DALLASFED **Working Paper**

Are Nonlinear Methods Necessary at the Zero Lower Bound?

Alexander W. Richter and Nathaniel A. Throckmorton

Federal Reserve Bank of Dallas
Research Department
Working Paper 1606

Are Nonlinear Methods Necessary at the Zero Lower Bound?*

Alexander W. Richter   Nathaniel A. Throckmorton

November 11, 2016

ABSTRACT

This paper examines the importance of using nonlinear methods to account for the zero lower bound (ZLB) on the Fed's policy rate. We estimate three models with a particle filter: (1) a nonlinear model with a ZLB constraint; (2) a constrained linear model that imposes the constraint in the filter but not the solution; and (3) an unconstrained linear model that never imposes the constraint. The linear models have a lower likelihood than the nonlinear model when the Fed is constrained and predict large monetary policy shocks during the ZLB period. We also compare the predictions from our nonlinear model to the quasi-linear solution with OccBin. OccBin captures the ZLB much better than the linear solutions but it still generates less endogenous volatility than the nonlinear model and it is not as conducive to estimation. Finally, we extend the baseline model to include a banking sector. We find larger differences between the predictions from the nonlinear model and both the linear and quasi-linear models.

Keywords: Bayesian Estimation; Model Comparison; Zero Lower Bound; Particle Filter
JEL Classifications: C11; E43; E58

*Richter, Research Department, Federal Reserve Bank of Dallas, and Auburn University, 2200 N. Pearl Street, Dallas, TX 75201 (alex.richter@dal.frb.org); Throckmorton, Department of Economics, College of William & Mary, P.O. Box 8795, Williamsburg, VA 23187 (nathrockmorton@wm.edu). We thank our discussant, Pablo Cuba-Borda, as well as Ben Keen, Mike Plante, and Todd Walker for suggestions that improved the paper. We also thank Auburn University and Bradley Morgan supporting the Hopper computer cluster. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 INTRODUCTION

Long after central banks raise their policy rates, the extended period of near zero rates will remain an important and challenging feature of the data. Of central importance to both current and future work on monetary policy is the question of how to deal with the zero lower bound (ZLB) constraint on the nominal interest rate. Researchers have three options when estimating a model: ignore the constraint, account for the constraint only in the prediction step, or impose the constraint in both the prediction and estimation phases. Recent studies of the ZLB period use a variety of models, solution methods, and estimation procedures, which differ in their treatment of the ZLB constraint.

This paper examines the importance of the ZLB constraint by estimating three versions of a small-scale New Keynesian model with a particle filter: (1) a nonlinear model with an occasionally binding ZLB constraint; (2) a constrained linear model, which imposes the constraint in the filter but not the solution; and (3) an unconstrained linear model, which never imposes the constraint. The first model serves as our benchmark because it is the most comprehensive in its treatment of the constraint. In that model, households' expectations account for the possibility of going to and leaving the ZLB, which depends on the state of the economy as well as current and future shocks. The drawback with this method is that the cost of evaluating the likelihood function increases with the size of the model. In the second model, households do not account for the constraint in their decision rules, but it is imposed in simulations of the model to prevent negative realizations of the policy rate. By disregarding the effects of the constraint on households' decisions, it is easy to solve the model using linear methods, which makes it quicker to estimate. The third model completely ignores the constraint when solving and simulating the model. We estimate each model with U.S. data by embedding the particle filter into a Metropolis-Hastings algorithm.

A few papers examine the impact of the solution in a calibrated model [Braun et al. (2012); Fernández-Villaverde et al. (2015); Gavin et al. (2015); Nakata (2012)]. Those papers compare the policy functions and impulse responses across a variety of solution methods. While they find meaningful differences between linear and nonlinear models, it is hard to assess their quantitative importance without taking the models to the data. Using artificial data from a linear New Keynesian model that uses news shocks to impose the ZLB, Hirose and Inoue (2016) find that ignoring the ZLB biases the parameter estimates as the frequency and duration of ZLB events increases. Gust et al. (2016) estimate a nonlinear model with U.S. data to show the empirical implications of the ZLB constraint. They stress the importance of using nonlinear estimation techniques by comparing the posterior distributions and impulse responses from the nonlinear and linear models. We build on their work by providing a detailed account of the shocks that kept the economy at the ZLB and showing how well each model fits the data in the ZLB period. To conduct our analysis, we compare the posterior distributions, the marginal data densities, the filtered observables and shocks, the likelihoods, and the posterior predictive distributions of various moments from our three models.

Surprisingly, the posterior distributions and marginal likelihoods from the nonlinear and constrained linear models are similar, but important differences arise in their predictions at the ZLB. The nonlinear model has a higher likelihood in periods when the Fed is constrained and primarily attributes the ZLB to a reduction in demand due to discount factor shocks. In both linear models, large contractionary monetary policy shocks are also needed to explain data in the ZLB period, which is at odds with the Fed's policy during the Great Recession. A comparison of the posterior predictive distributions shows the three models match the data equally well in the pre-ZLB period. In the ZLB period, however, the nonlinear model predicts higher output volatility and negative skewness in output and inflation, two key features of the data when the Fed was constrained. In contrast, neither of the linear models predict a change in volatility or skewness in the ZLB period.

Several papers examine the financial crisis and the ensuing Great Recession through the lens of an unconstrained linear model. For example, Ireland (2011) compares the shocks that caused the 1991, 2001, and 2007-2009 recessions using maximum likelihood estimates from an unconstrained, small-scale, linear New Keynesian model. He finds that a major difference between the three recessions is that the most recent recession was plagued by large monetary policy shocks. Suh and Walker (2016) use Bayesian methods to estimate an unconstrained, medium-scale, linear New Keynesian model with financial frictions. They find monetary policy shocks played a major role in explaining the changes in consumption and investment during the Great Recession. Unconstrained linear models are also widely used by the Fed for policy analysis and forecasting [e.g., Del Negro et al. (2013)]. Although the ZLB period lasted for seven years, our estimates indicate that the Fed was only constrained from 2008Q4 to 2011Q1. If the data used to estimate a model includes that period, it is crucial to use nonlinear estimation techniques. A linear model will lead to incorrect predictions about the causes and consequences of the ZLB because it does not capture the endogenous effects of the constraint, even if the constraint is imposed when filtering the data.

The nonlinear solution provides the most accurate way to characterize the dynamics just before and after the ZLB binds, but it is computationally expensive. An alternative solution developed by Guerrieri and Iacoviello (2015) is based on a quasi-linear version of the nonlinear model, where the constraint binds in one regime and is slack in the other regime. The benefit of this method is that it is as fast as linear methods that ignore the ZLB constraint. The authors also developed a toolbox called OccBin that is compatible with Dynare and easy to use. Unfortunately, those advantages come with two key drawbacks. First, households do not account for the expectational effects of going to the ZLB in their decisions, which causes households to act as if the constraint will never bind even if it is likely to bind in the near future. It also lowers the frequency of ZLB events, so households' decisions are less sensitive to the constraint. Second, the economy must return to the regime where the ZLB is slack when simulating the model, which makes it prohibitively expensive to estimate the model with a particle filter. As an alternative, Guerrieri and Iacoviello (2016) follow Fair and Taylor (1983) and use a deterministic filter that solves for the shocks that best match the data each period. The filter is fast and it admits a closed-form solution for the likelihood function.

We compare the impulse responses and filtered shocks from our nonlinear solution to the quasi-linear solution based on OccBin and then compare the deterministic filter to the particle filter. The quasi-linear model is better able to capture the increased volatility at the ZLB than the linear model, but it does not generate as much endogenous volatility as the nonlinear model. The quasi-linear model also requires larger shocks than the nonlinear model to explain the start of the ZLB period, just like the linear models. To test the accuracy of the deterministic filter against the particle filter, we estimate our unconstrained linear model with a Kalman filter, since it is optimal for linear systems. We find the particle filter produces nearly identical parameter estimates as the Kalman filter when the two filters have the same measurement error variance. In contrast, the deterministic filter produces parameter estimates that are outside the credible sets implied by the Kalman filter.

We conclude our analysis by extending our model to include additional states, shocks, and observables by adding a banking sector following Cúrdia and Woodford (2010). There are larger differences between the linear and nonlinear posterior estimates and the nonlinear model now has a significantly higher data density. Also, the differences between the impulse responses and filtered shocks are larger when comparing the nonlinear model to both the linear and quasi-linear models.

The paper proceeds as follows. Section 2 describes our model. Section 3 outlines our solution and estimation procedures. Section 4 shows the differences between the nonlinear and linear models by comparing the posterior distributions and filtered paths. Section 5 compares the dynamics from our nonlinear model to the quasi-linear model with OccBin. Section 6 adds a banking sector to our model and compares the nonlinear, linear, and quasi-linear solutions. Section 7 concludes.

2 STRUCTURAL MODEL

This section lays out a small-scale New Keynesian model with two endogenous state variables and three shocks. We estimate three versions of the model: (1) the nonlinear version, which imposes the ZLB constraint; (2) a linear analogue of the nonlinear model, which includes the constraint in the filter but not the solution; and (3) a linear analogue that completely removes the ZLB constraint.

2.1 HOUSEHOLDS A representative household chooses $\{c_t, n_t, b_t\}_{t=0}^{\infty}$ to maximize expected lifetime utility, $E_0 \sum_{t=0}^{\infty} \tilde{\beta}_t [\log(c_t - hc^a_{t-1}) - \chi n_t^{1+\eta}/(1+\eta)]$, where $\chi > 0$, $1/\eta$ 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 an expectation operator conditional on information in period 0, $\tilde{\beta}_0 \equiv 1$, and $\tilde{\beta}_t = \prod_{j=1}^{t>0} \beta_j$. To introduce fluctuations in the real interest rate, the discount factor, β, is time-varying and follows

$$\log \beta_t = (1-\rho_\beta) \log \bar{\beta} + \rho_\beta \log \beta_{t-1} + \sigma_\upsilon \upsilon_t, \ 0 \le \rho_\beta < 1, \ \upsilon \sim \mathbb{N}(0,1), \tag{1}$$

where $\bar{\beta}$ is the discount factor along the steady state growth path. The choices are constrained by $c_t + b_t = w_t n_t + i_{t-1} b_{t-1}/\pi_t + d_t$, where π is the gross inflation rate, w is the real wage rate, i is the gross nominal interest rate, and d is a real dividend. The household's optimality conditions imply

$$w_t = \chi n_t^\eta (c_t - hc^a_{t-1}) \quad \text{and} \quad 1 = i_t E_t[q_{t,t+1}/\pi_{t+1}],$$

where $q_{t,t+1} \equiv \beta_{t+1}(c_t - hc^a_{t-1})/(c_{t+1} - hc^a_t)$ is the pricing kernel between periods t and $t+1$.

2.2 FIRMS The production sector consists of a continuum of monopolistically competitive intermediate goods firms owned by households 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. The deviations from the steady state growth rate, $\bar{g}$, follow

$$\log g_t = (1-\rho_g) \log \bar{g} + \rho_g \log g_{t-1} + \sigma_\varepsilon \varepsilon_t, \ 0 \le \rho_g < 1, \ \varepsilon \sim \mathbb{N}(0,1). \tag{2}$$

Each intermediate firm chooses its labor to minimize its costs, $w_t n_t(f)$, subject to its production function. The final goods firm purchases $y_t(f)$ units from each intermediate firm to produce the final good, $y_t \equiv [\int_0^1 y_t(f)^{(\theta-1)/\theta} df]^{\theta/(\theta-1)}$, where $\theta > 1$ measures the elasticity of substitution between intermediate goods. It then maximizes dividends to determine its demand function for intermediate good f, $y_t(f) = (p_t(f)/p_t)^{-\theta} y_t$, where $p_t = [\int_0^1 p_t(f)^{1-\theta} df]^{1/(1-\theta)}$ is the price level.

Following Rotemberg (1982), each intermediate firm pays a cost to adjust its price, $adj_t(f) = \varphi[p_t(f)/(\bar{\pi} p_{t-1}(f)) - 1]^2 y_t/2$, where $\varphi > 0$ scales the cost and $\bar{\pi}$ is the gross inflation rate along the steady state growth path. Firm f chooses its price, $p_t(f)$, to maximize the expected present value of future dividends, $E_t \sum_{k=t}^{\infty} q_{t,k} d_k(f)$, where $q_{t,t} \equiv 1$, $q_{t,k} \equiv \prod_{j=t+1}^{k>t} q_{j-1,j}$, and $d_t(f) = (p_t(f)/p_t) y_t(f) - w_t n_t(f) - adj_t(f)$. In symmetric equilibrium, the optimality condition implies

$$\varphi(\hat{\pi}_t - 1)\hat{\pi}_t = 1 - \theta + \theta(w_t/z_t) + \varphi E_t[q_{t,t+1}(\hat{\pi}_{t+1} - 1)\hat{\pi}_{t+1}(y_{t+1}/y_t)],$$

where $\hat{\pi}_t \equiv \pi_t/\bar{\pi}$. When $\varphi = 0$, $w_t/z_t = (\theta-1)/\theta$, which is the inverse of the gross price markup.

