, the nonlinear model accounts for the expectation of entering, staying at, and exiting the ZLB. Two, a lower preference shock state generates simultaneous declines in inflation and real GDP growth that make the ZLB more likely to bind and for a longer duration, which is exacerbated by the fact that it is highly persistent. Therefore, historical estimates of the state variables are the most accurate with a method that sets small ME variances and uses both a nonlinear solution and filter.

4.6 IMPULSE RESPONSES To show the economic implications of the parameter bias and solution method, we compare generalized impulse response functions (GIRFs) to the true GIRFs across our estimation methods. The advantage of using GIRFs over traditional impulse response functions is that we can more accurately compare the effects of shocks in different states of the economy. To compute GIRFs, we follow the procedure in Koop et al. (1996). Given an initial state, we first calculate the mean of 10,000 model simulations using 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 (e.g., a preference shock) is replaced with the desired shock size. The GIRF is defined as the difference between the two mean paths. See Appendix D for further details.

Figure 2 shows the responses of real GDP growth to a negative 2 standard deviation preference shock. A preference shock is a proxy for a change in demand because it determines households' degree of patience. Negative shocks cause households to postpone consumption to future periods,

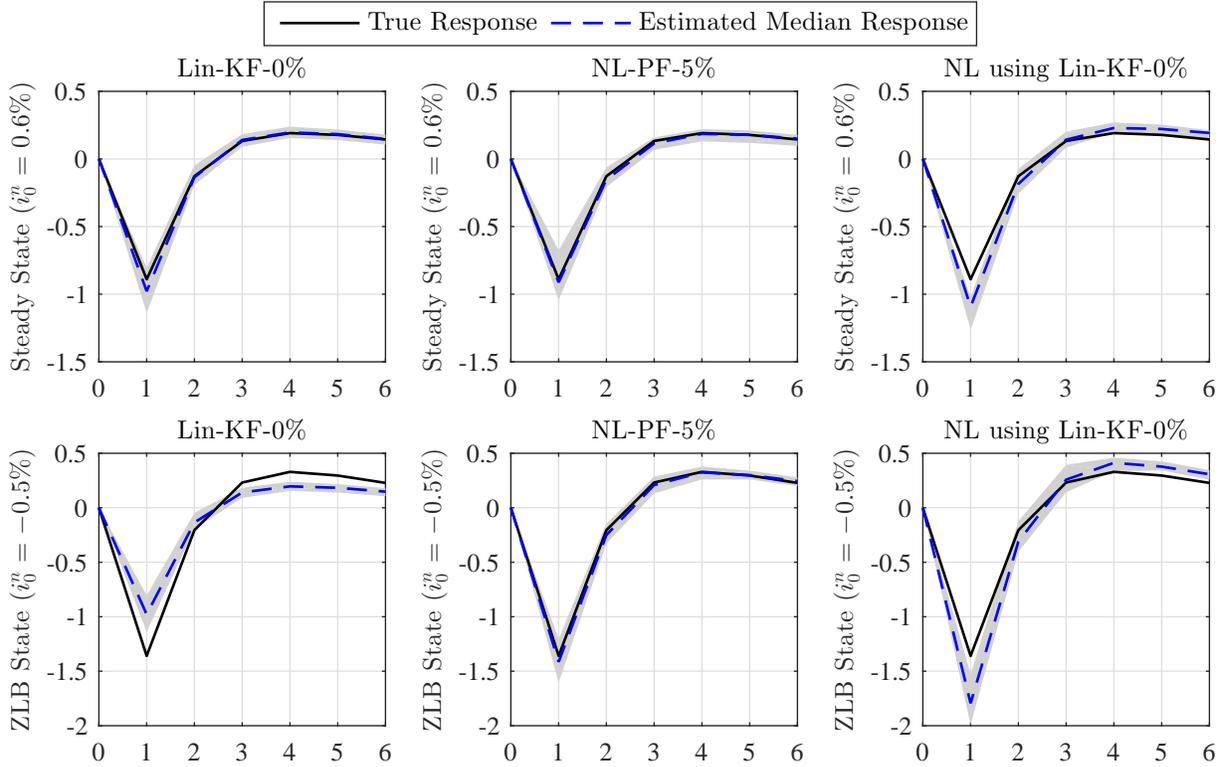


Figure 2: Generalized impulse responses of real GDP growth (in percentage point differences) to a -2 standard deviation preference shock. Shaded regions are the (5%, 95%) credible sets of the estimated responses across the 53 datasets that spend at least 15 quarters at the ZLB. The GIRFs are conditional on the posterior means from Lin-KF-0% (left column), NL-PF-5% (middle column), and Lin-PF-0% solved with the nonlinear solution method (right column).

which reduces current real GDP growth. We focus on this shock, rather than the technology growth or monetary policy shocks, because it is the primary mechanism for generating ZLB events in the model. GIRFs are calculated with the true nonlinear model and the estimated linear and nonlinear models, which are parameterized at the posterior mean estimates for each of the 53 datasets that spend at least 15 quarters at the ZLB. We use that subset since those datasets look more like recent U.S. data and better highlight the potential shortfalls of linear estimation methods. The Lin-KF-0% estimates are shown in the left column and the NL-PF-5% estimates are provided in the middle column. We also show what happens if estimates are obtained with Lin-KF-0% and then the responses are computed with the nonlinear solution method (right column), as some other papers have done.

