

Piecewise-Linear Approximations and Filtering for DSGE Models with Occasionally-Binding Constraints *

S. Boragan Aruoba
University of Maryland

Pablo Cuba-Borda
Federal Reserve Board

Kenji Higa-Flores
University of Maryland

Frank Schorfheide
University of Pennsylvania
CEPR, NBER, PIER

Sergio Villalvazo
University of Pennsylvania

Current Version: January 15, 2020

Abstract

We develop an algorithm to construct approximate decision rules that are piecewise-linear and continuous for DSGE models with an occasionally-binding constraint. The functional form of the decision rules allows us to derive a conditionally-optimal particle filter (COPF) for the evaluation of the likelihood function that exploits the structure of the solution. We document the accuracy of the likelihood approximation and embed it into a particle Markov chain Monte Carlo algorithm to conduct Bayesian estimation. Compared to a standard bootstrap particle filter, the COPF significantly reduces the persistence of the Markov chain, improves the accuracy of Monte Carlo approximations of posterior moments, and drastically speeds up computations. We use the techniques to estimate a small-scale DSGE model to assess the effects of the government spending portion of the American Recovery and Reinvestment Act in 2009 when interest rates reached the zero lower bound.

JEL CLASSIFICATION: C5, E4, E5

KEY WORDS: Bayesian Estimation, Nonlinear Filtering, Nonlinear Solution Methods, Particle MCMC, ZLB

* Correspondence: B. Aruoba (aruoba@umd.edu), K. Higa-Flores (kenjihf@umd.edu): Department of Economics, University of Maryland, College Park, MD 20742. P. Cuba-Borda (pablo.a.cubaborda@frb.gov): Division of International Finance, Federal Reserve Board, 20th Street & Constitution Ave., NW, Washington, DC 20551. F. Schorfheide (schorf@ssc.upenn.edu), S. Villalvazo (vservio@sas.upenn.edu): Department of Economics, University of Pennsylvania, 133 S. 36th Street, Philadelphia, PA 19104. We are thankful for helpful comments and suggestions from participants of the 2018 and 2019 MFM conferences, the 2019 conference of the Society for Nonlinear Dynamics, and the Alejandro Justiniano Memorial conference. Higa-Flores and Villalvazo gratefully acknowledge financial support from the Becker-Friedman Institute under the Macro-Financial Modeling Initiative. Aruoba and Schorfheide gratefully acknowledge financial support from the National Science Foundation under Grant SES 1851634. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of anyone else associated with the Federal Reserve System.

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models with financial frictions are widely used in central banks, by regulators, and in academia to study the effects of monetary and macroprudential policies and the propagation of shocks in the macro economy. The most recent vintage of these models involves occasionally-binding constraints arising from financial frictions and the zero lower bound (ZLB) on nominal interest rates. In order for these models to be useable for a quantitative analysis they need to be solved numerically and their parameters need to be estimated based on historical data.

Two types of solution approaches for models with occasionally-binding constraints have been used in the literature. The first group of solution algorithms can be broadly classified as global methods. Agents' decision rules (or value functions associated with optimization problems) are represented by a family of flexible functions, e.g., Chebyshev polynomials, or by a discrete mapping on a finite state space domain. The flexible functions are parameterized by coefficients that are chosen such that the resulting decision rules (approximately) satisfy the model's equilibrium conditions and solve the underlying intertemporal optimization problems.

The second type of solution approaches are variants of the extended perfect-foresight path (EPFP) method that build on Fair and Taylor (1983). These algorithms rely on the assumption that after H periods the system reverts back to the steady state in which the constraint, say, is non-binding. Making an initial guess about whether the constraint is binding in periods $t + h$, $h = 1, \dots, H$, it is possible to solve the dynamic system for the values of the endogenous variables. One then compares the initial guess about the duration of the binding regime to the backward solution and iterates until consistency is achieved. Because the computations are based on the initial state, the previously-described steps need to be repeated for every t in a multi-period simulation.

From the model solution one can construct a state-space representation for an estimable empirical model. The solution itself generates the state transition equations. A set of measurement equations can then be specified that links the state variables with the observables. Because the model solution is nonlinear, so is the state space representation. Thus, a nonlinear filter is required to compute the likelihood function. For instance, in the context of DSGE models with a ZLB constraint, Gust, Herbst, Lopez-Salido, and Smith (2017) and Aruoba, Cuba-Borda, and Schorfheide (2018) use a particle filter in combination with a global solution to construct likelihood functions, whereas Guerrieri and Iacoviello (2017) use an EPFP

solution for a model in which the number of observables equals the number of structural shocks and combine it with an inversion filter that essentially solves for the innovations as a function of the observables conditional on an initial state.

The contribution of our paper is to construct an alternative model solution that (i) is able to capture an important aspect of the decision rule nonlinearity generated by an occasionally-binding constraint, (ii) can be solved quickly, and (iii) allows us to derive an accurate and fast filter for the evaluation of the likelihood function that exploits the structure of the solution. This likelihood approximation can then be embedded into an estimation procedure. One of our goals is to make the procedure efficient enough that it can be run on a desktop computer in a reasonable amount of time. The small-scale New Keynesian model in our empirical application is estimated on U.S. data from 1984 to 2018 in about 13.5 hours on a single core.