2.3 MONETARY POLICY The central bank sets the gross nominal interest rate according to

$$i_t = \max\{\underline{\imath}, i_t^*\},\ i_t^* = (i_{t-1}^*)^{\rho_i}(\bar{\imath}\hat{\pi}_t^{\phi_\pi}(c_t/(\bar{g}c_{t-1}))^{\phi_c})^{1-\rho_i}\exp(\sigma_\nu\nu_t),\ 0 \leq \rho_i < 1,\ \nu \sim \mathbb{N}(0,1),$$

where $\underline{\imath}$ is the lower bound, i^* is the notional rate, ϕ_π and ϕ_c are the responses to deviations of inflation from target and deviations of consumption growth from its steady state, and $\bar{\imath}$ and $\bar{\pi}$ are the inflation and interest rate targets, which equal their values along the steady state growth path.

The treatment of the ZLB constraint will influence the estimated monetary policy shocks that explain the data. It is important to note that policy shocks affect the notional rate, not the nominal rate. In the nonlinear model, a positive policy shock at the ZLB leads to a higher notional rate and, due to smoothing, higher than expected future notional rates, which can exceed the ZLB. That shock causes a decrease in current real GDP growth and inflation, which affects the model likelihood. In the linear models, monetary policy shocks have similar effects on real GDP growth and inflation regardless of the notional rate. If the ZLB is imposed in the filter, then the filter is able to distinguish between the notional and nominal rates. In other words, if the ZLB binds in the data and the model predicts a negative notional rate, then the filter may be able to match the prediction to the data without requiring a positive policy shock to explain the difference, although other factors may still lead to alternative sequences of shocks. The unconstrained linear model, however, cannot distinguish between the two rates. In that case, if the model predicts a negative nominal rate, positive policy shocks are required to make the nominal rate consistent with the data.

2.4 COMPETITIVE EQUILIBRIUM To make the model stationary, we redefine all of the variables that grow in terms of technology (i.e., $\tilde{x}_t \equiv x_t/z_t$). The detrended equilibrium system consists of

$$\tilde{\lambda}_t = \tilde{c}_t - h\tilde{c}_{t-1}/g_t, \tag{3}$$

$$\tilde{w}_t = \chi\tilde{y}_t^\eta\tilde{\lambda}_t, \tag{4}$$

$$1 = i_t E_t[\beta_{t+1}(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(1/(g_{t+1}\bar{\pi}\hat{\pi}_{t+1}))], \tag{5}$$

$$i_t^* = (i_{t-1}^*)^{\rho_i}(\bar{\imath}\hat{\pi}_t^{\phi_\pi}(g_t\tilde{c}_t/(\bar{g}\tilde{c}_{t-1}))^{\phi_c})^{1-\rho_i}\exp(\sigma_\nu\nu_t), \tag{6}$$

$$\tilde{c}_t = [1 - \varphi(\hat{\pi}_t - 1)^2/2]\tilde{y}_t, \tag{7}$$

$$\varphi(\hat{\pi}_t - 1)\hat{\pi}_t = (1-\theta) + \theta\tilde{w}_t + \varphi E_t[\beta_{t+1}(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(\hat{\pi}_{t+1} - 1)\hat{\pi}_{t+1}(\tilde{y}_{t+1}/\tilde{y}_t)], \tag{8}$$

the ZLB constraint, and the stochastic processes, which impose the bond market clearing condition, $b_t = 0$, and the aggregation rule, $\tilde{c}_t = \tilde{c}_t^a$. A competitive equilibrium includes sequences of quantities, $\{\tilde{\lambda}_t, \tilde{c}_t, \tilde{y}_t\}_{t=0}^\infty$, prices, $\{w_t, i_t, i_t^*, \hat{\pi}_t\}_{t=0}^\infty$, and exogenous variables, $\{\beta_t, g_t\}_{t=0}^\infty$, that satisfy the detrended system, given the initial conditions, $\{\tilde{c}_{-1}, i_{-1}^*, \beta_0, g_0, \nu_0\}$, and the shocks, $\{\varepsilon_t, \upsilon_t, \nu_t\}_{t=1}^\infty$.

3 SOLUTION METHODS AND ESTIMATION PROCEDURE

This section concisely describes our solution methods and outlines the estimation procedure applied to all three models. See Plante et al. (2016) for a more detailed description of both algorithms.

3.1 SOLUTION METHODS We first solve the log-linear version of our nonlinear model with Sims's (2002) algorithm. Using that solution as an initial conjecture, we then 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). Each iteration, we minimize the Euler equation errors on every node in the state space. We then compute the maximum distance between the policy functions on any node and continue iterating until that distance falls below the tolerance criterion. We approximate the three exogenous processes with an N-state Markov chain following Rouwenhorst (1995) and use piecewise linear interpolation to calculate future variables.

Steady-State Discount Factor	$\bar{\beta}$	0.9987	Real GDP Growth Rate Measurement Error SD	$\sigma_{me,\hat{y}}$	0.00190
Frisch Elasticity of Labor Supply	$1/\eta$	3	Inflation Rate Measurement Error SD	$\sigma_{me,\pi}$	0.00077
Elasticity of Substitution between Goods	θ	6	Federal Funds Rate Measurement Error SD	$\sigma_{me,i}$	0.00210
Steady-State Labor	$\bar{n}$	0.33	Number of Particles	N_p	40,000
Nominal Interest Rate Lower Bound	$\underline{\imath}$	1.00035	Number of Posterior Draws	N_d	100,000

Table 1: Calibrated parameters for the model and estimation procedure.

3.2 ESTIMATION PROCEDURE

We estimate our models with quarterly data on per capita real GDP ($RGDP/CNP$), the GDP deflator (DEF), and the federal funds rate (FFR) from 1986Q1 to 2015Q4. Our sources are provided in Plante et al. (2016). The vector of observables is given by

$$\hat{\mathbf{x}}^{data} \equiv [\log(RGDP_t/CNP_t) - \log(RGDP_{t-1}/CNP_{t-1}), \log(DEF_t/DEF_{t-1}), \log(1 + FFR_t/100)/4].$$

We calibrate five poorly identified parameters (table 1). The steady-state discount factor, $\bar{\beta}$, is set to 0.9987, which equals $(1/T)\sum_{t=1}^{T}(1 + G_t/400)(1 + \Pi_t)/(1 + FFR_t/100)^{1/4}$ where T is the sample size, G_k is the annual utilization-adjusted growth rate of technology from Fernald (2012) and $\Pi_k = \log(DEF_k/DEF_{k-1})$. The leisure preference parameter, χ, is set so that steady-state labor equals $1/3$ of the available time. The elasticity of substitution between intermediate goods, θ, is set to 6, which corresponds to an average markup over marginal cost equal to 20%. The lower bound on the nominal interest rate, $\underline{\imath}$, is calibrated to 1.00035, which is the average federal funds rate. The Frisch labor supply elasticity, $1/\eta$, is set to 3, to match the estimate in Peterman (2016).

We use a random walk Metropolis-Hastings algorithm and a particle filter to evaluate the model likelihood following Fernández-Villaverde and Rubio-Ramírez (2007). However, we follow Herbst and Schorfheide (2016) and adapt the filter to include information from the current period, which helps the model match outliers during the Great Recession. The filter uses $40{,}000$ particles and systematic resampling with replacement following Kitagawa (1996). We convert the predictions of the linear models to levels, so we can apply the exact same filter to all three models. Given the simulated paths from each model, we transform the predictions for real GDP growth, inflation, and the policy rate according to $\hat{\mathbf{x}}_t^{model} = [\log(g_t\tilde{c}_t/\tilde{c}_{t-1}), \log(\pi_t), \log(i_t)]$. The observables contain measurement error (ME), so $\hat{\mathbf{x}}_t^{data} = \hat{\mathbf{x}}_t^{model} + \xi_t$, where $\xi \sim \mathbb{N}(0, \Sigma)$ is a vector of MEs and $\Sigma = \mathrm{diag}([\sigma_{me,\hat{y}}^2, \sigma_{me,\pi}^2, \sigma_{me,i}^2])$. The variance of each ME is set to 10% of the variance of the data (table 1). We obtain $100{,}000$ draws from the joint posterior distribution and keep every 100th draw.

The entire algorithm is programmed in Fortran using Open MPI and executed on a cluster. We parallelize the nonlinear solution by distributing the nodes in the state space across the available processors. To increase the accuracy of the filter, we calculate the posterior likelihood on each processor and evaluate whether to accept or reject a candidate draw based on the median likelihood.

4 MODEL COMPARISON

This section evaluates the performance of the nonlinear, constrained linear, and unconstrained linear models across a number of dimensions. Specifically, we show the posterior distributions, the predicted observables and shocks, the impulse responses, a time series of the filter densities, the posterior predictive distributions of three key moments, and the predicted duration of ZLB events.

	Prior		Posterior Mean $(5\%, 95\%)$		
Parameter	Dist	Mean (SD)	Nonlinear	Constrained Linear	Unconstrained Linear
φ	Gam	80.000 (20.000)	96.46409 (65.73921, 130.39517)	90.61107 (59.96913, 124.07097)	89.75551 (60.05718, 124.99871)
h	Beta	0.500 (0.200)	0.46334 (0.33583, 0.58604)	0.44752 (0.30823, 0.58102)	0.43851 (0.29796, 0.56827)
ϕ_π	Norm	2.500 (1.000)	4.07825 (3.32372, 4.85669)	4.12383 (3.30709, 5.01204)	3.74194 (3.02972, 4.53073)
ϕ_c	Norm	1.000 (0.400)	1.46414 (1.10608, 1.85105)	1.37493 (1.01308, 1.79434)	1.24805 (0.90772, 1.63401)
$\bar{g}$	Norm	1.004 (0.001)	1.00377 (1.00259, 1.00494)	1.00371 (1.00251, 1.00491)	1.00371 (1.00256, 1.00480)
$\bar{\pi}$	Norm	1.006 (0.001)	1.00613 (1.00558, 1.00669)	1.00613 (1.00549, 1.00678)	1.00596 (1.00534, 1.00658)
ρ_g	Beta	0.500 (0.200)	0.19499 (0.05748, 0.37778)	0.20000 (0.06326, 0.37596)	0.19183 (0.05145, 0.37168)
ρ_β	Beta	0.500 (0.200)	0.90287 (0.86884, 0.92903)	0.93649 (0.89556, 0.97084)	0.92928 (0.88839, 0.96493)
ρ_i	Beta	0.500 (0.200)	0.81450 (0.75752, 0.86385)	0.83577 (0.78382, 0.87841)	0.83983 (0.78855, 0.88403)
σ_ε	IGam	0.010 (0.010)	0.00972 (0.00747, 0.01235)	0.00944 (0.00722, 0.01198)	0.00938 (0.00729, 0.01189)
σ_υ	IGam	0.010 (0.010)	0.00214 (0.00158, 0.00286)	0.00212 (0.00160, 0.00278)	0.00191 (0.00143, 0.00257)
σ_ν	IGam	0.010 (0.010)	0.00196 (0.00146, 0.00258)	0.00189 (0.00141, 0.00247)	0.00177 (0.00133, 0.00227)
$\log(ML)$			1586.01	1586.25	1578.04

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

4.1 POSTERIOR DISTRIBUTIONS The first three columns of table 2 show the parameters and prior distributions, which are relatively diffuse and broadly consistent with other papers that use Bayesian methods. The remaining columns compare the posterior means and 90% credible sets for our three models. Interestingly, most of the posterior means for the nonlinear model are well within the credible sets of both linear models, and there is little difference in the extreme quantiles. One exception is the persistence of the discount factor. The tail of the credible set in the nonlinear model barely includes the posterior mean from the linear models, which is important since it affects the frequency and duration of ZLB events. There are also differences in the two policy parameters.

The last row in table 2 reports the marginal data density, which is based on Geweke's (1999) harmonic mean estimator. The exponential of the difference between any two of the values is the Bayes factor, which equals the posterior odds ratio since the prior distribution over our three models is uniform. The Bayes factor between the nonlinear and constrained linear models indicates that neither model provides a superior empirical fit over the entire sample, although, as we will show, there are meaningful differences in likelihoods during the ZLB period. These models, however, are superior to the unconstrained linear model, which has a smaller likelihood over the entire sample.

Given the similarities in the posterior distributions, it might be tempting to estimate the constrained linear model, either on the entire sample or a subsample of the data that does not include the ZLB period, and then solve and simulate a nonlinear version of the model conditional on the mean parameter estimates [e.g., Christiano et al. (2015); Cuba-Borda (2014)]. While that approach is much less computationally expensive and will likely provide a decent approximation in a small-scale New Keynesian model, larger differences between the posterior means will arise in more complicated models. For example, Gust et al. (2016) compare the posterior distributions from a constrained linear and nonlinear version of a medium-scale model with capital. Consistent with our results, they find many of the parameter estimates are similar across the two models, but there are large differences in the persistence and SD of the marginal efficiency of investment as well as important differences in the SDs of the other shocks and monetary policy parameters. Also, as we show below, larger differences between the posterior means arise in a model with a banking sector.