The true response (solid line) is compared to the estimated median response (dashed line). The 90% credible sets (shaded regions) show variation in the estimated GIRFs across the datasets. In the top row, the GIRFs are initialized at the true stochastic steady state and in the bottom row they are initialized in a ZLB state. The ZLB state is the average state in quarters when the notional rate is between -0.6% and -0.4% in a long simulation of the true model. In the ZLB state, the average notional rate is -0.5% , which equals the 2008Q4 estimate for the U.S. from Plante et al. (2018).

In steady state (top panel), while not significant, the median Lin-KF-0% response on impact is

bigger than the truth since the estimates must compensate for the absence of the ZLB in the solution. However, the median NL-PF-5% response closely follows the truth. If researchers instead computed GIRFs with the nonlinear solution using estimates from Lin-KF-0%, they would significantly overpredict the effect of the shock on impact and be worse off than using the linear solution.

The effect of the shock is independent of the state of the economy in a linear model, so Lin-KF-0% predicts the same 1% decline in real GDP growth in the ZLB state (bottom panel) as it does in steady state. However, the true response of real GDP growth is much larger (-1.4%) because the ZLB amplifies the effect of the shock in the data. When the nominal interest rate is far from the ZLB, the drop in real GDP growth is damped by the monetary policy response. When the ZLB binds, the central bank cannot respond by lowering its policy rate, which leads to the larger decline in real GDP growth. Therefore, the true response is far outside the Lin-KF-0% credible set. In contrast, the median NL-PF-5% response coincides with the true response of real GDP growth in the ZLB state. If estimates are obtained with Lin-KF-0% and the GIRFs are computed with the nonlinear solution, the median impact (-1.8%) is significantly different from the truth. Researchers are equally worse off using the linear and nonlinear solution with the Lin-KF-0% parameter estimates.

The key takeaway from these simulations is that the errors in estimating the parameters and latent states, as shown in previous tables, will generate significant errors in the model's predictions.

5 CONCLUSION

This paper examines the accuracy of linear and nonlinear estimation methods for DSGE models using artificial datasets generated from a nonlinear New Keynesian model with an occasionally binding ZLB constraint. We find the accuracy of each method crucially depends on the size of the ME variances in the observation equation and the number of quarters the ZLB binds in the data. As long as the ME variances are not too large (e.g., less than 5% of the variance in the data) and there is a sufficient presence of the ZLB in the data (more than 15% of quarters in the sample), the nonlinear model estimated with a particle filter is statistically more accurate than the linear model estimated with a Kalman filter. These findings are particularly important in light of the recent ZLB experience in the U.S., Japan, and the Euro Area and the increased likelihood of future ZLB events.

Over the last decade, researchers have used a variety of posterior samplers, solution methods, and filters for policy analysis. Our framework provides a unified way to assess the relative performance of different methods applied to our datasets. For example, a promising alternative to the random walk Metropolis-Hastings (RWMH) algorithm is the sequential Monte Carlo (SMC) method first applied to a DSGE model by Creal (2007). Herbst and Schorfheide (2014) find SMC is better suited for multimodal and irregular posterior distributions than RWMH, and parallelization helps reduce the numerical cost of SMC even though it performs significantly more evaluations of the likelihood function than a typical RWMH algorithm. A popular alternative to our global solution

method is to use the Smolyak method to discretize the state space, low-order Chebyshev polynomials to approximate the policy functions, and fixed-point iteration (e.g., Fernández-Villaverde et al. (2015); Gust et al. (2017); Judd et al. (2014)). Our results can also serve as a benchmark for continued improvements to the particle filter, such as the tempered particle filter proposed by Herbst and Schorfheide (2017). Finally, while the UKF is a significant improvement over the Kalman filter, the central difference Kalman filter has the potential to further improve accuracy (Andreasen (2013)).

Aside from studying our model with an occasionally binding ZLB constraint, researchers could compare estimation methods applied to artificial datasets generated by models with other potentially important nonlinearities, such as asymmetric adjustment costs, firm default, borrowing constraints, search frictions, and stochastic volatility. Another line of research could examine the role of model misspecification. For example, one could generate datasets with a global solution to a medium-scale model and then ask how well a small-scale model can recover key parameters. We believe this line of work will uncover key nonlinear features and deliver methodological advances.