The basic idea of the proposed solution method is to approximate agents' decision rules globally by piecewise-linear functions that are continuous but have a kink in the region of the state-space in which a constraint becomes binding. The decision rules account for the fact that in the next period the constraint could either be binding or non-binding and thus they capture precautionary behavior. Moreover, the decision rules only have to be computed once (as opposed to for each period t separately as in the EPFP approach.) Aruoba, Cuba-Borda, and Schorfheide (2018) used more flexible approximations for the decision rules, stitching together higher-order Chebyshev polynomials along the locus in the state-space where the ZLB constraint becomes binding. However, the decision rules on both sides of the kink remained approximately linear. On the one hand, our solution generalizes the widely-used log-linearization techniques which generate log-linear decision rules. On the other hand, compared to higher-order Chebyshev polynomials, the piecewise-linearity and continuity at the kink drastically reduce the number of coefficients that need to be determined and hence simplifies computations.

The resulting solution can be cast into a two-regime vector autoregressive representation with an endogenous regime switching that is triggered when structural shock innovations cross a particular threshold. To solve the nonlinear filtering problem, we derive a conditionally-optimal proposal distribution for a particle filter. A particle filter approximates the distribution of a vector of hidden states s_t conditional on the sequence of observations $Y_{1:t}$ available in time t by a swarm of M particle values and weights $\{s_t^j, W_t^j\}_{j=1}^M$. The filter operates recursively and turns the swarm $\{s_{t-1}^j, W_{t-1}^j\}_{j=1}^M$ into a swarm $\{s_t^j, W_t^j\}_{j=1}^M$, which requires a change in the particle values (mutation) and a change in the particle weights (correction). It has been shown, that conditionally on $\{s_{t-1}^j, W_{t-1}^j\}_{j=1}^M$ it is optimal (see Herbst

and Schorfheide (2015) for a more detailed exposition) to sample $s_t^j \sim p(s_t|y_t, s_{t-1}^j)$. If draws from this distribution can be obtained by direct sampling, the filter becomes very efficient. We show that for state transitions that are piecewise-linear, the conditionally-optimal proposal is a mixture of truncated normal distributions that allows for direct sampling.

In a sequence of numerical illustrations based on a small-scale New Keynesian DSGE model with a ZLB constraint, we document important properties of our solution algorithm and the COPF likelihood approximation. We show that compared to a naive bootstrap particle filter (BSPF) which samples s_t^j from the state-transition density $p(s_t|s_{t-1}^j)$, our conditionally-optimal particle filter (COPF) drastically reduces the variance of the likelihood approximation holding the runtime fixed. In practice, this allows us to run the COPF with far fewer particles than the BSPF (150 in our empirical application), which in turn speeds up the computations. When we embed the more accurate COPF into a random walk Metropolis-Hastings (RWMH) algorithm, we are able to significantly reduce the persistence of the resulting Markov chain and therefore improve the accuracy of Monte Carlo approximations of moments of the posterior distribution.

Using U.S. data from 1984 to 2018 on output growth, inflation, interest rates, and the government-spending to GDP ratio, we estimate the small-scale DSGE model using our proposed piecewise-linear and continuous (PLC) solution in combination with the COPF. From the estimated model we compute dollar-for-dollar government spending multipliers associated with the increase in government spending that was part of the 2009 American Recovery and Reinvestment Act (ARRA). The counterfactual output levels are computed by lowering the exogenous government spending process in the model by an amount that is commensurable to the ARRA intervention and keeping all other exogenous processes at their historical levels. We find that the ex-post ARRA multiplier, computed by comparing realized output and government spending to counterfactual levels, is approximately 0.8.

Our paper is related to several strands of the literature. Global solutions for models with occasionally-binding constraints are computed, for instance, in Christiano and Fisher (2000), Fernández-Villaverde, Gordon, Guerrón-Quintana, and Rubio-Ramírez (2015), Maliar and Maliar (2015), Gust, Herbst, Lopez-Salido, and Smith (2017), and Aruoba, Cuba-Borda, and Schorfheide (2018). The decision rules can also be obtained iteratively through value function or policy function iteration. Examples of the latter methodology include Adam and Billi (2007), Nakata (2016) and Atkinson, Richter, and Throckmorton (2020). Mendoza and Villalvazo (2019) propose a fixed-point iteration algorithm (FiPIt) that iterates on intertemporal equilibrium conditions which is faster than traditional global solution methods. Under

all of these approaches, the decision rules only have to be computed once conditional on a vector of model parameters, and can then easily be used for multi-step model simulation.

EPFP solutions have been developed by Eggertsson and Woodford (2003), Christiano, Eichenbaum, and Trabandt (2015), and Guerrieri and Iacoviello (2015) who provide a software package called *OccBin*. The algorithms are typically designed for models in which the equilibrium conditions, with the exception of the occasionally-binding constraint, are linearized.¹ Kulish, Morley, and Robinson (2017) focus on the ZLB application and treat the duration of the fixed-interest-rate regime as a parameter to be estimated. Holden (2019) implements the EPFP method using endogenous news shocks and provides a software called *DynareOBC*. Boehl (2019) transforms the linearized equilibrium conditions into an extended reduced-form system that depends only on the initial states and the expected number of periods at the constraint and allows for a more efficient computation of the solution. In approaches that utilize EPFP a multi-step simulation requires running the shooting algorithm for every period, making the implementation of a standard particle filter difficult and time consuming.

de Groot, Durdu, and Mendoza (2019) compare the performance of local approximations and EPFP methods to the global solution of open-economy models with incomplete markets and occasionally-binding borrowing constraints. Their results show that simulated moments from local and EPFP methods can differ substantially from those obtained using global solutions; and advice in favor of solving this class of models using the original nonlinear equilibrium conditions. Our methodology follows a similar strategy, with the approximated piecewise-linear decision rules pinned down by the nonlinear system of the equilibrium conditions.