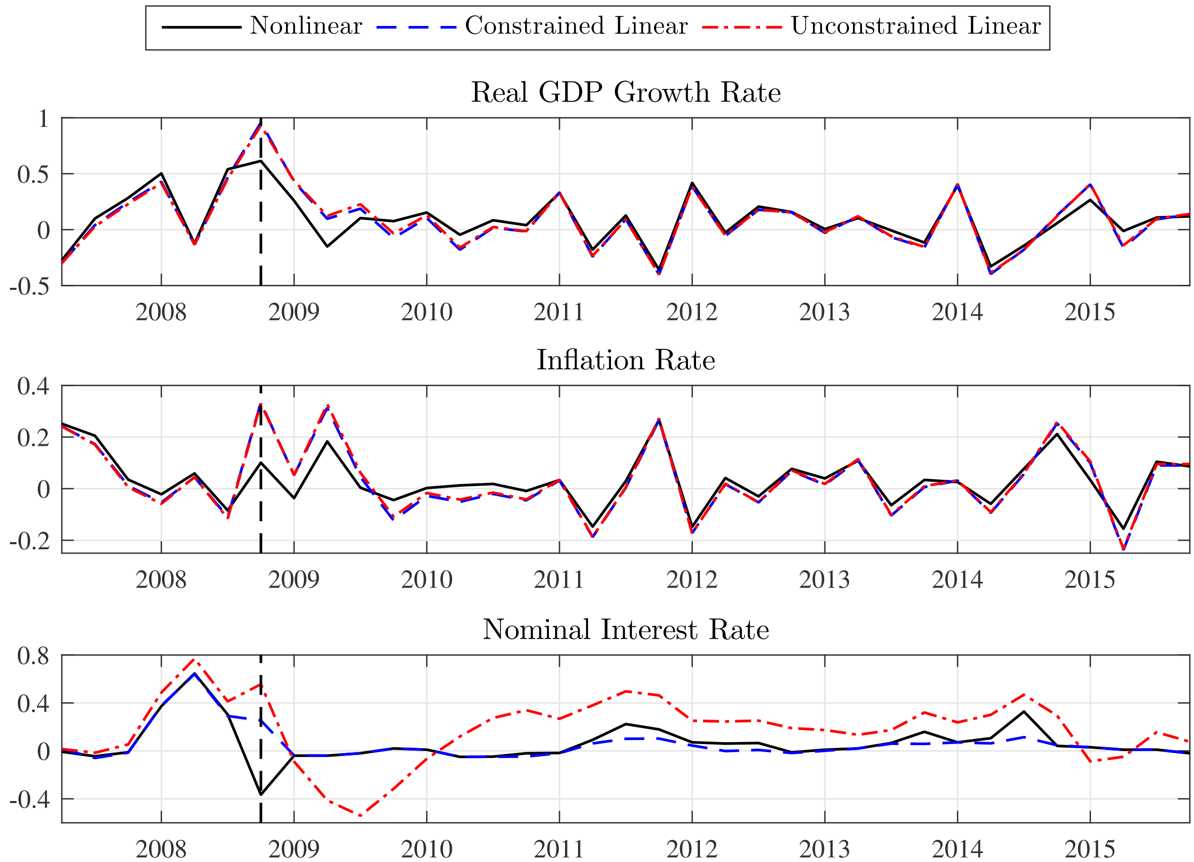

Figure 1: Median filtered observables minus the data in annual percentage points. The vertical dashed line is 2008Q4.

4.2 FILTERED OBSERVABLES AND SHOCKS

Despite similar posterior distributions, the particle filter shows the nonlinear and linear models fit the data differently and lead to competing explanations of how the economy arrived and stayed at the ZLB. Figure 1 shows the level differences between the median filtered observables and the data. The difference is shown in annual percentage points, so 0.1 means the filtered observable is 0.1 percentage points higher than the data. As a reference, we show the filtered observables and shocks from the nonlinear model in the appendix (figures 10 and 11). The vertical line in 2008Q4 indicates when the federal funds rate first fell below 0.25%. In that quarter, the nonlinear model provides a closer fit to the data on real GDP growth and inflation than the linear models. Also, both linear models over-predict the policy rate, while the nonlinear model under-predicts the rate. The nonlinear model continues to better match output and inflation until 2011Q1, although the differences diminish over time. After 2011Q1, the paths implied by the nonlinear and constrained linear models are similar, but the path of the nominal rate from the unconstrained linear model is often higher than in the other models.

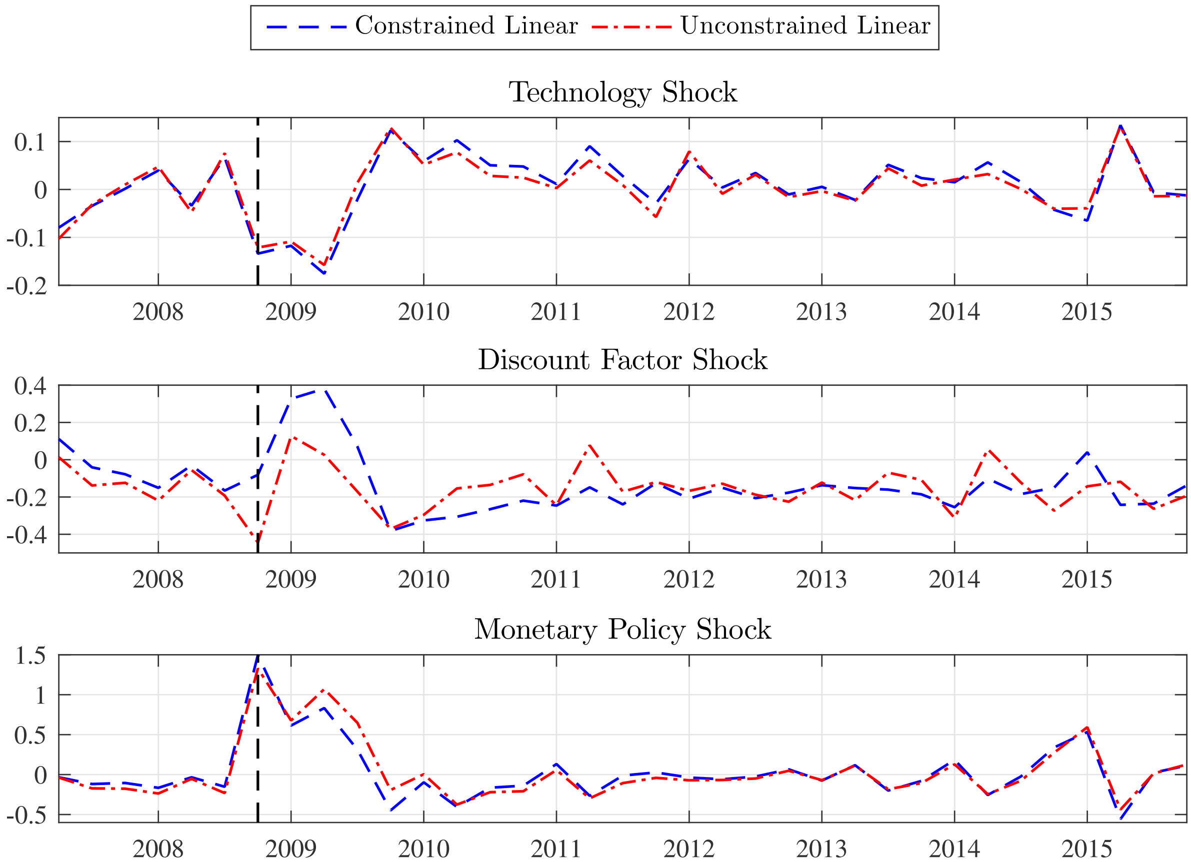

Figure 2: Median filtered shocks in standard deviations from the nonlinear model. The vertical dashed line is 2008Q4.

The models differ most in their predictions of the shocks during the ZLB period. Figure 2 plots the median filtered shocks from the linear models minus the shocks from the nonlinear model. The differences are shown relative to the posterior mean estimate of the shock SD in the nonlinear model, so 2 means a linear model predicted a shock 2SDs larger than in the nonlinear model. In 2008Q4, the differences are stark. Although the nonlinear and constrained linear models predict similar discount factor shocks, the constrained linear model predicts that a larger negative technology shock (-0.1SD) and a much larger monetary policy shock ($+1.5$SD) are necessary to explain the data in 2008Q4. The reason for staying at the ZLB is also quite different. In 2009Q2, the constrained linear model predicts a larger discount factor ($+0.4$SD) and monetary policy shock ($+0.8$SD) than the nonlinear model. The unconstrained linear model predicts implausibly large policy shocks from 2008Q4 to 2009Q4, but the discount factor shocks are also slightly different.

The differences in the filtered paths stem from the negative notional rate the nonlinear model predicts from 2008Q4 to 2011Q1 (figure 10). During that period, the Fed could not lower its policy rate even though it would have preferred to set the rate to -5.1% (-1.3% quarterly) in 2009Q3 because of the large contraction in real GDP growth and the negative inflation rate. In the linear models, households believe the central bank will set a negative policy rate in a severe recession regardless of whether the constraint is imposed in the particle filter, which means the shocks produce different dynamics than in the nonlinear model. After 2011Q1, the nonlinear model predicts a near-zero notional rate, so the filtered paths are similar across the two models.

Gust et al. (2016) estimate a similar path for the notional rate, despite the differences between our nonlinear models. The key takeaway from the notional rate is that the Fed was only constrained for two years, even though the federal funds rate remained below 0.25% until the end of our sample. If the Fed was constrained for a longer period, the differences in the shocks would have been larger.

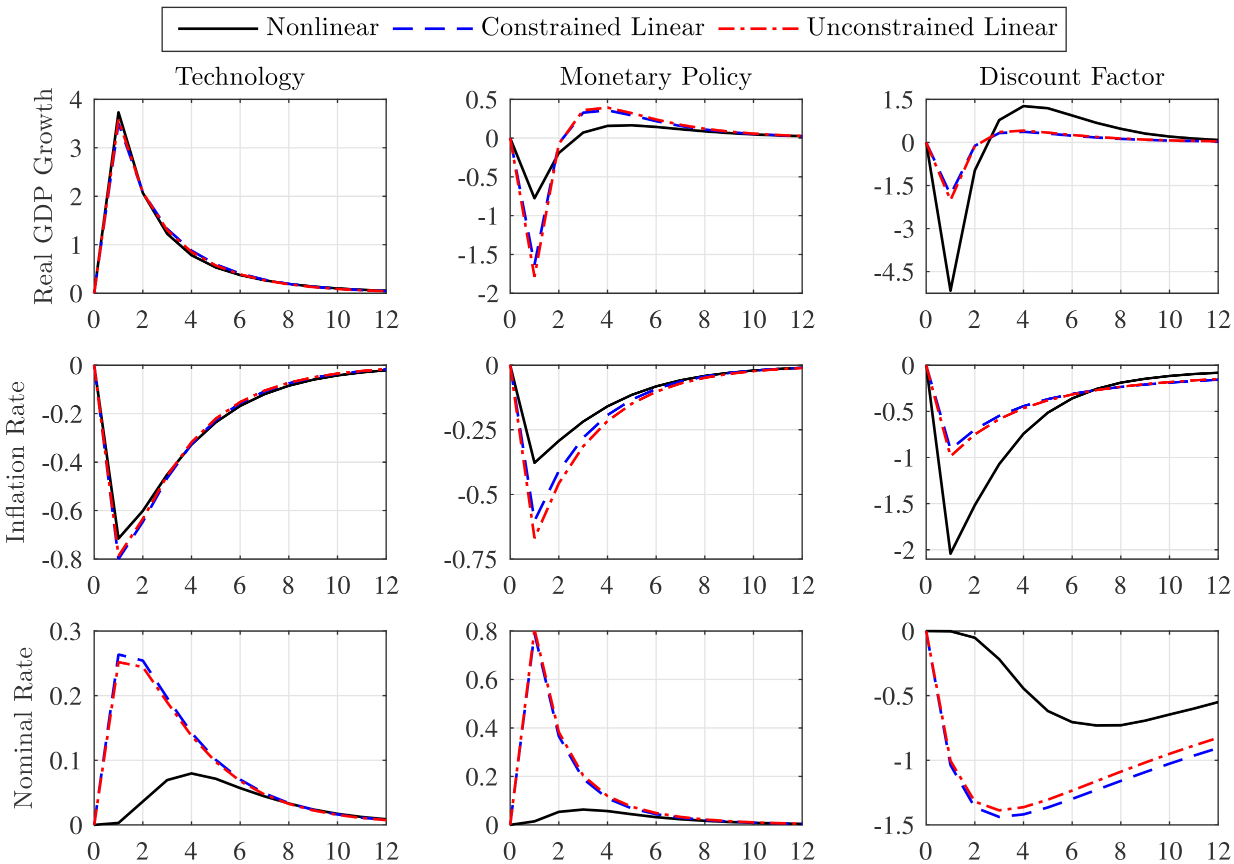

Figure 3: Generalized impulse responses to a 2 standard deviation positive shock in each model. The simulations are initialized at the filtered state corresponding to 2008Q4 using the posterior mean of the nonlinear model. The vertical axis is the annualized difference in the rate from the baseline simulation. The horizontal axis displays the time period.