REFERENCES

- AN, S. AND F. SCHORFHEIDE (2007): “Bayesian Analysis of DSGE Models,” *Econometric Reviews*, 26, 113–172.
- ANDREASEN, M. M. (2013): “Non-Linear DSGE Models And The Central Difference Kalman Filter,” *Journal of Applied Econometrics*, 28, 929–955.
- 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.
- ARUOBA, S. B., J. FERNANDEZ-VILLAVARDE, AND J. F. RUBIO-RAMIREZ (2006): “Comparing Solution Methods for Dynamic Equilibrium Economies,” *Journal of Economic Dynamics and Control*, 30, 2477–2508.
- BOCOLA, L. (2016): “The Pass-Through of Sovereign Risk,” *Journal of Political Economy*, 124, 879–926.
- CANOVA, F., F. FERRONI, AND C. MATTHES (2014): “Choosing The Variables To Estimate Singular Dsge Models,” *Journal of Applied Econometrics*, 29, 1099–1117.
- CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, 113, 1–45.
- CHRISTIANO, L. J., M. EICHENBAUM, AND M. TRABANDT (2015): “Understanding the Great Recession,” *American Economic Journal: Macroeconomics*, 7, 110–67.
- COLEMAN, II, W. J. (1991): “Equilibrium in a Production Economy with an Income Tax,” *Econometrica*, 59, 1091–1104.
- CREAL, D. (2007): “Sequential Monte Carlo samplers for Bayesian DSGE models,” Manuscript, University of Chicago, Booth.

- CUBA-BORDA, P. (2014): “What Explains the Great Recession and the Slow Recovery?” Manuscript, University of Maryland.
- CUBA-BORDA, P., L. GUERRIERI, M. IACOVIELLO, AND M. ZHONG (2017): “Likelihood Evaluation of Models with Occasionally Binding Constraints,” Manuscript, Federal Reserve Board.
- DEL NEGRO, M., M. P. GIANNONI, AND F. SCHORFHEIDE (2015): “Inflation in the Great Recession and New Keynesian Models,” *American Economic Journal: Macroeconomics*, 7, 168–196.
- DOH, T. (2011): “Yield Curve in an Estimated Nonlinear Macro Model,” *Journal of Economic Dynamics and Control*, 35, 1229 – 1244.
- 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.
- FERNÁNDEZ-VILLAVERDE, J. AND J. F. RUBIO-RAMÍREZ (2005): “Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood,” *Journal of Applied Econometrics*, 20, 891–910.
- (2007): “Estimating Macroeconomic Models: A Likelihood Approach,” *Review of Economic Studies*, 74, 1059–1087.
- FERNÁNDEZ-VILLAVERDE, J., J. F. RUBIO-RAMÍREZ, AND F. SCHORFHEIDE (2016): “Solution and Estimation Methods for DSGE Models,” in *Handbook of Macroeconomics, Vol. 2*, ed. by J. B. Taylor and H. Uhlig, North-Holland, Handbook of Macroeconomics, 527–724.
- GALÍ, J., F. SMETS, AND R. WOUTERS (2012): “Unemployment in an Estimated New Keynesian Model,” in *NBER Macroeconomics Annual 2012, Vol. 26*, ed. by D. Acemoglu and M. Woodford, University of Chicago Press, 329–360.
- 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.
- GEWEKE, J. (1999): “Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication,” *Econometric Reviews*, 18, 1–73.
- 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.
- GUERRÓN-QUINTANA, P. A. (2010): “What You Match Does Matter: The Effects of Data on DSGE Estimation,” *Journal of Applied Econometrics*, 25, 774–804.
- 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.
- HERBST, E. AND F. SCHORFHEIDE (2014): “Sequential Monte Carlo Sampling For DSGE Models,” *Journal of Applied Econometrics*, 29, 1073–1098.

- (2017): “Tempered Particle Filtering,” NBER Working Paper 23448.
- 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.
- 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.
- (2011): “A New Keynesian Perspective on the Great Recession,” *Journal of Money, Credit and Banking*, 43, 31–54.
- JUDD, K. L., L. MALIAR, S. MALIAR, AND R. VALERO (2014): “Smolyak Method for Solving Dynamic Economic Models: Lagrange Interpolation, Anisotropic Grid and Adaptive Domain,” *Journal of Economic Dynamics and Control*, 44, 92–123.
- JULIER, S. J. AND J. K. UHLMANN (1997): “A New Extension of the Kalman Filter to Non-linear Systems,” in *Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defense Sensing, Simulations and Controls*, 182–193.
- KEEN, B. D., A. W. RICHTER, AND N. A. THROCKMORTON (2017): “Forward Guidance and the State of the Economy,” *Economic Inquiry*, 55, 1593–1624.
- 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.
- KOOP, G., M. H. PESARAN, AND S. M. POTTER (1996): “Impulse Response Analysis in Non-linear Multivariate Models,” *Journal of Econometrics*, 74, 119–147.
- KOPECKY, K. AND R. SUEN (2010): “Finite State Markov-chain Approximations to Highly Persistent Processes,” *Review of Economic Dynamics*, 13, 701–714.
- 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.
- NAKATA, T. (2017): “Uncertainty at the Zero Lower Bound,” *American Economic Journal: Macroeconomics*, 9, 186–221.
- 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.
- OTROK, C. (2001): “On Measuring the Welfare Cost of Business Cycles,” *Journal of Monetary Economics*, 47, 61–92.
- PETERMAN, W. B. (2016): “Reconciling Micro and Macro Estimates of the Frisch Labor Supply Elasticity,” *Economic Inquiry*, 54, 100–120.
- PLANTE, M., A. W. RICHTER, AND N. A. THROCKMORTON (2018): “The Zero Lower Bound and Endogenous Uncertainty,” *Economic Journal*, forthcoming.
- 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.
- (2016): “Is Rotemberg Pricing Justified by Macro Data?” *Economics Letters*, 149, 44–48.
- RICHTER, A. W., N. A. THROCKMORTON, AND T. B. WALKER (2014): “Accuracy, Speed and Robustness of Policy Function Iteration,” *Computational Economics*, 44, 445–476.
- ROTEMBERG, J. J. (1982): “Sticky Prices in the United States,” *Journal of Political Economy*, 90, 1187–1211.
- 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.
- SIMS, C. A. (2002): “Solving Linear Rational Expectations Models,” *Computational Economics*, 20, 1–20.
- SMETS, F. AND R. WOUTERS (2007): “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach,” *American Economic Review*, 97, 586–606.
- 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.
- SUH, H. AND T. B. WALKER (2016): “Taking Financial Frictions to the Data,” *Journal of Economic Dynamics and Control*, 64, 39–65.
- VAN BINSBERGEN, J. H., J. FERNÁNDEZ-VILLAYERDE, 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.
- WOLMAN, A. L. (2005): “Real Implications of the Zero Bound on Nominal Interest Rates,” *Journal of Money, Credit and Banking*, 37, 273–96.