Our solution takes the form of an endogenous regime-switching VAR. Chen (2017) imposes the regime-switching structure in the model specification. In her setup, an exogenous two-state Markov process determines two policy regimes: a normal regime in which the interest rate is positive and follow a Taylor-type monetary policy rule, and a regime in which the interest rate is constrained by the ZLB. Because the regimes are exogenous, the model can be solved using the tools proposed by Farmer, Waggoner, and Zha (2011). Similarly, Bianchi and Melosi (2017) use an exogenous regime-switching process to characterize the ZLB dynamics during the Great Recession. Benigno, Foerster, Otrok, and Rebucci (2016) endogenize the regime switching probability in a model of financial crisis, but, unlike in

¹The resulting solution is often called piecewise-linear which is a bit misleading. It is the equilibrium conditions that are piecewise-linear. In contrast, our approach generates piecewise-linear decision rules.

our paper, the transition from one to the other regime remains partly decoupled from the realization of the fundamental shocks.

A key feature of our paper is that it integrates model solution, likelihood approximation, and Bayesian estimation. There are a few papers that assess the interplay of existing model solution and likelihood evaluation techniques in Monte Carlo experiments. Cuba-Borda, Guerrieri, Iacoviello, and Zhong (2019) take a simple consumption-savings model subject to a borrowing constraint. They illustrate that less accurate solution methods affect inference even when the inversion filter is available. They also show that as one increases the measurement error variance in the BSPF, the likelihood misspecification becomes more problematic, making it harder to retrieve the parameter values that govern the DGP. In their setting, measurement error and solution approximation error makes it difficult for the econometrician to identify the model regime that generates the data, and this incorrect classification of regimes leads to a bias in parameter estimates. In our empirical application one of the observed time series allow us to exactly identify the regime and we modify the COPF to capture this feature.

Atkinson, Richter, and Throckmorton (2020) compare the performance of a fully nonlinear approach and using OccBin with the Inversion Filter in estimation. They simulate data from a model that includes more frictions and shocks than their test model where the latter is close to the model we use in this paper in terms of size. As such, their model is misspecified with respect to the data generating process. Their results show that the nonlinear approach performs slightly better than the OccBin approach but the differences are small. Moreover, relative to the pseudo-true parameters the estimates from both approaches show biases in some key parameters such as the degree of price rigidities. Since the OccBin-Inversion Filter approach can be scaled up easily and is faster, they argue that building a bigger and less misspecified model using this approach may be preferable. Similarly, our method offers scalability even without multicore processing or distributed computing.

Boehl (2019) combines his model solution with a variant of an ensemble Kalman filter. His paper does not focus on accuracy comparisons of solution and estimation methods. Instead, it applies the proposed techniques to estimate a small-scale New Keynesian model on U.S. data. Similarly, our paper uses U.S. data to estimate the parameters of a small-scale New Keynesian model and quantifies the role of fiscal policy in the aftermath of the Great Recession.

There is a large literature that studies fiscal spending multipliers using methodologies

ranging from using DSGE models, structural vector autoregressions (SVARs), to more traditional macroeconometric models; see Batini, Eyraud, Forni, and Weber (2014) for a survey and references. Results from DSGE models tend to be quite sensitive to the size of the fiscal shock and modeling assumptions with respect to monetary and fiscal policy rules and labor supply (Hills and Nakata (2018)). A fairly robust finding across most studies using DSGE models is that multipliers are larger at the ZLB than away from the ZLB. See, for example, Eggertsson (2011) and Christiano and Eichenbaum (2012). Using a fairly stylized NK model with passive fiscal policy and preferences that are additively separable in terms of consumption and leisure, we find that the ex-post multiplier during the Great Recession when the U.S. was at the ZLB, was larger than in normal times when interest rates are positive. One novel result we show through counterfactuals is that in 2009 and 2010 there was very little room for the Federal Reserve to stimulate the economy with conventional monetary policy over and above what the policy rule implied because adverse shocks kept the desired interest rate near zero despite the large expansionary fiscal policy due to ARRA.

The remainder of the paper is organized as follows. Section 2 describes the small-scale New Keynesian DSGE model with ZLB constraint used in the subsequent analysis. In Section 3 we describe how to impose continuity on piecewise-linear decision rules and derive a canonical form for the DSGE model solution. Section 4 discusses how the decision rule coefficients are determined to approximately satisfy the model's equilibrium conditions. The COPF is derived in Section 5. Section 6 presents some numerical experiments to document the accuracy of the likelihood approximation through the COPF and Section 7 contains the empirical analysis. Finally, Section 8 concludes. Derivations and further implementation details are provided in the Online Appendix.

2 A Prototypical New Keynesian DSGE Model

We will illustrate our solution and filtering methods based on a prototypical New Keynesian DSGE model. The model is identical to the one used in Aruoba, Cuba-Borda, and Schorfheide (2018). Variants of this model have been widely studied in the literature and its properties are discussed in detail in Woodford (2003). The model economy consists of perfectly competitive final-goods-producing firms, a continuum of monopolistically competitive intermediate goods producers, a continuum of identical households, and a government that engages in monetary and fiscal policy. To make this paper self-contained and introduce some

important notation, we describe the preferences and technologies of the agents in Section 2.1, and summarize the equilibrium conditions in Section 2.2.