4.3 GENERALIZED IMPULSE RESPONSES

Generalized impulse responses (GIRFs) help us understand why our models produce different results when the ZLB binds. To compute a GIRF, we follow the procedure in Koop et al. (1996). We first calculate the mean of $10{,}000$ simulations of a given model, conditional on random shocks in every quarter. We then calculate a second mean from another set of $10{,}000$ simulations, but this time the shock in the first quarter is replaced with the shock of interest. The GIRF reports the difference between the two mean paths in each model.

Figure 3 plots the responses to a 2SD positive technology (first column), monetary policy (middle column), and discount factor (last column) shock in each model. To compare the responses, we initialize each simulation at the filtered state corresponding to 2008Q4 in the nonlinear model.

A positive technology shock decreases the marginal cost of production, which generates a tradeoff between output and inflation just like a typical supply shock. The competing effects on real GDP growth and inflation cause technology shocks to have a relatively small impact on the notional rate and whether the ZLB binds compared to discount factor shocks. Therefore, the responses are similar across our models. A positive monetary policy shock, however, directly affects the notional rate. In the linear models, the nominal and notional rates are equal so the shock affects households' decisions even when the ZLB binds. In the nonlinear model the shock only affects the economy if it is large enough to push the nominal interest rate above its ZLB. Since the economy begins in a deep recession and several simulations never exit the ZLB, the monetary policy shock has a much smaller effect on real GDP growth and inflation in the nonlinear model than in the linear models.

A large discount factor indicates that households have a strong desire to save. Elevated savings depresses demand, which reduces output, inflation, and the notional interest rate. In the nonlinear model, any further reduction in expected inflation is offset by an equal increase in the real interest rate since the nominal interest rate is constrained. The higher real rate raises the cost of current consumption, which further lowers demand relative to the linear model such that the responses of real GDP growth and inflation are more than twice as large. Both linear models must compensate for the damped responses to discount factor shocks at the ZLB, which explains why the linear models need a larger negative technology shock and a larger positive policy shock to explain the data in 2008Q4 and both larger discount factor and larger policy shocks to explain the data in 2009.

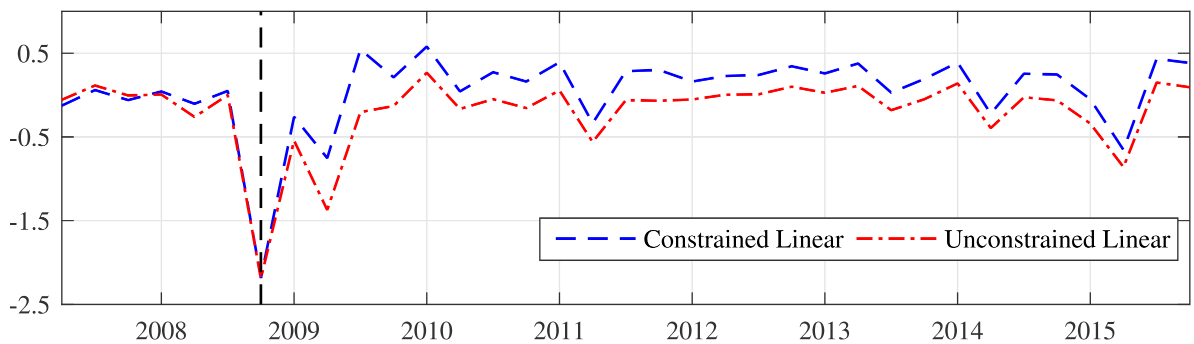

Figure 4: Median filter density in each linear model minus the density in the nonlinear model.

4.4 FILTER DENSITIES The predicted notional rate and the volatility of real GDP growth and inflation during the ZLB period affects each model's ability to fit the data. One measure of the fit is shown in figure 4, which plots a time series of the filter density in both linear models minus the density in the nonlinear model. The linear model's fit is poorest when the notional rate is negative and declining since it endogenously generates much less volatility. For example, the differences in 2008Q4 imply a Bayes factor of about 8. Eventually, both linear models fit the data marginally better than the nonlinear model once the notional rate begins to rise in 2009Q3, which is likely due to the negative notional rate creating persistent volatility in the nonlinear model well after the effects of the crisis subsided. The higher likelihood coming from the linear models, however, is not significant enough to generate a noticeable difference in how well the observables fit the data.

4.5 POSTERIOR PREDICTIVE ANALYSIS Posterior predictive analysis provides another way to identify the strengths and weaknesses of each model. We first compute moments of interest in the data in two subsamples: the pre-ZLB period (1986Q1-2008Q3) and the ZLB period (2008Q4-2011Q1), which we define as the period when the notional rate was negative and the central bank was constrained. We then compute their posterior predictive distributions in each model following the methods in Geweke (2005) and Faust and Gupta (2012). For a given model, we conduct 10,000 simulations for each draw from the posterior distribution. We initialize the simulations in the pre-ZLB period with a state vector drawn from the model's ergodic distribution. Each simulation has the same length as the data, and we condition on periods when the ZLB does not bind. To compute the distributions in the ZLB sample, we initialize each simulation at the filtered state corresponding to 2008Q4. We then condition on simulations with a minimum ZLB event of 8 quarters. We chose that value because it produces a median ZLB duration of 10 quarters, which is the same number of quarters in our ZLB sample. Given the simulated paths, we calculate time averages of the statistics of interest in each sample and then compute the means and quantiles across the simulations, so the distributions account for the uncertainty surrounding both the shocks and the parameter estimates.

	Mean		SD		Skewness	
Real GDP Growth	Pre-ZLB	ZLB	Pre-ZLB	ZLB	Pre-ZLB	ZLB
Data	1.75	−0.75	2.21	3.96	−0.41	−1.35
Nonlinear	1.56 (0.70, 2.43)	0.02 (−2.10, 2.04)	2.44 (1.99, 2.98)	4.60 (3.25, 6.43)	0.08 (−0.32, 0.49)	−0.71 (−1.77, 0.29)
Constrained Linear	1.50 (0.62, 2.39)	1.02 (−1.08, 3.04)	2.54 (2.09, 3.07)	2.55 (1.54, 3.83)	0.04 (−0.36, 0.45)	−0.19 (−1.43, 0.87)
Unconstrained Linear	1.51 (0.65, 2.36)	1.07 (−1.01, 3.12)	2.54 (2.10, 3.05)	2.45 (1.48, 3.62)	0.04 (−0.35, 0.44)	−0.10 (−1.17, 0.96)
Inflation Rate	Pre-ZLB	ZLB	Pre-ZLB	ZLB	Pre-ZLB	ZLB
Data	2.43	1.09	0.91	0.90	0.56	−0.96
Nonlinear	2.49 (2.02, 2.96)	0.96 (−0.24, 2.02)	0.94 (0.72, 1.20)	1.14 (0.57, 2.05)	0.00 (−0.49, 0.50)	−0.08 (−1.13, 0.99)
Constrained Linear	2.54 (1.95, 3.18)	1.21 (0.34, 2.11)	1.03 (0.79, 1.33)	0.80 (0.43, 1.24)	0.02 (−0.47, 0.52)	0.09 (−0.91, 1.10)
Unconstrained Linear	2.44 (1.88, 3.03)	1.11 (0.21, 2.04)	1.02 (0.78, 1.31)	0.79 (0.42, 1.25)	0.02 (−0.47, 0.52)	0.09 (−0.93, 1.12)

Table 3: Comparison between various moments of interest in the data and their posterior predictive distributions in each model. The values shown in parentheses are $(5\%, 95\%)$ credible sets. All of the values are annualized net rates.

Table 3 shows the median and 90% credible sets for the mean, SD, and skewness of real GDP growth and inflation in the pre-ZLB and ZLB periods of each model. Unsurprisingly, the three models produce similar distributions in the pre-ZLB period. The distributions of the means and SDs are wide but the median values are near the values in the data. The skewness in the pre-ZLB period, however, is unlikely to occur in our models. In the data, real GDP growth is negatively skewed and inflation is positively skewed, but the median skewness is near zero in the three models.

In the ZLB period, mean real GDP growth and inflation are lower, the SD of real GDP growth is higher, the SD of inflation is about the same, and both real GDP growth and inflation are negatively skewed. The nonlinear model better matches those features than the linear models in several ways. First, average real GDP growth is much lower than its value in the pre-ZLB period, whereas it is only slightly lower in the linear models. Second, real GDP growth is far more volatile than in the pre-ZLB period. In contrast, neither of the linear models predict much change in volatility. Third, real GDP growth and inflation are negatively skewed, while the linear models generate very little skewness. Overall, the values in the data are close to the median estimates from our nonlinear model, whereas the data is often in the tails or outside of the 90% credible sets in the linear models.

Figure 5 plots the posterior predictive distribution of ZLB event durations. The unconditional durations (top panel) are similar across the three models. The most frequent ZLB event is only one

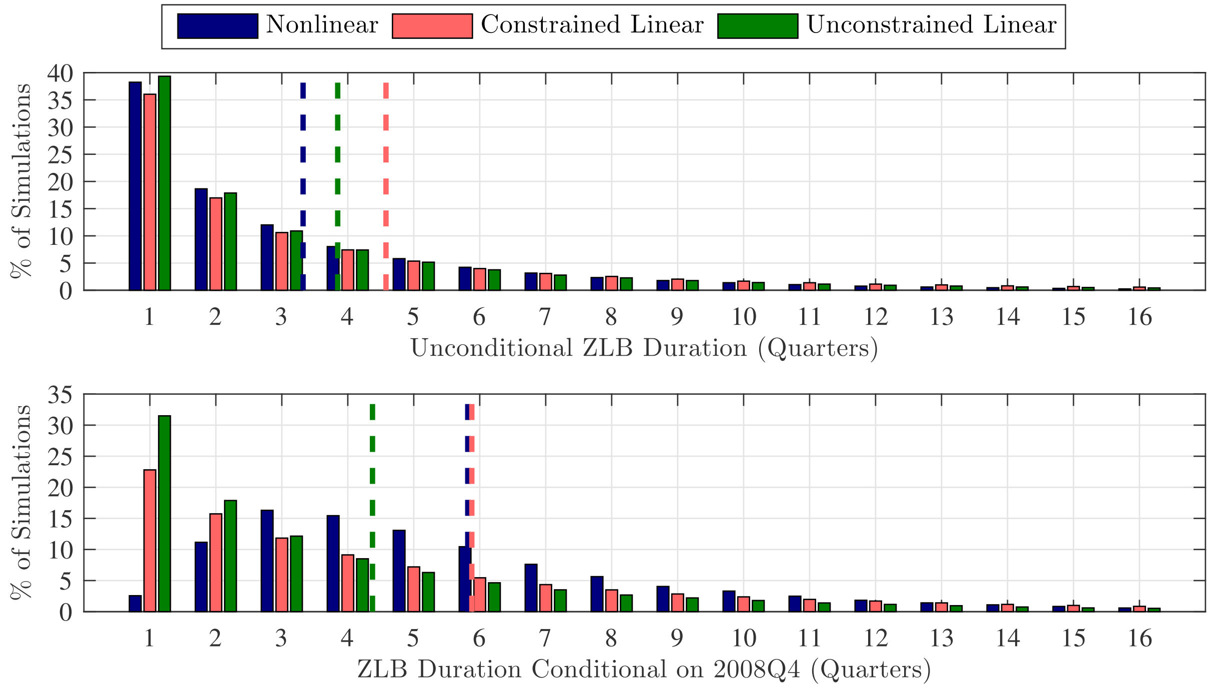

Figure 5: Distribution of ZLB events in each model. The vertical dashed line represents the expected ZLB duration.

quarter and the average duration ranges from 3.2 quarters in the nonlinear model to 4.6 quarters in the constrained linear model. The higher mean and slightly higher likelihood of a lengthy ZLB event in the linear model is mostly due to the higher estimated persistence of the discount factor. The durations conditional on the filtered state in 2008Q4 (bottom panel), however, are significantly different. In both linear models, the most frequent ZLB event remains one quarter, whereas the most common ZLB event duration in the nonlinear model is between 3 and 4 quarters. Also, the average duration in the nonlinear model increases to 5.9 quarters and lengthy ZLB events become more common than in the linear models, despite the differences in the discount factor persistence.

The durations implied by the nonlinear model are far more consistent with survey data. For example, in the January 2009 Blue Chip survey of financial forecasters, 59% of the forecasters predicted the federal funds rate would exceed 0.25% after 3 or 4 quarters (26% predicted 3 quarters and 33% predicted 4), but only 9% predicted that it would exceed 0.25% in 2 or fewer quarters. In the nonlinear model, 41.2% of ZLB events are six or more quarters, whereas 26% of survey respondents thought the federal funds rate would remain below 0.25% for six or more quarters. Evidently, most forecasters expected the Fed to maintain its zero interest rate policy for longer than a couple quarters but not an extended period of time, despite the severity of the Great Recession.