A ADDITIONAL RESULTS

Table 8 shows the accuracy of several unconditional moments in the data. We obtain 1,000 simulations of 100 quarters for each estimation method, posterior draw, and dataset. We then take the median of each moment across the 1,000 simulations and compute the *RMSE* across the posterior draws. Finally, we normalize the *RMSE* for a given moment by the error with Lin-KF-0% and report the median and (5%, 95%) credible sets across the datasets. Similar to previous tables, a number less than one indicates an alternative method predicts a moment closer to the truth than the one predicted by Lin-KF-0%. Credible sets that include one imply statistically insignificant differences.

Moment	Unconditional (640 datasets)		ZLB quarters ≥ 15 (53 datasets)		ZLB quarters ≥ 25 (8 datasets)	
	Lin-KF-5%	NL-PF-5%	Lin-KF-5%	NL-PF-5%	Lin-KF-5%	NL-PF-5%
mean(\tilde{c}^g)	0.942* (0.90, 0.99)	0.749 (0.57, 1.38)	0.940* (0.90, 0.98)	0.740 (0.56, 1.48)	0.943 (0.90, 1.01)	0.831 (0.53, 1.93)
mean(π)	0.998 (0.99, 1.01)	0.504* (0.42, 0.68)	0.999 (0.99, 1.01)	0.504* (0.42, 0.79)	0.997 (0.99, 1.00)	0.506* (0.46, 0.89)
mean(i)	0.998 (0.99, 1.01)	0.382* (0.31, 0.57)	0.998 (0.99, 1.01)	0.383* (0.32, 0.56)	0.995 (0.99, 1.01)	0.392* (0.35, 0.57)
std(\tilde{c}^g)	1.197 (0.71, 1.49)	1.169 (0.69, 1.50)	0.984 (0.68, 1.52)	1.059 (0.62, 1.60)	1.004 (0.65, 1.44)	1.062 (0.46, 1.45)
std(π)	1.048 (0.71, 1.38)	1.066 (0.56, 1.50)	0.787 (0.65, 1.18)	0.690 (0.44, 1.49)	0.752* (0.65, 0.80)	0.686* (0.39, 0.79)
std(i)	1.008 (0.66, 1.41)	0.883 (0.30, 1.42)	0.763 (0.56, 1.11)	0.394* (0.23, 0.82)	0.702* (0.56, 0.79)	0.330* (0.19, 0.42)
corr($\tilde{c}_t^g, \tilde{c}_{t-1}^g$)	1.040 (0.96, 1.13)	0.992 (0.83, 1.20)	1.044 (0.96, 1.14)	0.953 (0.71, 1.22)	1.013 (0.94, 1.20)	0.968 (0.65, 1.39)
corr(π_t, π_{t-1})	1.003 (0.77, 1.22)	0.846 (0.68, 1.03)	1.127* (1.01, 1.28)	0.833 (0.59, 1.04)	1.131* (1.05, 1.19)	0.749 (0.60, 1.01)
corr(i_t, i_{t-1})	1.094 (0.76, 1.41)	0.834 (0.64, 1.12)	1.231* (1.07, 1.54)	0.714 (0.49, 1.11)	1.154* (1.04, 1.35)	0.603* (0.45, 0.75)
corr(\tilde{c}^g, i)	1.123 (0.88, 1.40)	1.133 (0.87, 1.46)	1.169 (0.98, 1.37)	1.166 (0.89, 1.57)	1.122* (1.02, 1.35)	1.241 (0.87, 1.66)
corr(\tilde{c}^g, π)	1.079 (0.89, 1.29)	1.086 (0.90, 1.37)	1.110 (0.88, 1.32)	1.164 (0.99, 1.54)	1.111 (0.89, 1.27)	1.176 (0.88, 1.66)
corr(i, π)	1.028 (0.92, 1.17)	0.952 (0.65, 1.30)	1.035 (0.92, 1.20)	0.825 (0.52, 1.10)	1.013 (0.93, 1.15)	0.694 (0.55, 1.01)

Table 8: Medians and (5%, 95%) credible sets of the root-mean square-error of the unconditional moments relative to Lin-KF-0%. An asterisk indicates the differences in the errors (values different from 1) are significant at a 10% level.