2.1 Preferences and Technologies

Households. Households derive utility from consumption C_t relative to an exogenous habit stock and disutility from hours worked H_t . We assume that the habit stock is given by the level of technology A_t , which ensures that the economy evolves along a balanced growth path. We also assume that the households value transaction services from real money balances, detrended by A_t , and include them in the utility function. The households maximize

$$\mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s d_{t+s} \left(\frac{(C_{t+s}/A_{t+s})^{1-\tau} - 1}{1-\tau} - \chi_H \frac{H_{t+s}^{1+1/\eta}}{1+1/\eta} + \chi_M V \left(\frac{M_{t+s}}{P_{t+s}A_{t+s}} \right) \right) \right], \quad (1)$$

subject to the budget constraint

$$P_t C_t + T_t + M_t + B_t = P_t W_t H_t + M_{t-1} + R_{t-1} B_{t-1} + P_t D_t + P_t S C_t.$$

Here β is the discount factor, d_t is a shock to the discount factor, $1/\tau$ is the intertemporal elasticity of substitution, η is the Frisch labor supply elasticity, and P_t is the price of the final good. The shock d_t captures frictions that affect intertemporal preferences in a reduced-form way. Fluctuations in d_t affect households patience and their desire to postpone consumption. As is commonly exploited in the literature, a sufficiently large negative shock to d_t makes the central bank cut interest rates all the way to the ZLB. The households supply labor services to the firms in a perfectly competitive labor market, taking the real wage W_t as given. At the end of period t , households hold money in the amount of M_t . They have access to a bond market where nominal government bonds B_t that pay gross interest R_t are traded. Furthermore, the households receive profits D_t from the firms and pay lump-sum taxes T_t . SC_t is the net cash inflow from trading a full set of state-contingent securities.

Firms. The final-goods producers aggregate intermediate goods, indexed by $j \in [0, 1]$, using the technology

$$Y_t = \left(\int_0^1 Y_t(j)^{1-\nu} dj \right)^{\frac{1}{1-\nu}}.$$

The firms take input prices $P_t(j)$ and output prices P_t as given. Profit maximization implies

that the demand for inputs is given by

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-1/\nu} Y_t.$$

Under the assumption of free entry into the final-goods market, profits are zero in equilibrium, and the price of the aggregate good is given by

$$P_t = \left(\int_0^1 P_t(j)^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}}. \quad (2)$$

We define inflation as $\pi_t = P_t/P_{t-1}$.

Intermediate good j is produced by a monopolist who has access to the production technology

$$Y_t(j) = A_t H_t(j), \quad (3)$$

where A_t is an exogenous productivity process that is common to all firms and $H_t(j)$ is the firm-specific labor input. Intermediate-goods-producing firms face quadratic price adjustment costs of the form

$$AC_t(j) = \frac{\phi}{2} \left(\frac{P_t(j)}{P_{t-1}(j)} - \bar{\pi} \right)^2 Y_t(j),$$

where ϕ governs the price stickiness in the economy and $\bar{\pi}$ is a baseline rate of price change that does not require the payment of any adjustment costs. In our quantitative analysis, we set $\bar{\pi} = \pi_*$, where π_* is the target inflation rate of the central bank. Firm j chooses its labor input $H_t(j)$ and the price $P_t(j)$ to maximize the present value of future profits

$$\mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s Q_{t+s|t} \left(\frac{P_{t+s}(j)}{P_{t+s}} Y_{t+s}(j) - W_{t+s} H_{t+s}(j) - AC_{t+s}(j) \right) \right]. \quad (4)$$

Here, $Q_{t+s|t}$ is the time t value to the household of a unit of the consumption good in period $t+s$, which is treated as exogenous by the firm.

Government Policies. Monetary policy is described by an interest rate feedback rule. Because the ZLB constraint is an important part of our analysis we introduce it explicitly as follows:

$$R_t = \max \{1, R_t^* e^{\sigma_{R^e, t}}\}. \quad (5)$$

Here R_t^* is the systematic part of monetary policy which reacts to the current state of the

economy

$$R_t^* = \left[r\pi_* \left(\frac{\pi_t}{\pi_*} \right)^{\psi_1} \left(\frac{Y_t}{\gamma Y_{t-1}} \right)^{\psi_2} \right]^{1-\rho_R} R_{t-1}^{\rho_R},$$

where r is the steady-state real interest rate and π_* is the target-inflation rate. $\epsilon_{R,t}$ is a monetary policy shock. Provided that the ZLB is not binding, the central bank reacts to deviations of inflation from the target rate π_* and deviations of output growth from its long-run value γ .

The government consumes a stochastic fraction of aggregate output. We assume that government spending evolves according to

$$G_t = \left(1 - \frac{1}{g_t} \right) Y_t \quad (6)$$

where g_t is an exogenous process. The government levies a lump-sum tax T_t (or provides a subsidy if T_t is negative) to finance any shortfalls in government revenues (or to rebate any surplus). Its budget constraint is given by

$$P_t G_t + M_{t-1} + R_{t-1} B_{t-1} = T_t + M_t + B_t. \quad (7)$$

Exogenous shocks. The model economy is perturbed by four exogenous processes. Aggregate productivity evolves according to

$$\log A_t = \log \gamma + \log A_{t-1} + \log z_t, \text{ where } \log z_t = \rho_z \log z_{t-1} + \sigma_z \epsilon_{z,t}. \quad (8)$$

Thus, on average, the economy grows at the rate γ , and z_t generates exogenous stationary fluctuations of the technology growth rate around this long-run trend. We assume that the government spending shock follows the AR(1) law of motion

$$\log g_t = (1 - \rho_g) \log g_* + \rho_g \log g_{t-1} + \sigma_g \epsilon_{g,t}. \quad (9)$$

The shock to the discount factor evolves according to

$$\log d_t = \rho_d \log d_{t-1} + \sigma_d \epsilon_{d,t} \quad (10)$$

The monetary policy shock $\epsilon_{R,t}$ is assumed to be serially uncorrelated. We stack the four innovations into the vector $\epsilon_t = [\epsilon_{z,t}, \epsilon_{g,t}, \epsilon_{d,t}, \epsilon_{R,t}]'$ and assume that $\epsilon_t \sim iidN(0, I)$.