Another important determinant of the dynamics in each model is the frequency of ZLB events. A lower (higher) frequency means households place a smaller (larger) weight on the ZLB in expectation, so the more volatile dynamics at the ZLB have a smaller (larger) impact on households' decisions at and away from the ZLB. Unfortunately, there is no reliable measure of this statistic in the data. The data contains one ZLB event that lasted for 27, or 23%, of the 120 quarters in our sample. With a longer sample, the frequency would decline in the data, but the model-implied distributions would likely remain unchanged. Survey data is unreliable because of how quickly the Fed cut its policy rate. Therefore, there is no clear way to determine the frequency of ZLB events.

5 NONLINEAR VS. QUASI-LINEAR MODEL

We focus on how the nonlinear model compares with the linear models because the nonlinear provides the most accurate predictions of the dynamics at the ZLB and linear models are commonly used for policy analysis. One alternative is to solve a quasi-linear version of the nonlinear model using the OccBin toolbox developed by Guerrieri and Iacoviello (2015). To introduce the constraint, they break the problem into two regimes, one where the constraint binds and one where it is slack, and then use backward induction from the last period that the ZLB binds to solve the model. Their key assumption is that households do not account for the possibility that the constraint may bind in the future when it does not bind in the current period. That means households' decisions away from the ZLB are unaffected by their decisions at the ZLB. Two potentially important implications stem from this assumption. First, households ignore the effects of the constraint in states of the economy where the ZLB is likely to bind in the near future. Second, the ZLB has a smaller impact on households' decisions when it binds because it has a lower frequency. In the nonlinear model, households form expectations over the entire distribution of shocks at and away from the ZLB. The question is whether the differences in the solutions are empirically significant.

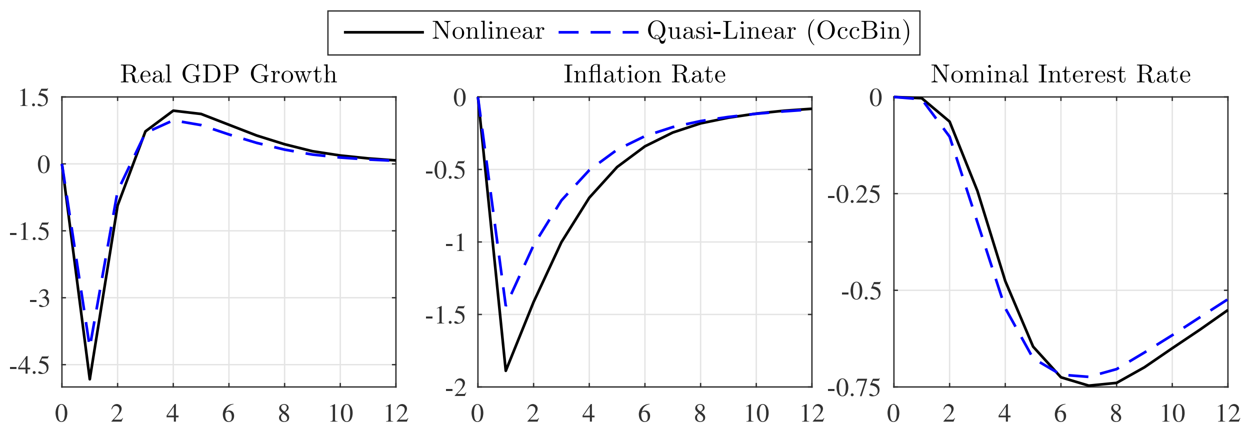

Figure 6: Generalized impulse responses to a 2 standard deviation positive discount factor shock. The simulations are initialized at the filtered state corresponding to 2008Q4 using the posterior mean of the nonlinear model. The vertical axis is the annualized difference in the rate from the baseline simulation. The horizontal axis displays the time period.

Figure 6 shows generalized impulse responses to a discount factor shock in the nonlinear model (solid line) and the quasi-linear model (dashed line). Those responses are initialized at the median filtered state corresponding to 2008Q4 from the nonlinear model. The quasi-linear model does a better job matching the dynamics of the nonlinear model than the linear model. For example, the responses to a technology shock are similar and the differences in the responses to both a monetary policy and discount factor shock are much smaller. However, there is still a key difference—the quasi-linear model does not produce as much volatility when the ZLB binds as the nonlinear model. In response to a 2SD positive discount factor shock in the quasi-linear model, real GDP growth declines by about 4.1% and inflation falls by roughly 1.4%. In the nonlinear model, those rates fall 17% and 31% more, even though we use the same parameters. Similar differences occur when we initialize at the filtered states in 2009 or 2010. The differences in the responses mean that the quasi-linear model must explain the data with larger and different shocks than the nonlinear model.

Figure 7 shows the median filtered shocks in the quasi-linear model minus the shocks in the nonlinear model. The data is filtered conditional on the posterior mean parameters from the nonlin-

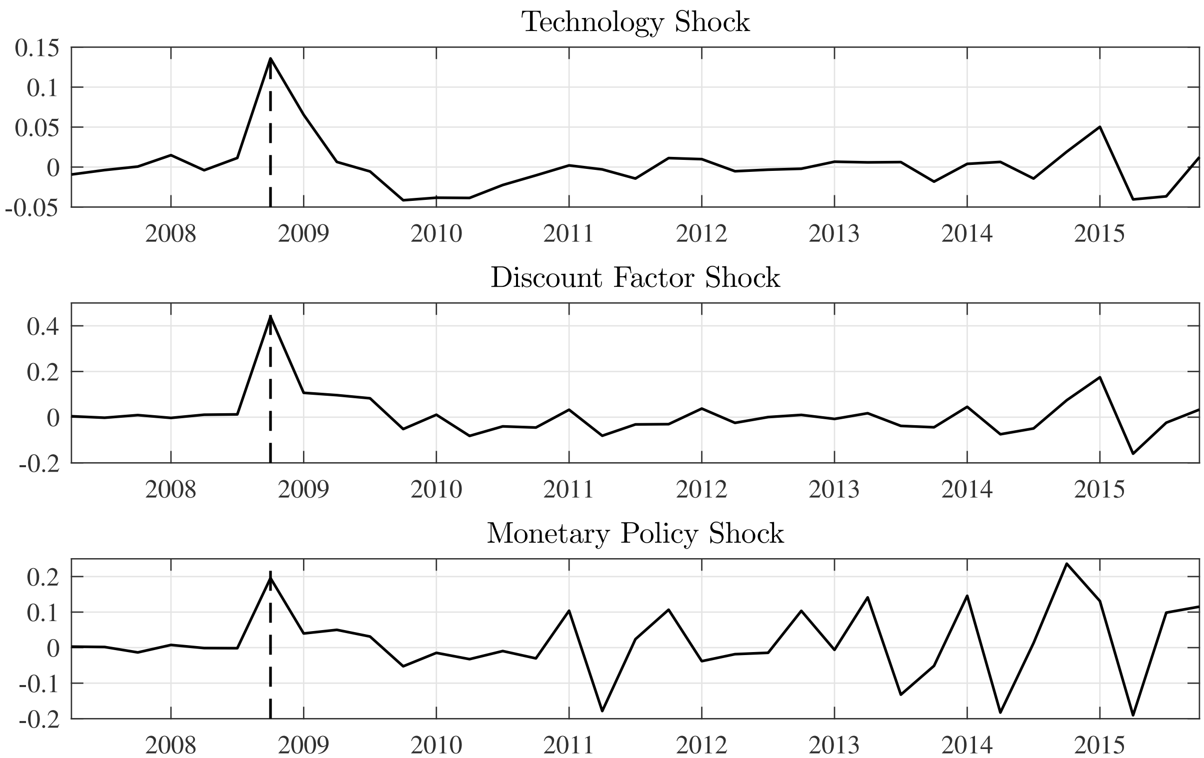

Figure 7: Median filtered shocks in standard deviations from the nonlinear model. The vertical dashed line is 2008Q4.

ear model and the shocks are normalized by their posterior mean SDs. In 2008Q4, the quasi-linear model still requires a larger discount factor ($+0.4$SD) and monetary policy ($+0.2$SD) shock to explain the data, and throughout 2009, it predicts larger discount factor shocks just like the linear models. However, these predictions are much closer to the nonlinear model than the predictions from the linear models because the quasi-linear model endogenously generates far more volatility.

The advantage of using the quasi-linear model over the nonlinear model is that it is quicker to solve. For example, with 64 processors it takes 1-2 seconds to solve the nonlinear model with our policy function iteration algorithm, whereas it takes a fraction of a second to solve the quasi-linear model with OccBin. Furthermore, the nonlinear solution time increases exponentially with the size of the model, whereas the size of the model has little effect on the solution time in the quasi-linear model. Unfortunately, the quasi-linear model is costly to filter, because OccBin requires a long enough simulation to return to the regime where the ZLB does not bind in order to filter each period in the sample. For example, it takes about 3-4 seconds to run the particle filter given the nonlinear solution and at least 30 seconds to filter the quasi-linear model. Therefore, a simulation-based filter, such as the particle filter, is a very costly way to estimate the quasi-linear model since a large number of simulations are necessary to get a good approximation of the model's likelihood.

To speed up the filter, Guerrieri and Iacoviello (2016) follow Fair and Taylor (1983) and use a nonlinear filter that does not depend on a large number of simulations. Each period the filter solves for the shocks that minimize the distance between the observables and the model predictions. To test the performance of that filter, we use the unconstrained linear model because we can use the estimates from the Kalman filter as a benchmark. Table 4 shows the posterior means and credible sets based on the particle filter, the Kalman filter with and without ME, and the Fair and Taylor

	Linear Unconstrained Model Posterior Mean $(5\%, 95\%)$			
Parameter	Particle	Kalman ME	Kalman No ME	Fair & Taylor
φ	89.75551 $(60.0572, 124.9987)$	88.32964 $(58.9214, 122.5572)$	79.63849 $(52.7309, 110.7193)$	115.27029 $(81.4358, 157.3756)$
h	0.43851 $(0.29796, 0.56827)$	0.43056 $(0.29310, 0.56472)$	0.43644 $(0.31774, 0.55465)$	0.54844 $(0.43274, 0.65539)$
ϕ_π	3.74194 $(3.02972, 4.53073)$	3.73074 $(3.01440, 4.54203)$	3.75178 $(3.07512, 4.52112)$	1.99119 $(1.79762, 2.21518)$
ϕ_c	1.24805 $(0.90772, 1.63401)$	1.25149 $(0.88588, 1.64862)$	1.35632 $(1.00980, 1.73330)$	0.73057 $(0.51654, 0.94298)$
ρ_i	0.83983 $(0.78855, 0.88403)$	0.84123 $(0.79044, 0.88300)$	0.89440 $(0.87116, 0.91642)$	0.85140 $(0.82349, 0.87751)$
ρ_g	0.19183 $(0.05145, 0.37168)$	0.19402 $(0.05692, 0.38591)$	0.12382 $(0.03770, 0.25779)$	0.22251 $(0.06236, 0.43357)$
ρ_β	0.92928 $(0.88839, 0.96493)$	0.93240 $(0.89520, 0.96652)$	0.94715 $(0.91585, 0.97348)$	0.97460 $(0.96040, 0.98621)$
σ_ε	0.00938 $(0.00729, 0.01189)$	0.00929 $(0.00718, 0.01178)$	0.01122 $(0.00911, 0.01359)$	0.01356 $(0.01025, 0.01710)$
σ_υ	0.00191 $(0.00143, 0.00257)$	0.00186 $(0.00140, 0.00244)$	0.00160 $(0.00124, 0.00206)$	0.00109 $(0.00087, 0.00134)$
σ_ν	0.00177 $(0.00133, 0.00227)$	0.00174 $(0.00133, 0.00226)$	0.00123 $(0.00106, 0.00141)$	0.00125 $(0.00108, 0.00144)$

Table 4: Prior distributions, means, standard deviations, and credible sets of the estimated parameters.

filter. Most researchers specify zero ME when using the Kalman filter, but that makes it difficult to assess the performance of the particle filer, which requires ME to avoid degeneracy—a situation when all but a few particles receive near zero weight. When the particle filter and Kalman filter have the same ME variance, the posterior means and credible sets are similar (columns 1 and 2). When there is no ME, the Kalman filter produces different posterior means (column 3), but in all but one case the credible sets include the mean parameter estimates from the particle filter. Moreover, those differences would decline if we estimated the model with more particles and less ME. When the Fair and Taylor filter is used, the differences are significant (column 4). The posterior means are outside the credible sets from the Kalman filter for most parameters, which means the Fair and Taylor filter is too inaccurate to estimate the quasi-linear model using OccBin.

6 EXTENDED MODEL

This section extends our model to show how the differences between the nonlinear and linear models change with additional state variables and shocks. The model is based on Cúrdia and Woodford (2010). It includes two types of households, lenders and borrowers, that differ in their preferences and a perfectly competitive commercial bank that intermediates loans between the two households. The production sector and the monetary policy rule are identical to the baseline model.

6.1 HOUSEHOLDS There is a continuum of households, of which a fraction μ are lending households and a fraction $1-\mu$ are borrowing households. Both households have the same preferences as the baseline model, except borrowing households have a lower discount factor (i.e., $\beta_{bh,t} = \bar{\beta}_{bh} - \bar{\beta}_{lh} + \beta_{lh,t}$) and we removed habit formation to keep the model numerically tractable.