Unconditionally, NL-PF-5% predicts means that are closer to the truth than the estimates with the linear models, since the nonlinear solution accounts for differences between the deterministic and stochastic steady states. In datasets with many ZLB quarters, NL-PF-5% is also closer to the truth for other moments, such as some of the standard deviations, serial-correlations, and cross-correlations. For example, when conditioning on datasets with at least 25 ZLB quarters, NL-PF-5% is significantly closer to the truth than Lin-KF-0% for $mean(\pi)$, $mean(i)$, $std(\pi)$, $std(i)$, and $corr(i_t, i_{t-1})$ because it accounts for changes in volatility and persistence induced by the ZLB.

Table 9 shows how the number of particles in NL-PF-5% affects the posterior estimates. We report the medians and (5%, 95%) credible sets of the marginal data densities and estimated param-

Ptr	Unconditional (128 datasets)			ZLB quarters ≥ 15 (13 datasets)		
	NL-PF-5%-20k	NL-PF-5%-40k	NL-PF-5%-80k	NL-PF-5%-20k	NL-PF-5%-40k	NL-PF-5%-80k
φ	106.3 (71.8, 144.3)	106.1 (71.9, 144.4)	106.2 (71.9, 144.3)	113.2 (78.3, 151.2)	113.1 (78.6, 150.6)	113.5 (79.5, 151.5)
ϕ_π	2.037 (1.655, 2.432)	2.037 (1.655, 2.430)	2.037 (1.658, 2.429)	2.056 (1.668, 2.444)	2.052 (1.663, 2.448)	2.053 (1.674, 2.452)
ϕ_y	0.513 (0.239, 0.823)	0.511 (0.239, 0.822)	0.512 (0.238, 0.824)	0.511 (0.251, 0.812)	0.510 (0.250, 0.811)	0.514 (0.254, 0.811)
h	0.487 (0.334, 0.622)	0.487 (0.334, 0.621)	0.487 (0.334, 0.622)	0.484 (0.352, 0.614)	0.485 (0.351, 0.614)	0.484 (0.354, 0.616)
ρ_a	0.814 (0.679, 0.885)	0.813 (0.677, 0.885)	0.814 (0.678, 0.885)	0.850 (0.761, 0.898)	0.849 (0.762, 0.898)	0.849 (0.764, 0.898)
ρ_i	0.815 (0.749, 0.866)	0.815 (0.749, 0.866)	0.815 (0.749, 0.866)	0.812 (0.752, 0.866)	0.811 (0.750, 0.866)	0.813 (0.751, 0.867)
σ_g	0.0136 (0.0104, 0.0178)	0.0136 (0.0104, 0.0178)	0.0136 (0.0104, 0.0178)	0.0137 (0.0110, 0.0183)	0.0137 (0.0109, 0.0183)	0.0137 (0.0109, 0.0184)
σ_a	0.0143 (0.0109, 0.0185)	0.0143 (0.0109, 0.0185)	0.0143 (0.0109, 0.0185)	0.0147 (0.0111, 0.0193)	0.0147 (0.0111, 0.0192)	0.0147 (0.0111, 0.0192)
σ_i	0.0022 (0.0018, 0.0028)	0.0022 (0.0018, 0.0028)	0.0022 (0.0018, 0.0028)	0.0023 (0.0017, 0.0029)	0.0023 (0.0017, 0.0029)	0.0023 (0.0017, 0.0029)
ℓ	1283.2 (1266.6, 1303.0)	1283.2 (1266.7, 1302.9)	1283.1 (1266.7, 1303.0)	1292.3 (1269.7, 1307.0)	1292.3 (1269.7, 1307.0)	1292.3 (1269.7, 1307.1)

Table 9: Posterior medians and (5%, 95%) credible sets of the estimated parameters and marginal data densities.

eters with 20,000, 40,000 (baseline specification), and 80,000 particles. We find almost no variation in the medians or credible sets, even when conditioning on datasets that spend at least 15 quarters at the ZLB. Although these results clearly depend on the model specification and the number of observables, they indicate that it is safe to use a smaller number of particles than what is traditionally used in the literature. Adapting the particle filter to the current observation and performing multiple runs of the filter for each draw most likely improved accuracy and played a major role in our results.