2.2 Equilibrium Conditions

Because the exogenous productivity process has a stochastic trend, it is convenient to characterize the equilibrium conditions of the model economy in terms of detrended consumption $c_t \equiv C_t/A_t$ and detrended output $y_t \equiv Y_t/A_t$.

It is well known that the New Keynesian model features multiple equilibria; see Aruoba and Schorfheide (2016). In the remainder of the paper we focus on what the literature has called the targeted-inflation equilibrium. This is essentially the equilibrium that arises in linearized DSGE models, adjusted for the presence of the ZLB constraint. The corresponding steady state is given by

$$\pi_*, \quad r_* = \frac{\gamma}{\beta}, \quad R_* = r_* \pi_*, \quad y_* = \left(\frac{1 - \nu}{\chi_H} g_*^\tau \right)^{\frac{1}{\tau+1/\eta}}, \quad c_* = \frac{y_*}{g_*}.$$

Without loss of generality, for any variable x_t we can define the percentage deviations from the steady state as $\hat{x}_t = \ln x_t - \ln x_*$. Using this notation we can substitute x_t by $x_* e^{\hat{x}_t}$.² Our goal is to write the equilibrium conditions as a system of expectational difference equations of the form

$$\mathbb{E}_t[\mathcal{R}(\hat{y}_t, \hat{c}_t, \hat{\pi}_t, \hat{R}_t, \hat{y}_{t+1}, \hat{c}_{t+1}, \hat{\pi}_{t+1}, \hat{R}_{t+1}, \dots)] = 0, \quad (11)$$

where $\mathcal{R}(\cdot)$ captures residuals in the equilibrium conditions.

The residual function comprises of the following elements. The consumption Euler equation leads to

$$\mathcal{R}_c(\cdot) = \hat{d}_{t+1} - \hat{d}_t - \tau(\hat{c}_{t+1} - \hat{c}_t) + \hat{R}_t - \hat{\pi}_{t+1} - \hat{z}_{t+1}. \quad (12)$$

In a symmetric equilibrium, in which all firms set the same price $P_t(j)$, the price-setting decision of the firms leads to

$$\begin{aligned} \mathcal{R}_\pi(\cdot) = & \ln \left[\frac{1}{\nu} \left(\frac{1}{\nu} + \left(1 - \frac{1}{\nu} \right) e^{\tau \hat{c}_t + \hat{y}_t / \eta} \right) - \phi \pi_*^2 (e^{\hat{\pi}_t} - 1) \left[\left(1 - \frac{1}{2\nu} \right) e^{\hat{\pi}_t} + \frac{1}{2\nu} \right] \right. \\ & \left. + \phi \beta \pi_*^2 \left(e^{\hat{d}_{t+1} - \hat{d}_t} \right) \left(e^{-\tau(\hat{c}_{t+1} - \hat{c}_t)} \right) \left(e^{\hat{y}_{t+1} - \hat{y}_t} \right) \left(e^{\hat{\pi}_{t+1}} - 1 \right) e^{\hat{\pi}_{t+1}} \right]. \end{aligned} \quad (13)$$

²Introducing \hat{x}_t does not imply that we are log-linearizing all of the equilibrium conditions. It is foremost a reparameterization. However, in our model it happens to be the case that the consumption Euler equation and (abstracting from the max operator) the monetary policy rule are log-linear.

The aggregate resource constraint leads to

$$\mathcal{R}_y(\cdot) = \hat{y}_t - \hat{c}_t + \ln \left[\frac{1}{e^{\hat{g}_t}} - \frac{\phi}{2} g_* (\pi_* e^{\hat{\pi}_t} - \bar{\pi})^2 \right]. \quad (14)$$

It reflects both government spending as well as the resource cost (in terms of output) caused by price changes. The monetary policy rule generates the residual function

$$\mathcal{R}_R(\cdot) = \hat{R}_t - \max \left\{ (1 - \rho_R) [\psi_1 \hat{\pi}_t + \psi_2 (\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t)] + \rho_R \hat{R}_{t-1} + \sigma_R \epsilon_{R,t}, -\log(r\pi_*) \right\}. \quad (15)$$

We do not use a measure of money in the subsequent analysis and therefore drop the equilibrium condition that determines money demand. We stack the residual functions for the exogenous shocks as follows:

$$\mathcal{R}_{exo}(\cdot) = \begin{bmatrix} \hat{z}_t - \rho_z \hat{z}_{t-1} - \sigma_z \epsilon_{z,t} \\ \hat{d}_t - \rho_d \hat{d}_{t-1} - \sigma_d \epsilon_{d,t} \\ \hat{g}_t - \rho_g \hat{g}_{t-1} - \sigma_g \epsilon_{g,t} \\ e_{R,t} - \sigma_R \epsilon_{R,t} \end{bmatrix}. \quad (16)$$

3 PLC Decision Rules and the Canonical Form

Throughout this paper we consider solutions for models with a single occasionally-binding constraint. In this section we show how we construct our piecewise-linear decision rules and how we impose continuity at the kink, where the constraint changes from being slack to being binding. In Sections 3.1 and 3.2 we do so using a generic notation that can be used with any model with an occasionally-binding constraint. We provide an illustration in the context of the New Keynesian DSGE model in Section 3.3. Finally, in Section 3.4 we construct what we call the canonical form of the model, which will make the subsequent discussion about filtering and estimation easier.