Lending households (lh), choose $\{c_{lh,t}, n_{lh,t}, d_t\}_{t=0}^{\infty}$ to maximize expected utility subject to

$$c_{lh,t} + d_t = w_t n_{lh,t} + i_{t-1}^d d_{t-1}/\pi_t + (div_t^b + div_t^f)/\mu,$$

where d is the real value of a nominal bank deposit that pays interest i^d and div^b and div^f are real dividends from banks and intermediate firms that are equally distributed across the μ households.

Borrowing households (bh) choose $\{c_{bh,t}, n_{bh,t}, \ell_t\}_{t=0}^{\infty}$ to maximize expected utility subject to

$$c_{bh,t} + i_{t-1}^{\ell}\ell_{t-1}/\pi_t = w_t n_{bh,t} + \ell_t + \psi_t L_t/(1-\mu),$$

where ℓ is the real value of a nominal bank loan issued at rate i^{ℓ} and ψL is the real value of unpaid loans, which is treated as a lump-sum transfer and equally distributed across the $1-\mu$ households.

6.2 BANKS The banking sector is perfectly competitive, but banks face a real and financial cost to issuing loans following Cúrdia and Woodford (2010). The real cost is $\Gamma(L) = \gamma_1 z(L/z)^{\gamma_0}$, which grows at the rate of technology and is a convex function of detrended loans. The financial cost, ψL, is due to the real value of unpaid loans, which is a proxy for credit risk. Banks know some loans will not be repaid and account for them in their decisions, so real bank dividends equal

$$div_t^b = (i_{t-1}^{\ell}L_{t-1} - i_{t-1}^{d}D_{t-1})/\pi_t + (D_t - L_t) - (\Gamma(L_t) + \psi_t L_t).$$

Each bank chooses its loans, L_t, to maximize the expected discounted value of future dividends, $E_0\sum_{t=0}^{\infty}\tilde{\beta}_{lh}(c_{lh,0}/c_{lh,t})div_t^b$, subject to its balance sheet, $D_t = L_t$. The optimality condition implies

$$i_t^{\ell} = i_t^d(1+s_t), \tag{9}$$

where $s \equiv \gamma_0\gamma_1(L/z)^{\gamma_0-1} + \psi$ is the interest rate spread. The fraction of unpaid loans, ψ, follows

$$\ln(\psi_t) = (1-\rho_\psi)\ln(\bar{\psi}) + \rho_\psi\ln(\psi_{t-1}) + \sigma_\upsilon\upsilon_t, \tag{10}$$

where $\bar{\psi}$ is the fraction of unpaid loans along the steady state growth path, $0 \le \rho_\psi < 1$ is the persistence of the fraction of unpaid loans, and υ is a standard normal shock. Shocks to ψ directly affect the interest rate spread, so they can account for exogenous factors that influenced the spread during the financial crisis (e.g., the collapse of Lehman Brothers) and the various forms of unconventional monetary policy that were used to put downward pressure on the spread (e.g., quantitative easing).

6.3 COMPETITIVE EQUILIBRIUM Total deposits and loans are $D_t = \mu d_t$ and $L_t = (1-\mu)\ell_t$. Aggregate consumption and labor are $c_t = \mu c_{lh,t} + (1-\mu)c_{bh,t}$ and $n_t = \mu n_{lh,t} + (1-\mu)n_{bh,t}$. The aggregate resource constraint is given by $c_t = y_t - adj_t - \Gamma(L_t)$. The detrended equilibrium system includes the ZLB constraint, the bank's optimality condition, the stochastic processes, and

$$\tilde{w}_t = \chi_j n_{j,t}^{\eta}\tilde{c}_{j,t}, \quad j \in \{lh, bh\}, \tag{11}$$

$$1 = i_t^k E_t[\beta_{j,t+1}(\tilde{c}_{j,t}/\tilde{c}_{j,t+1})(1/(g_{t+1}\bar{\pi}\hat{\pi}_{t+1}))], \quad (j,k) \in \{(lh,d),(bh,\ell)\}, \tag{12}$$

$$\tilde{y}_t = \mu n_{lh,t} + (1-\mu)n_{bh,t}, \tag{13}$$

$$\varphi(\hat{\pi}_t - 1)\hat{\pi}_t = (1-\theta) + \theta\tilde{w}_t + \varphi E_t[\beta_{lh,t+1}(\tilde{c}_{lh,t}/\tilde{c}_{lh,t+1})(\hat{\pi}_{t+1}-1)\hat{\pi}_{t+1}(\tilde{y}_{t+1}/\tilde{y}_t)], \tag{14}$$

$$\tilde{c}_{bh,t} + i_{t-1}^{\ell}\tilde{\ell}_{t-1}/(g_t\pi_t) = \tilde{w}_t n_{bh,t} + (1+\psi_t)\tilde{\ell}_t, \tag{15}$$

$$\tilde{c}_t = \mu\tilde{c}_{lh,t} + (1-\mu)\tilde{c}_{bh,t}, \tag{16}$$

$$\tilde{c}_t = [1 - \varphi(\pi_t/\bar{\pi} - 1)^2/2]\tilde{y}_t - \gamma_1[(1-\mu)\tilde{\ell}_t]^{\gamma_0}, \tag{17}$$

$$i_t^{dn} = (i_{t-1}^{dn})^{\rho_{id}}(\bar{i}^d(\pi_t/\bar{\pi})^{\phi_\pi}(g_t\tilde{c}_t/(\bar{g}\tilde{c}_{t-1}))^{\phi_c})^{1-\rho_{id}}\exp(\nu_t). \tag{18}$$

A competitive equilibrium includes sequences of quantities, $\{\tilde{c}_t, \tilde{c}_{lh,t}, \tilde{c}_{bh,t}, n_{lh,t}, n_{bh,t}, \tilde{\ell}_t, \tilde{y}_t\}_{t=0}^{\infty}$, prices, $\{w_t, i_t^d, i_t^{dn}, i_t^{\ell}, \hat{\pi}_t\}_{t=0}^{\infty}$, and exogenous variables, $\{\beta_{lh,t}, \beta_{bh,t}, g_t, \psi_t\}_{t=0}^{\infty}$, that satisfy the detrended system, given the state, $\{c_{-1}, i_{-1}^{dn}, i_{-1}\tilde{\ell}_{-1}, \beta_0, g_0, \nu_0, \psi_0\}$, and the shocks, $\{\varepsilon_t, \upsilon_t, \nu_t, \psi_t\}_{t=1}^{\infty}$.

6.4 ESTIMATION We estimate the extended model using the same procedure as the baseline model, except we add an observable: the spread between the Baa corporate bond yield and the 10-year treasury yield. In addition to the five calibrated parameters in the baseline model, we assume both types of households supply the same amount of labor in steady state. We also set the steady state loan-to-income ratio to 3.35, which equals the ratio of private (households, nonprofit organizations, and nonfinancial businesses liabilities, excluding mortgages) debt to nominal GDP.

	Prior		Posterior Mean $(5\%, 95\%)$		
Parameter	Dist	Mean (SD)	Nonlinear	Constrained Linear	Unconstrained Linear
φ	Gam	80.000 (20.000)	77.81684 (53.39802, 106.45351)	70.45353 (47.41162, 99.07662)	71.49784 (46.11284, 104.36066)
ϕ_π	Norm	3.000 (0.750)	3.80226 (3.16238, 4.50769)	3.86881 (3.16977, 4.61810)	3.63155 (2.99201, 4.28087)
ϕ_c	Norm	1.000 (0.400)	1.79446 (1.41407, 2.17936)	1.66688 (1.33652, 2.04541)	1.57240 (1.24957, 1.92761)
$\bar{g}$	Norm	1.004 (0.001)	1.00388 (1.00293, 1.00486)	1.00371 (1.00273, 1.00471)	1.00365 (1.00272, 1.00463)
$\bar{\pi}$	Norm	1.006 (0.001)	1.00615 (1.00553, 1.00673)	1.00609 (1.00547, 1.0067)	1.00591 (1.00527, 1.00655)
$\bar{\psi}$	Gam	0.006 (0.002)	0.00430 (0.00293, 0.00550)	0.00427 (0.00273, 0.00582)	0.00430 (0.00264, 0.00584)
$\bar{s}$	Gam	0.008 (0.002)	0.00575 (0.00527, 0.00628)	0.00603 (0.00543, 0.00667)	0.00599 (0.00538, 0.00665)
γ_0	Gam	15.000 (5.000)	13.85454 (7.27579, 21.21671)	16.67012 (9.08941, 23.83093)	16.64823 (9.29994, 24.52005)
μ	Beta	0.500 (0.200)	0.30501 (0.12568, 0.51416)	0.35452 (0.13520, 0.62667)	0.37174 (0.14907, 0.65444)
ρ_β	Beta	0.500 (0.200)	0.90781 (0.87660, 0.93156)	0.94823 (0.91425, 0.97638)	0.94320 (0.90764, 0.97338)
ρ_ψ	Beta	0.500 (0.200)	0.91674 (0.86818, 0.95349)	0.92732 (0.86223, 0.97353)	0.91809 (0.84885, 0.97196)
ρ_g	Beta	0.500 (0.200)	0.21511 (0.06693, 0.42169)	0.21110 (0.06179, 0.40569)	0.19486 (0.05992, 0.38272)
ρ_{id}	Beta	0.500 (0.200)	0.77172 (0.69731, 0.82936)	0.79529 (0.73783, 0.84627)	0.80168 (0.73981, 0.85426)
σ_ε	IGam	0.010 (0.010)	0.00157 (0.00124, 0.00197)	0.00162 (0.00126, 0.00203)	0.00139 (0.00110, 0.00173)
σ_υ	IGam	0.100 (0.100)	0.14209 (0.10900, 0.18598)	0.17936 (0.11941, 0.26234)	0.18259 (0.12172, 0.28526)
σ_ν	IGam	0.010 (0.010)	0.00215 (0.00153, 0.00297)	0.00204 (0.00147, 0.00273)	0.00203 (0.00145, 0.00277)
σ_ξ	IGam	0.010 (0.010)	0.00673 (0.00534, 0.00819)	0.00673 (0.00536, 0.00808)	0.00675 (0.0053, 0.00811)
$\log(ML)$			2263.30	2245.68	2237.84

Table 5: Prior distributions, means, standard deviations, and credible sets of the estimated parameters.

6.5 MODEL COMPARISON Table 5 reports the posterior parameter estimates for the extended model. Once again, most of the estimates are similar across the three models. However, there are two key exceptions. One, the SD of the interest rate spread shock, σ_υ, is much smaller in the nonlinear model than the linear models. The spread is difficult for the model to match because it is heavily skewed. In our sample, the spread never fell below 1.25% and the average was only 2.33% annually, but during the financial crisis it peaked at 5.58%. In the nonlinear model, the spread has a log normal distribution, which has a long right tail and captures the asymmetry in the data much better than the normal distribution in the linear models. Therefore, the nonlinear model does not need as large of shocks to explain the changes in the spread. Second, the persistence of discount factor, ρ_β, is smaller in the nonlinear model than the linear models by a wider margin than in the baseline model. We also find the marginal data density for the nonlinear model is significantly higher than in both linear models, in contrast with the baseline model. The linear models do not endogenously generate enough volatility to match the declines in real GDP growth and inflation while also matching the sharp increase in the spread that occurred at the start of the ZLB period.

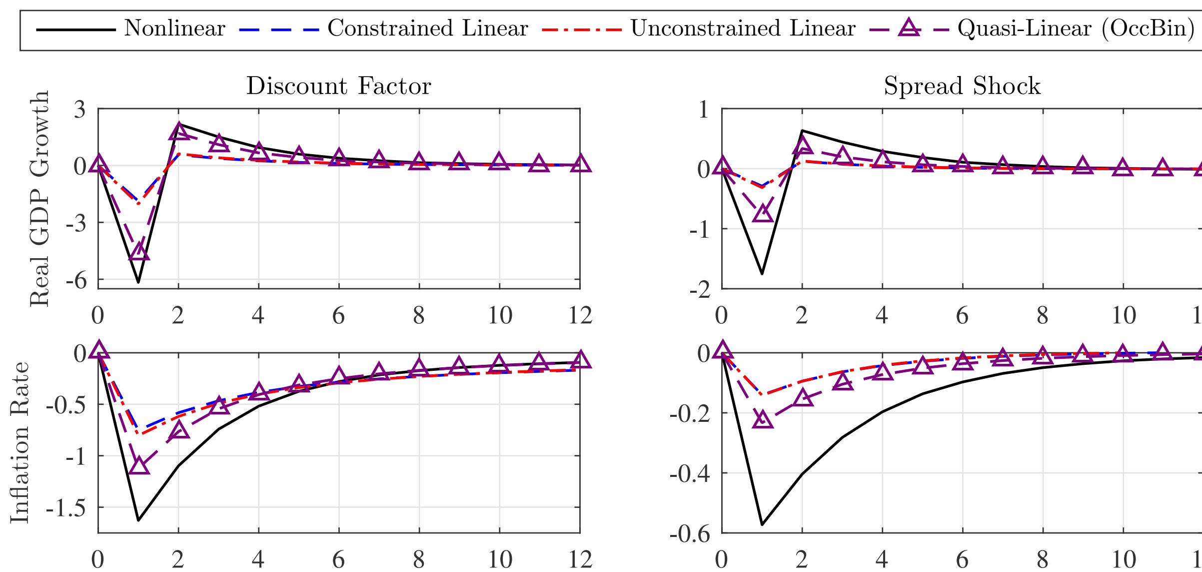

Figure 8: Generalized impulse responses to a 2 standard deviation positive shock in each model. The simulations are initialized at the filtered state corresponding to 2008Q4 using the posterior mean of the nonlinear model. The vertical axis is the annualized difference in the rate from the baseline simulation. The horizontal axis displays the time period.