B NONLINEAR SOLUTION METHOD

We begin by compactly writing the detrended nonlinear equilibrium system from [section 2.1](#) 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, $\mathbf{s}_t \equiv [\tilde{c}_t, \tilde{y}_t^g, \tilde{y}_t, \tilde{\lambda}_t, w_t, i_t, i_t^n, \pi_t, a_t, g_t]'$ is a vector of variables, $\varepsilon_t \equiv [\varepsilon_{g,t}, \varepsilon_{a,t}, \varepsilon_{i,t}]'$ is a vector of shocks, $\mathbf{z}_t \equiv [g_t, a_t, \varepsilon_{i,t}, i_{t-1}^n, \tilde{c}_{t-1}]'$ is a vector of state variables, and $\vartheta \equiv [\beta, \eta, \epsilon, \bar{n}, \bar{s}, \bar{g}, \bar{\pi}, \varphi, \phi_\pi, \phi_y, h, \rho_a, \rho_i, \sigma_g, \sigma_a, \sigma_i]'$ is a vector of all the model parameters.

There are many ways to discretize the exogenous state variables, g_t , a_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 i_{t-1}^n and c_{t-1} are respectively set to $\pm 2.5\%$ and $\pm 4\%$ of the deterministic steady state to contain the filtered state variables based on posterior draws. We discretize the state variables into (8, 12, 8, 7, 7) evenly-spaced points. In

total, there are $D = 37,632$ nodes in the state space, and the realization of \mathbf{z}_t on node d is denoted $\mathbf{z}_t(d)$. The Rouwenhorst method provides integration nodes, $[g_{t+1}(m), a_{t+1}(m), \sigma_{i,t+1}(m)]$, with weights, $\phi(m)$, for $m \in \{1, \dots, M\}$. Since the exogenous variables evolve according to a Markov chain, the number of future realizations is the same as the state variables, $(8, 12, 8)$, or $M = 768$.

For the policy functions, we approximate $\tilde{c}_t(\mathbf{z}_t)$ and $\pi_t^{gap}(\mathbf{z}_t) \equiv \pi_t(\mathbf{z}_t)/\bar{\pi}$. 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. Using Sims's (2002) `gensys` algorithm to solve the log-linear model without the ZLB imposed. Then map that solution onto the discretized state space to initialize \tilde{c}_0 and π_0^{gap} .
2. On iteration $j \in \{1, \dots, N_{iter}\}$ and for each $d \in \{1, \dots, D\}$, use Chris Sims's `csolve` to find \tilde{c}_t and π_t^{gap} to satisfy $\mathbb{E}[f(\cdot)|\mathbf{z}_t(d), \vartheta] \approx 0$, where N_{iter} is the number of iterations. Guessing $\tilde{c}_t = \tilde{c}_{j-1}(d)$ and $\pi_t^{gap} = \pi_{j-1}^{gap}(d)$, approximate $\mathbb{E}[f(\cdot)|\mathbf{z}_t(d), \vartheta]$ as follows:
 - (a) Solve for $\{\tilde{y}_t, i_t^n, i_t, \tilde{\lambda}_t, \tilde{w}_t\}$ given \tilde{c}_t, π_t^{gap} , and $\mathbf{z}_t(d)$.
 - (b) Linearly interpolate the policy functions, \tilde{c}_{j-1} and π_{j-1}^{gap} , at the updated state variables, $\mathbf{z}_{t+1}(m)$, to obtain $\tilde{c}_{t+1}(m)$ and $\pi_{t+1}^{gap}(m)$ on every integration node, $m \in \{1, \dots, M\}$.
 - (c) Given $\{\tilde{c}_{t+1}(m), \pi_{t+1}^{gap}(m)\}_{m=1}^M$, solve for the other elements of $\mathbf{s}_{t+1}(m)$ and compute:

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

When `csolve` converges, set $\tilde{c}_j(d) = \tilde{c}_t$ and $\pi_j^{gap}(d) = \pi_t^{gap}$.

3. Repeat step 2 until $\text{maxdist}_j < 10^{-6}$, where $\text{maxdist}_j \equiv \max\{|\tilde{c}_j - \tilde{c}_{j-1}|, |\pi_j^{gap} - \pi_{j-1}^{gap}|\}$. When that occurs, the algorithm has converged to an approximate nonlinear solution.

C ESTIMATION ALGORITHM

We use a random walk Metropolis-Hastings algorithm to estimate our model with artificial data of 100 quarters. To measure how well the model fits the data, we use the adapted particle filter described in Algorithm 12 in Herbst and Schorfheide (2016), which modifies the basic filter in Stewart and McCarty (1992) and Gordon et al. (1993) to better account for the outliers in the data.

C.1 METROPOLIS-HASTINGS ALGORITHM The following steps outline the algorithm:

1. Specify the prior distributions, means, variances, and bounds of each element of the vector of N_e estimated parameters, $\theta \equiv [\varphi, \phi_\pi, \phi_y, h, \rho_a, \rho_i, \sigma_g, \sigma_a, \sigma_i]'$.