3.1 PLC Decision Rules

Let $\mathbb{X} = [x_1, X_2'] \in \mathcal{X}$ be an $n \times 1$ vector of state variables. Here x_1 is one particular element of \mathbb{X} and we assume that the vector of the remaining state variables, X_2 , also contains a constant. Let y denote a $k \times 1$ vector of control variables. We assume that there is a

linear(ized) function $h(x_1, X_2, y)$ that determines whether the constraint is binding:

$$h(x_1, X_2, y) = \begin{cases} > 0 & \text{if constraint is non-binding (n)} \\ \leq 0 & \text{if constraint is binding (b)} \end{cases} . \quad (17)$$

The $h(\cdot)$ function may depend on the state variables (x_1, X_2) and some of the elements in y . Because the function is assumed to be linear, we write it as

$$h(x_1, X_2, y) = \gamma_1 x_1 + \gamma'_2 X_2 + \gamma'_y y. \quad (18)$$

The γ 's are not free coefficients. They are obtained from the equilibrium conditions of the DSGE model. We define the kink function $x_1 = \ell(X_2)$ such that $[\ell(X_2), X'_2]' \in \mathcal{X}$ characterizes the locus of points in the state-space for which $h(\ell(X_2), X_2, y(\ell(X_2), X_2)) = 0$, i.e., the constraint is just binding.

The linearity of $h(\cdot)$ in (18) and the assumed piecewise-linearity of the decision rules for y imply that $\ell(X_2)$ is a linear function and we parameterize it as

$$\ell(X_2) = \delta' X_2. \quad (19)$$

Here δ is an $(n - 1)$ -dimensional vector. The δ coefficients are currently free and will be determined as functions of the decision rule coefficients and the coefficients of the constraint function $h(\cdot)$ as we explain below. Note that so far we have not yet made a determination whether the constraint is slack if $x_1 < \delta' X_2$.

Returning to the control variables, we assume the decision rules for each y^i are of the piecewise-linear form

$$y^i(x_1, X_2) = \begin{cases} \alpha_{1,1}^i x_1 + \alpha_{1,2}^{i'} X_2 & \text{if } x_1 \geq \ell(X_2) \\ \alpha_{2,1}^i x_1 + \alpha_{2,2}^{i'} X_2 & \text{if } x_1 < \ell(X_2) \end{cases} \quad i = 1, \dots, k, \quad (20)$$

where each decision rule has $2n$ unknown coefficients. The decision rules are exactly linear if $\alpha_{1,1}^i = \alpha_{2,1}^i$ and $\alpha_{1,2}^i = \alpha_{2,2}^i$. The specification in (20) makes the benefit of using the kink function $\ell(\cdot)$ clear: given the state variables x_1 and X_2 , we can easily determine which side of the constraint we need to be, even when the constraint contains some control variables.

3.2 Imposing Continuity on Piecewise-Linear Decision Rules

We now turn to imposing continuity on the decision rules at the kink, which means we impose the restriction that the two parts of each decision rule are equal to each other along the kink. Doing so will restrict a subset of the unknown α and δ coefficients.

More specifically, continuity at $x_1 = \delta' X_2$ requires that for each $i = 1, \dots, k$

$$\alpha_{1,1}^i \delta' X_2 + \alpha_{1,2}^{i'} X_2 = \alpha_{2,1}^i \delta' X_2 + \alpha_{2,2}^{i'} X_2 \quad \forall X_2,$$

which generates $(n - 1)$ restrictions for each i :

$$\alpha_{1,1}^i \delta' + \alpha_{1,2}^{i'} = \alpha_{2,1}^i \delta' + \alpha_{2,2}^{i'}. \quad (21)$$

Next, we impose restrictions that make the $\ell(\cdot)$ and $y(\cdot)$ functions consistent with the constraint in (17). The condition $h[g(X_2), X_2, y(\ell(X_2), X_2)] = 0$, which represents the kink in terms of the $h(\cdot)$ function, can be written as

$$\gamma_1 \delta' X_2 + \gamma_2' X_2 + \sum_{i=1}^k \gamma_y^i (\alpha_{1,1}^i \delta' X_2 + \alpha_{1,2}^{i'} X_2) = 0 \quad \forall X_2,$$

which leads to another $(n - 1)$ restrictions:

$$\gamma_1 \delta' + \gamma_2' + \sum_{i=1}^k \gamma_y^i (\alpha_{1,1}^i \delta' + \alpha_{1,2}^{i'}) = 0. \quad (22)$$

Counting all unknowns and restrictions, we have $k(n + 1)$ degrees of freedom.³ Let us assume the coefficients $\alpha_{1,1}^i$, $\alpha_{1,2}^{i'}$ and $\alpha_{2,1}^i$ for each decision rule are free and collect them in the vector ϑ of size $k(n + 1)$

$$\vartheta = [\alpha_{1,1}^1, \dots, \alpha_{1,1}^k, \alpha_{1,2}^{1'}, \dots, \alpha_{1,2}^{k'}, \alpha_{2,1}^1, \dots, \alpha_{2,1}^k]'. \quad (23)$$

In other words, we treat all the decision rule coefficients for the “1” regime and the coefficient in front of x_1 in the “2” regime as free. The remaining decision rule coefficients in the “2” regime, $\alpha_{2,2}^i$, $i = 1, \dots, k$, as well as all of the δ coefficients are determined as functions of these free coefficients, which we now turn to.