Figure 8 shows generalized impulse responses to each shock in the nonlinear, constrained linear, unconstrained linear, and quasi-linear model. The responses are initialized at the median filtered state corresponding to 2008Q4 from the nonlinear model. Relative to the baseline model, there are even larger differences between the nonlinear and linear models in response to a discount factor shock. The responses to discount factor shocks in the quasi-linear model are closer to the nonlinear model, but they still generate far less volatility. The differences are equally well pronounced for the spread shocks, which are log-normally distributed in the nonlinear model. Neither the linear or quasi-linear models can account for asymmetric shock distributions. The differences in these responses indicate that the models will predict very different shocks when the ZLB binds.

In the appendix, figures 11 and 13 plot the median filtered shocks from the baseline and extended nonlinear models relative to the posterior mean SD. The baseline model requires a 3.5SD positive shock to the discount factor to explain the data in 2008Q4. The extended model, however, only requires a 2.5SD shock because a large positive spread shock helps explain the sharp contraction in economy activity. Figure 9 compares the shocks predicted by the extended nonlinear model to those from the linear models. The differences are even larger than we reported for the baseline model. For example, in 2008Q4 the constrained linear model predicts a monetary policy shock that is 1.8SDs larger than the extended nonlinear model, whereas in the baseline model the monetary policy shock was only 1.3SDs larger than the nonlinear model. Furthermore, those differences

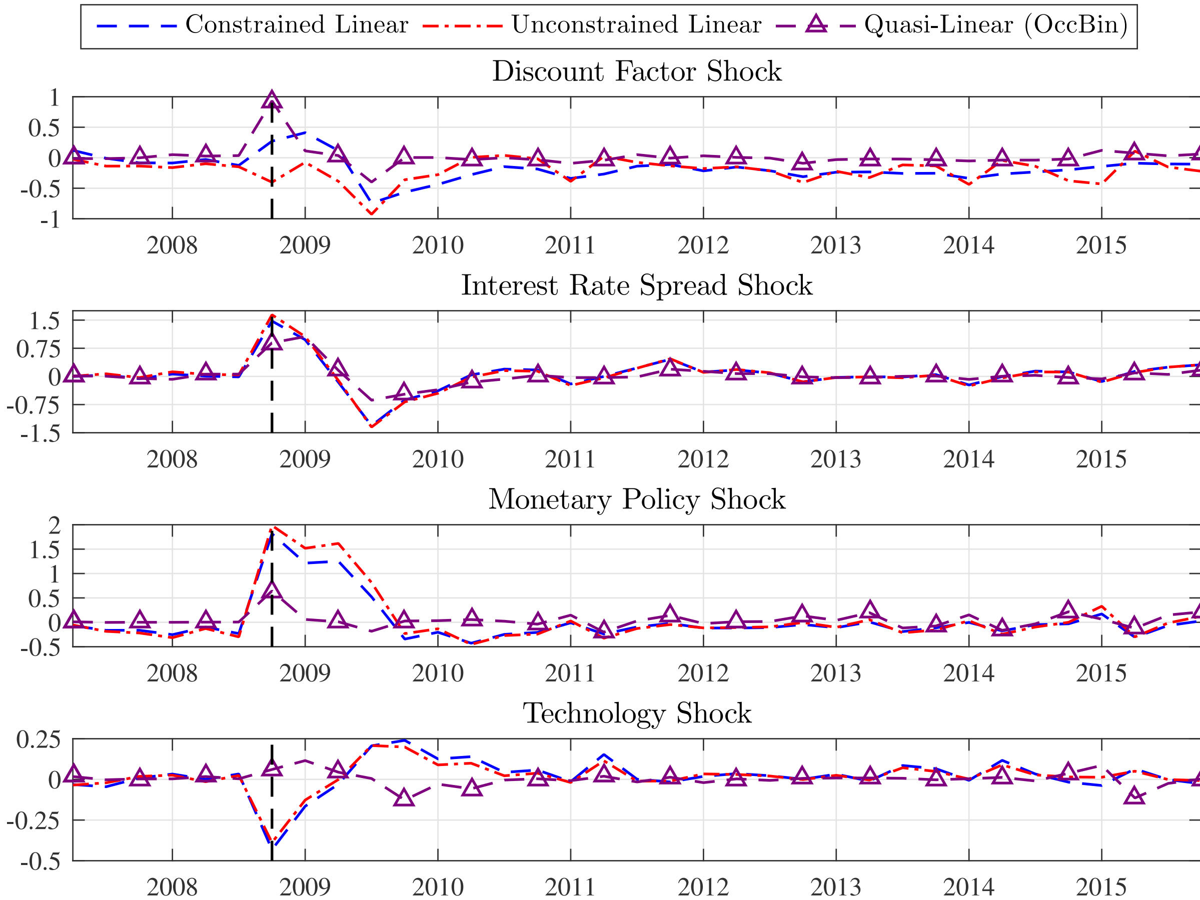

Figure 9: Median filtered shocks in standard deviations from the nonlinear model. The vertical dashed line is 2008Q4.

persist until 2010. In addition to the larger monetary policy shocks, the linear models also predict larger discount factor and spread shocks at the start of the ZLB period. These results indicate that the differences between the linear and nonlinear models increase with the complexity of the model.

Figure 9 also shows the filtered shocks from the quasi-linear model, which are conditional on the mean parameter estimates from the nonlinear model and do not include parameter uncertainty. Those estimates demonstrate that the spread, monetary policy, and technology shocks from the quasi-linear model are closer to the nonlinear model than the linear models. However, the quasi-linear model still predicts larger monetary policy and spread shocks than the nonlinear model when the Fed is constrained and an even larger discount factor shock in 2008Q4 than the linear models. The differences of each shock from the nonlinear model are also larger than in the baseline model.

7 CONCLUSION

This paper analyzes the importance of incorporating the ZLB constraint into households' expectations. To conduct our analysis, we compare the posterior distributions based on three models that differ in their treatment of the constraint: (1) a nonlinear model, which includes an occassionally binding ZLB constraint; (2) a constrained linear model, which imposes the constraint in the filter but not the solution; and (3) an unconstrained linear model, which never imposes the constraint.

We find that our models generate similar posterior distributions when estimated with U.S. data that includes the ZLB period. However, the models provide different explanations for why the U.S. economy arrived and stayed at the ZLB. The nonlinear model primarily attributes the ZLB to a reduction in household demand due to discount factor shocks, while the linear model incorrectly predicts that positive monetary policy shocks are needed to explain the data in the ZLB period. When we extend the model to include additional states, shocks, and observables by adding a banking sector, the differences between the linear and nonlinear models become even more significant.

We also compare the predictions from our nonlinear model to the quasi-linear model using OccBin. The dynamics in the quasi-linear model are closer to those in the nonlinear model than the linear model, but there is still less endogenous volatility when the ZLB binds. As a consequence, the quasi-linear model predicts different shocks than the nonlinear model when the Fed is constrained. The quasi-linear model is also not as conducive to estimation because the particle filter, which relies on repeated model simulations, is too slow and the Fair and Taylor filter is inaccurate.

Our results demonstrate that there are meaningful differences in the predictions from linear, quasi-linear, and nonlinear models when the ZLB binds. It stands to reason that larger differences in the parameter estimates and model predictions would occur in models that include other important nonlinearities, such as irreversible investment, borrowing constraints, and stochastic volatility. Future research is needed to further identify these differences and how much they matter for policy.

REFERENCES

BRAUN, R. A., L. M. KÖRBER, AND Y. WAKI (2012): "Some Unpleasant Properties of Log-Linearized Solutions When the Nominal Rate is Zero," Federal Reserve Bank of Atlanta Working Paper 2012-5a.

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.

CUBA-BORDA, P. (2014): "What Explains the Great Recession and the Slow Recovery?" Manuscript, University of Maryland.

CÚRDIA, V. AND M. WOODFORD (2010): "Credit Spreads and Monetary Policy," *Journal of Money, Credit and Banking*, 42, 3–35.

DEL NEGRO, M., S. EUSEPI, M. GIANNONI, A. M. SBORDONE, A. TAMBALOTTI, M. COCCI, R. B. HASEGAWA, AND M. H. LINDER (2013): "The FRBNY DSGE Model," Federal Reserve Bank of New York Staff Paper 647.

FAIR, R. C. AND J. B. TAYLOR (1983): "Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models," *Econometrica*, 51, 1169–1185.

FAUST, J. AND A. GUPTA (2012): "Posterior Predictive Analysis for Evaluating DSGE Models," NBER Working Paper 17906.

FERNALD, J. (2012): "Productivity and Potential Output before, during, and after the Great Recession," Federal Reserve Bank of San Francisco Working Paper 2012-18.

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 (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74, 1059–1087.

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.

——— (2005): *Contemporary Bayesian Econometrics and Statistics*, Hoboken, NJ: Wiley.

GUERRIERI, L. AND M. IACOVIELLO (2015): "OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily," *Journal of Monetary Economics*, 70, 22–38.

——— (2016): "Collateral Constraints and Macroeconomic Asymmetries," Manuscript, Federal Reserve Board.

GUST, C., E. HERBST, D. LÓPEZ-SALIDO, AND M. E. SMITH (2016): "The Empirical Implications of the Interest-Rate Lower Bound," Finance and Economics Discussion Series 2012-083.

HERBST, E. P. AND F. SCHORFHEIDE (2016): *Bayesian Estimation of DSGE Models*, 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.

IRELAND, P. N. (2011): "A New Keynesian Perspective on the Great Recession," *Journal of Money, Credit and Banking*, 43, 31–54.

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 Nonlinear Multivariate Models," *Journal of Econometrics*, 74, 119–147.

NAKATA, T. (2012): "Uncertainty at the Zero Lower Bound," Finance and Economics Discussion Series, 2013-09.

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 (2016): "The Zero Lower Bound and Endogenous Uncertainty," *Economic Journal*, forthcoming.

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 University Press, 294–330.

SIMS, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20, 1–20.

SUH, H. AND T. B. WALKER (2016): "Taking Financial Frictions to the Data," *Journal of Economic Dynamics and Control*, 64, 39–65.

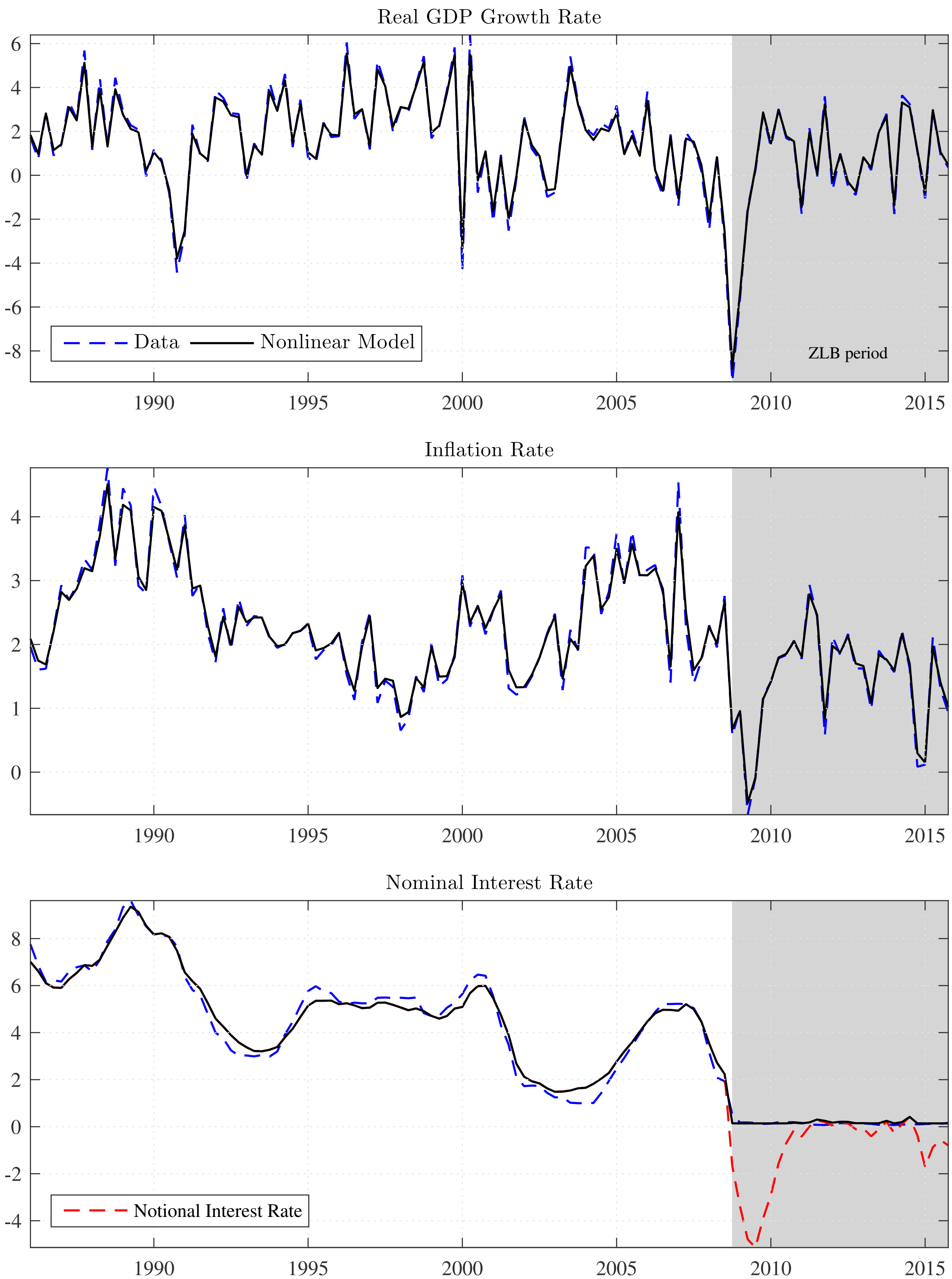

Figure 10: Annualized paths of the observables and their median filtered series from the baseline nonlinear model.