2. The artificial data consists of real GDP growth, the inflation rate, and the nominal interest rate, $\mathbf{x}_t \equiv [\tilde{y}_t^g, \pi_t, i_t]'$, where $N_x = 3$ is the number of observable variables.
3. Find the posterior mode to initialize the preliminary Metropolis-Hastings step.

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

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

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

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

- ii. Given $\hat{\theta}_i$, solve the model according to [Appendix B](#). If the algorithm converges, then compute the stochastic steady state, otherwise repeat step 3(a)i and redraw $\hat{\theta}_i$.
- iii. If the stochastic steady state exists, then use the particle filter in [section C.2](#) to obtain the log-likelihood value for the model, $\log \ell_i^{model}$, otherwise repeat step 3(a)i.
- iv. The posterior log-likelihood is $\log \ell_i^{post} = \log \ell_i^{prior} + \log \ell_i^{model}$

(b) Calculate $\max(\log \ell_1^{post}, \dots, \log \ell_{N_m}^{post})$ 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 $i \in \{0, \dots, N_d\}$, 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. Given $\hat{\theta}_i^{cand}$, solve the model according to [Appendix B](#). If the algorithm converges, compute the stochastic steady state, otherwise repeat 5(a)i and draw a new $\hat{\theta}_i^{cand}$.

- iv. If the stochastic steady state exists, then use the particle filter in [section C.2](#) to obtain the log-likelihood value for the model, $\log \ell_i^{model}$, otherwise repeat 5(a)i.
- 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 $N_b = 5000$ draws and use the remaining sample to calculate the mean draw, $\bar{\theta}^{preMH} = \sum_{i=N_b+1}^{N_{preMH}} \hat{\theta}_i$, and the covariance matrix, Σ^{preMH} . We follow step 4 to calculate Σ^{preMH} but use all $N_d - N_b$ draws instead of just the upper p th percentile.

6. Following the procedure in step 5, perform a final run of the Metropolis-Hastings algorithm, where $\hat{\theta}_0 = \bar{\theta}^{preMH}$ and $\Sigma = \Sigma^{preMH}$. We set $N_d = 100,000$ and keep every 100th draw. The remaining 1,000 draws form a representative sample from the joint posterior density.

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 $p \in \{0, \dots, N_p\}$ 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 $t \in \{1, \dots, T\}$, sequentially filter the data with the linear or nonlinear model as follows:
 - (a) For $p \in \{1, \dots, N_p\}$, 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 $R \equiv \text{diag}(\sigma_{me,\bar{y}^g}^2, \sigma_{me,\pi}^2, \sigma_{me,i}^2)$ is the measurement error covariance matrix.
 - iii. The probability of observing the current state, \mathbf{s}_t , given $\bar{\mathbf{s}}_{t-1}$, 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 $p(\xi_t|\mathbf{s}_t)p(\mathbf{s}_t|\bar{\mathbf{s}}_{t-1}) \propto \exp(-\xi_t'R^{-1}\xi_t/2) \exp(-\varepsilon_t'\bar{\varepsilon}_t/2)$ by solving for the optimal $\bar{\varepsilon}_t$. We converted MATLAB's `fminsearch` routine 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 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 log-likelihood is $\log \ell^{model} = \sum_{t=1}^T \log \ell_t^{model}$.

C.3 UNSCENTED KALMAN FILTER A well-known alternative to the Unscented Kalman Filter (UKF) is the Extended Kalman Filter (EKF). We focus on UKF instead of EKF because EKF does not perform well when the state equation is poorly approximated with linearization, which occurs in our model when the notional rate is negative. EKF relies on a linear approximation of the state equation to predict the covariance of the state vector, whereas UKF nonlinearly propagates a deterministic set of points, more accurately approximating the mean and covariance of the state vector.

1. Initialize the state vector, \mathbf{s}_0 , at the stochastic steady state. Initialize the state covariance matrix, P_0 , so that it solves the Lyapunov equation $P_0 - T'P_0T = MM'$, where T is the state transition matrix and M is the shock impact matrix of the solution to the linear model.
2. For $t \in \{1, \dots, T\}$, sequentially filter the data with the nonlinear model as follows:
 - (a) Augment the estimated state vector and covariance matrix.
 - i. Augment the state vector to include the expected value of the shocks and measurement error, $\mathbf{s}_{t-1}^a \equiv [\mathbf{s}_{t-1}', E(\varepsilon_t'), E(\xi_t)']' = [\mathbf{s}_{t-1}', \mathbf{0}, \mathbf{0}]'$, which has N_{s^a} elements.
 - ii. Similarly, augment the state covariance matrix, P , with the shocks covariance matrix, I (the identity matrix), and the measurement error covariance matrix, R ,

$$P_{t-1}^a \equiv \begin{bmatrix} P_{t-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & R \end{bmatrix}.$$

iii. Create a distribution of augmented state vectors with $2N_{\mathbf{s}^a} + 1$ sigma points,