³There are $2n$ α coefficients for each decision rule and $(n - 1)$ δ coefficients, which yield $2nk + n - 1$ unknowns. With $(n - 1)$ restrictions for each decision rule as derived in (21) and the $(n - 1)$ restrictions in (22), we get $(k + 1)(n - 1)$ restrictions. Subtracting the number of restrictions from the number of unknowns, we get $k(n + 1)$.

Conditional on ϑ , we can rewrite (22) as

$$\underbrace{\left(\gamma_1 + \sum_{i=1}^k \gamma_y^i \alpha_{1,1}^i\right)}_{a(\vartheta)} \delta'(\vartheta) + \underbrace{\left(\gamma_2 + \sum_{i=1}^k \gamma_y^i \alpha_{1,2}^i\right)}_{-B'(\vartheta)} = 0$$

and solve for δ as

$$\delta'(\vartheta) = \frac{1}{a(\vartheta)} B'(\vartheta), \quad (24)$$

where $a(\vartheta)$ is a scalar and $b(\vartheta)$ is a $(n-1)$ -dimensional vector. By combining (24) with (21) we obtain an expression for the constrained decision rule coefficients $\alpha_{2,2}^i$:

$$\alpha_{2,2}^{i'}(\vartheta) = (\alpha_{1,1}^i - \alpha_{2,1}^i) \left(\frac{1}{a(\vartheta)} B(\vartheta) \right) + \alpha_{1,2}^{i'}. \quad (25)$$

The last step is to determine which part of the decision rule in (20) corresponds to the part of the state space where the constraint is slack and which part where the constraint is binding. Take $h(x_1, X_2, y(x_1, X_2))$ for some x_1 and X_2 . Let us derive how its sign depends on the sign of $(x_1 - \delta(\vartheta)'X_2)$. First, assume $x_1 > \delta'(\vartheta)X_2$, then

$$\begin{aligned} h[x_1, X_2, y(x_1, X_2)] &= \gamma_1 x_1 + \gamma_2' X_2 + \sum_{i=1}^k \gamma_y^i (\alpha_{1,1}^i x_1 + \alpha_{1,2}^{i'} X_2) \\ &= \underbrace{\left(\gamma_1 + \sum_{i=1}^k \gamma_y^i \alpha_{1,1}^i\right)}_{c(\vartheta)} x_1 + \underbrace{\left(\gamma_2' + \sum_{i=1}^k \gamma_y^i \alpha_{1,2}^{i'}\right)}_{D'(\vartheta)} X_2, \end{aligned} \quad (26)$$

where $c(\vartheta)$ is a scalar and $D(\vartheta)'$ is a $(n-1)$ -dimensional vector, which can be evaluated given model parameters and free decision-rule coefficients. We can rewrite (22) as

$$\left[\gamma_1 + \sum_{i=1}^k \gamma_y^i \alpha_{1,1}^i \right] \delta' + \left[\gamma_2' + \sum_{i=1}^k \gamma_y^i \alpha_{1,2}^{i'} \right] = 0, \quad (27)$$

or simply $c(\vartheta)\delta'(\vartheta) + D'(\vartheta) = 0$ which implies $D'(\vartheta) = -c(\vartheta)\delta'(\vartheta)$. Using this on (26) we get

$$h[x_1, X_2, y(x_1, X_2)] = c(\vartheta)[x_1 - \delta'(\vartheta)X_2]. \quad (28)$$

Since we assumed $x_1 > \delta'(\vartheta)X_2$ above in the derivations, we conclude that $h(\cdot) > 0$ if and

only if $c(\vartheta) > 0$. In other words, if $c(\vartheta) > 0$, then the “1” regime in (20) corresponds to the constraint being slack.

3.3 Example: New Keynesian Model

In this section we adapt the generic notation in the previous sections to the New Keynesian model with the ZLB outlined in Section 2. We partition the state space as $x_{1,t} = \hat{R}_{t-1}$ and $X_{2,t} = [1, \hat{y}_{t-1}, \hat{z}_t, \hat{d}_t, \hat{g}_t, e_{R,t}]'$ and we have $n = 7$.⁴ As for the choice of control variables to approximate, we have a few options. We approximate the decision rules $\hat{\pi}(x_1, X_2)$ and $\hat{y}(x_1, X_2)$ directly and let the remaining control variables \hat{c} and \hat{R} follow exactly from the equilibrium conditions. Specifically, given x_1, X_2 , $\hat{y}_t = \hat{y}(x_1, X_2)$ and $\hat{\pi}_t = \hat{\pi}(x_1, X_2)$, $\hat{c}(x_1, X_2)$ follows from solving for \hat{c}_t in $\mathcal{R}_y(\cdot)$ in (14) and $\hat{R}(x_1, X_2)$ follows from solving for \hat{R}_t in $\mathcal{R}_R(\cdot)$ in (15). Thus, with a slight abuse of notation, we set $y(\cdot) = [\hat{y}(\cdot), \hat{\pi}(\cdot)]'$ and $k = 2$.