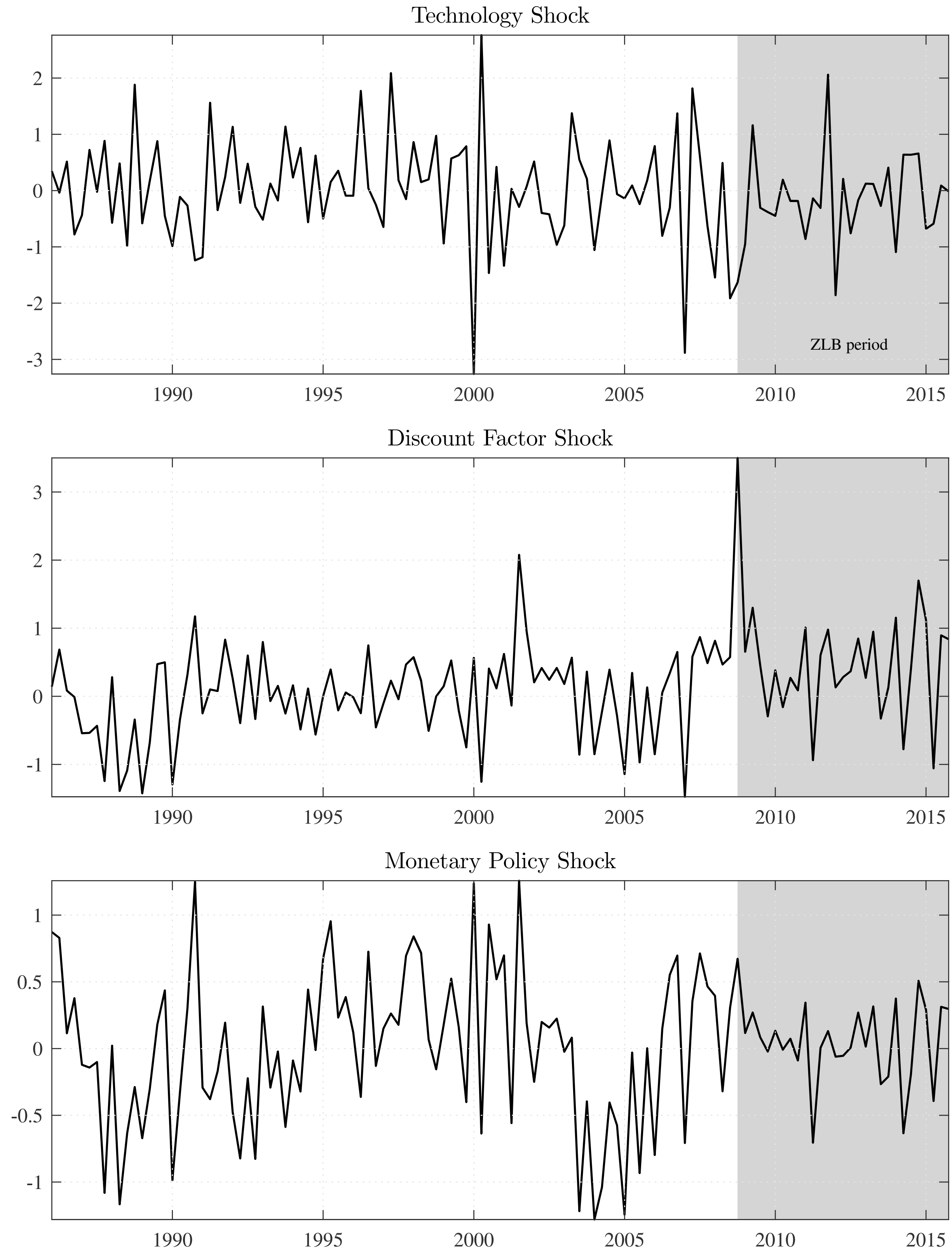

Figure 11: Median filtered shocks in the baseline nonlinear model relative to their posterior mean standard deviation.

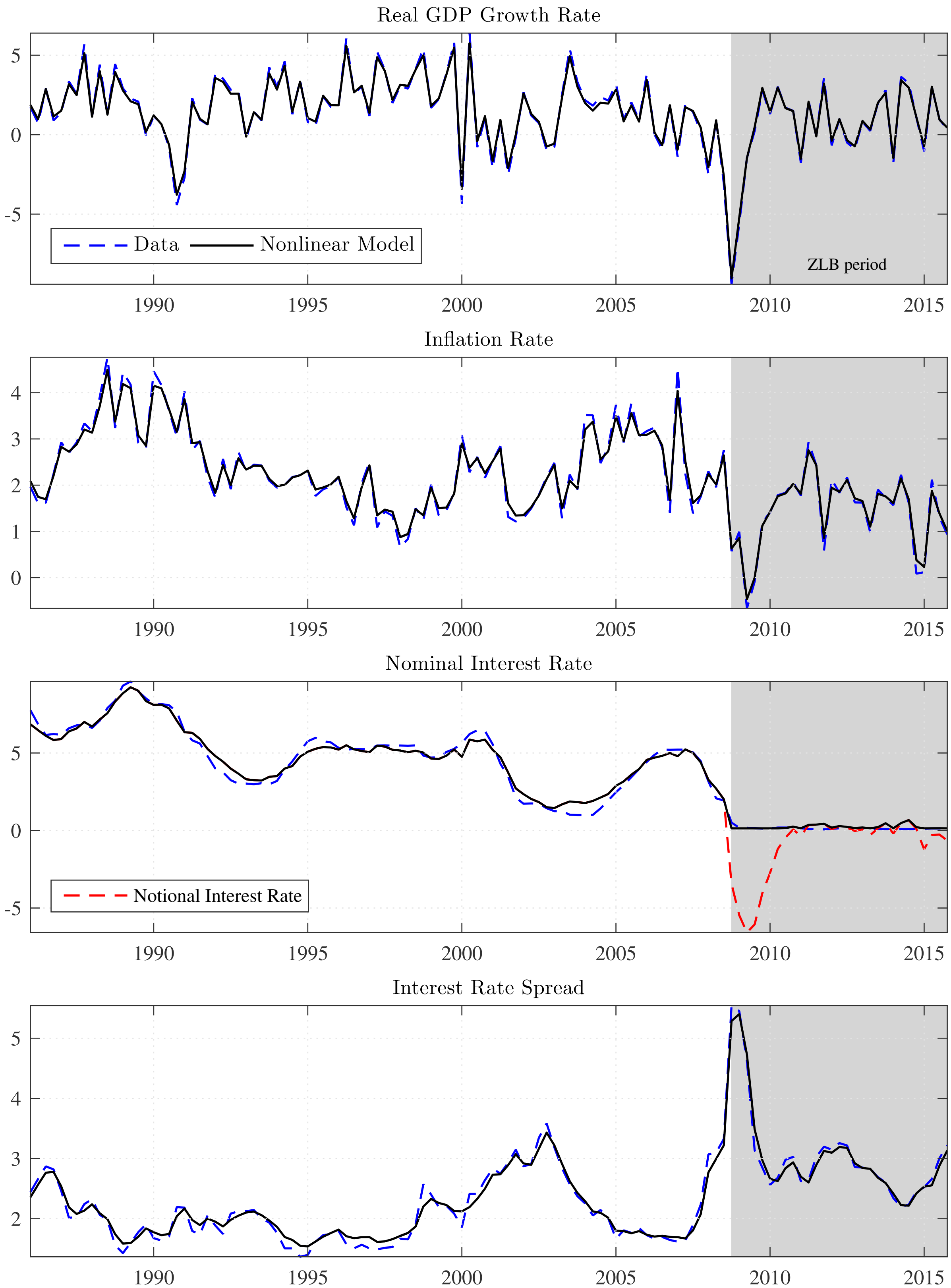

Figure 12: Annualized paths of the observables and their median filtered series from the extended nonlinear model.

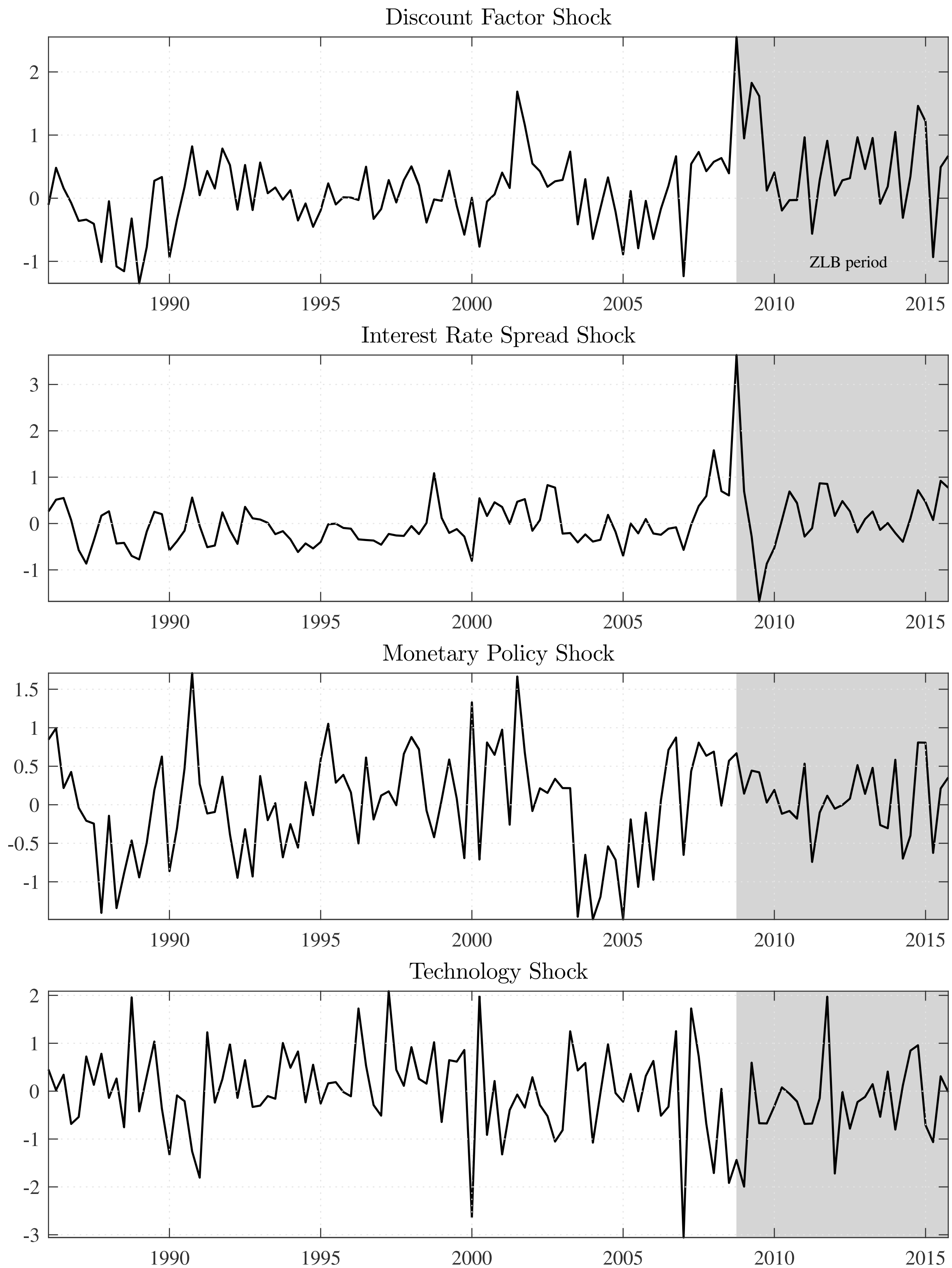

Figure 13: Median filtered shocks in the extended nonlinear model relative to their posterior mean standard deviation.