$$\mathbf{s}_{t-1,i}^a = \begin{cases} \mathbf{s}_{t-1}^a & \text{for } i = 0, \\ \mathbf{s}_{t-1}^a - \sqrt{3}L_i & \text{for } i = 1, \dots, N_{\mathbf{s}^a}, \\ \mathbf{s}_{t-1}^a + \sqrt{3}L_{i-N_{\mathbf{s}^a}} & \text{for } i = N_{\mathbf{s}^a} + 1, \dots, 2N_{\mathbf{s}^a}, \end{cases}$$

where L_i is i th column column of the lower triangular Cholesky factor L of P_{t-1}^a , such that $P_{t-1}^a = LL'$. The respective weights are

$$\omega_i = \begin{cases} 1 - N_{\mathbf{s}^a}/3 & \text{for } i = 0, \\ 1/6 & \text{for } i = 1, \dots, 2N_{\mathbf{s}^a}. \end{cases}$$

(b) For $i = 0, \dots, 2N_{\mathbf{s}^a}$, first use the nonlinear solution to update the state vector, $\mathbf{s}_{t|t-1,i}$ given $\mathbf{s}_{t-1,i}^a$, and then compute the predicted state vector and covariance matrix,

$$\mathbf{s}_{t|t-1} = \sum_{i=0}^{2N_{\mathbf{s}^a}} \omega_i \mathbf{s}_{t|t-1,i}, \quad P_{t|t-1} = \sum_{i=0}^{2N_{\mathbf{s}^a}} \omega_i (\mathbf{s}_{t|t-1,i} - \mathbf{s}_{t|t-1})(\mathbf{s}_{t|t-1,i} - \mathbf{s}_{t|t-1})'.$$

This prediction step mirrors the basic Kalman filter, except the predictions are weighted averages across the sigma points based on the augmented state vector and covariance matrix. Given the predictions, calculate the Kalman gain, $K_t = P_{t|t-1}H'S_t^{-1}$, where

$$S_t = \sum_{i=0}^{2N_{\mathbf{s}^a}} \omega_i (\mathbf{s}_{t|t-1,i}^h - \mathbf{s}_{t|t-1}^h)(\mathbf{s}_{t|t-1,i}^h - \mathbf{s}_{t|t-1}^h)',$$

$\mathbf{s}_{t|t-1,i}^h = H\mathbf{s}_{t|t-1,i}$, $\mathbf{s}_{t|t-1}^h = H\mathbf{s}_{t|t-1}$, and the measurement error, $\xi_t = \mathbf{x}_t - \mathbf{s}_{t|t-1}^h$.

(c) Update the state vector and covariance matrix estimates,

$$\begin{aligned} \mathbf{s}_t &= \mathbf{s}_{t|t-1} + K_t \xi_t, \\ P_t &= P_{t|t-1} - K_t S_t K_t'. \end{aligned}$$

This update step is the same as the basic Kalman filter. The time- t contribution to the log-likelihood is $\log \ell_t^{model} = -(\log |S_t| + \xi_t' S_t^{-1} \xi_t + N_x \log 2\pi)/2$.

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

D GENERALIZED IMPULSE RESPONSE FUNCTIONS

The general procedure for calculating GIRFs is described in Koop et al. (1996). The GIRFs are based on the average path from repeated model simulations and generated with the following steps:

1. Initialize each simulation with the desired state vector, \mathbf{s}_0 (e.g., the stochastic steady state). To calculate the stochastic steady state, we simulate the model without shocks. The stochastic steady state is defined as the state vector where $|\mathbf{s}_t - \mathbf{s}_{t-1}|$ converges to zero, which differs from the deterministic steady state since the policy functions embed higher order moments.
2. Draw random growth rate, preference, and monetary policy shocks, $\{\varepsilon_t\}_{t=0}^N$, for each simulation, where N is the simulation length. From the initial state vector, \mathbf{s}_0 , simulate N_{GIRF} equilibrium paths, $\{\mathbf{x}_t^j(\mathbf{s}_0)\}_{t=0}^N$, where $j \in \{1, 2, \dots, N_{GIRF}\}$ and $N_{GIRF} = 10,000$.
3. Using the same draws of shocks from step 2, replace the preference shock in period one with a 2 standard deviation shock (i.e., set $\varepsilon_{a,1} = 2\sigma_a$). Then simulate the model with these alternate sequences of shocks to obtain N_{GIRF} equilibrium paths, $\{\mathbf{x}_t^j(\mathbf{s}_0, \varepsilon_{a,1})\}_{t=0}^N$.
4. Average across the N_{GIRF} simulations from steps 2 and 3 to obtain average paths given by

$$\bar{\mathbf{x}}_t(\mathbf{s}_0) = N_{GIRF}^{-1} \sum_{j=1}^{N_{GIRF}} \mathbf{x}_t^j(\mathbf{s}_0), \quad \bar{\mathbf{x}}_t(\mathbf{s}_0, \varepsilon_{a,1}) = N_{GIRF}^{-1} \sum_{j=1}^{N_{GIRF}} \mathbf{x}_t^j(\mathbf{s}_0, \varepsilon_{a,1}).$$

5. The difference between the two average paths is a GIRF.