The constraint in this problem is the ZLB constraint which can be written in terms of the variables defined so far as $\hat{R}_t + \log(r_*\pi_*) \geq 0$, which leads to the $h(\cdot)$ function

$$h(x_{1,t}, X_{2,t}, y_t(\cdot)) = (1 - \rho)[\psi_1 \hat{\pi}(\cdot) + \psi_2(\hat{y}(\cdot) - \hat{y}_{t-1} + \hat{z}_t)] + \rho \hat{R}_{t-1} + e_{R,t} + \log(r_*\pi_*). \quad (29)$$

Thus, the γ coefficients in (18) are

$$\gamma_1 = \rho, \quad \gamma'_2 = [\log(r_*\pi_*), -(1 - \rho)\psi_2, (1 - \rho)\psi_2, 0, 0], \quad \gamma'_y = [(1 - \rho)\psi_1, (1 - \rho)\psi_2]$$

and we can write $c(\vartheta)$ in (26) as

$$c(\vartheta) = \rho + (1 - \rho)\psi_1\alpha_{1,1}^\pi + (1 - \rho)\psi_2\alpha_{1,1}^y.$$

If $\alpha_{1,1}^\pi$ and $\alpha_{1,1}^y$ are both positive, then $c(\vartheta)$ is also positive because the remaining structural parameters are positive under standard parameterizations. Thus, we label the “1” regime as the one where the ZLB is slack, ‘n’ (non-binding) and the “2” regime as the ‘b’ (binding) regime. The decision rules in the regime when the ZLB is slack are not identical but closely resemble the linear decision rules of the model linearized around the steady state.

⁴In principle, one of the other state variables could have been chosen as x_1 . However, we found it natural to use the lagged interest rates, because all else equal, higher lagged interest rates move the economy away from the ZLB constraint.

These decision rules do have the property that the coefficients on \hat{R}_{t-1} in the π and y decision rules are positive. In our implementation we check that $c(\vartheta)$ is indeed positive every time we solve the model.

3.4 Canonical Form

The final step in preparing the model solution for filtering is to cast the solution in terms of the following canonical form:

$$s_t = \begin{cases} \Phi_0(n) + \Phi_1(n)s_{t-1} + \Phi_\eta(n)\eta_t & \text{if } \eta_{1,t} < \zeta(s_{t-1}) \\ \Phi_0(b) + \Phi_1(b)s_{t-1} + \Phi_\eta(b)\eta_t & \text{otherwise,} \end{cases} \quad (30)$$

which is a vector autoregressive representation for a vector of model variables s_t . It will serve as a transition equation in a state-space model. Thus, the vector s_t needs to include all variables (whether or not they are directly approximated) that are necessary for the construction of the measurement equations and all variables necessary to determine the transition of such variables, which are all the state variables. The canonical form resembles a regime-switching VAR with a “binding” (b) and “non-binding” (n) regime. However, the regime transition is not determined by an exogenous Markov process. Instead, it is determined by the realization of the shock innovations. The vector η_t is a linear transformation of the structural shock innovations of the DSGE model.

Because the definition of s_t is model and application specific, we outline the construction of the canonical form in the context of the New Keynesian DSGE model with ZLB constraint. Define the vectors s_t

$$s_t = [\hat{y}_t, \hat{\pi}_t, \hat{R}_t, \hat{z}_t, \hat{d}_t, \hat{g}_t, e_{R,t}]' \quad (31)$$

and recall that $\epsilon_t = [\epsilon_{z,t}, \epsilon_{d,t}, \epsilon_{g,t}, \epsilon_{R,t}]'$. We will begin by expressing the law of motion of s_t as a function of the innovations ϵ_t and then, later on, we transform the ϵ_t 's into η 's:

$$s_t = \Phi_0(\cdot) + \Phi_1(\cdot)s_{t-1} + \Phi_\epsilon(\cdot)\epsilon_t.$$

Rather than providing detailed algebraic expressions for the elements of the $\Phi(\cdot)$ matrices, we will provide an outline of how the expressions can be derived.

Output, inflation, and interest rates. We use the first three rows of the $\Phi(\cdot)$ matrices to represent the decision rules for \hat{y}_t and $\hat{\pi}_t$ and the monetary policy rule that determines \hat{R}_t .

Note that the decision rules in (20) are expressed in terms of $\mathbb{X}_t = [\hat{R}_{t-1}, 1, \hat{y}_{t-1}, \hat{z}_t, \hat{d}_t, \hat{g}_t, e_{R,t}]'$, whereas the canonical form is written in terms of s_t ; see (31). Thus, in order to generate the equations for \hat{y}_t , $\hat{\pi}_t$, and \hat{R}_t for the canonical form, we have to express \mathbb{X}_t as a linear function of s_{t-1} and ϵ_t .

Exogenous shocks. The remaining four rows of the $\Phi(\cdot)$ matrices reproduce the law of motion of the exogenous shock processes in (16).

From ϵ_t 's to η_t 's and defining the threshold condition. To express the threshold condition in the canonical form and transform the ϵ_t into η_t innovations, define

$$\zeta(s_{t-1}) = \log(r_*\pi_*) + \phi_0(n) + \phi'_1(n)s_{t-1}, \quad \eta_{1,t} = -\frac{1}{\|\phi_\epsilon(n)\|}\phi'_\epsilon(n)\epsilon_t$$

such that the ZLB constraint is non-binding if and only if

$$\eta_{1,t} < \zeta(s_{t-1})$$

as in (30). Let $\text{Null}(x)$ be an orthogonal basis for the null space for the vector x . We define the vector η_t as

$$\eta_t = \begin{bmatrix} \phi'_\epsilon(n)/\|\phi_\epsilon(n)\| \\ \text{Null}(\phi_\epsilon(n)/\|\phi_\epsilon(n)\|)' \end{bmatrix} \epsilon_t.$$

The transformation has the property that if $\mathbb{E}[\epsilon_t\epsilon_t'] = I$, then $\mathbb{E}[\eta_t\eta_t'] = I$ as well. The definition of η_t as function of ϵ_t allows us to convert the $\Phi_\epsilon(\cdot)$ matrix into $\Phi_\eta(\cdot)$ and completes the derivation of the canonical form.

3.5 Measurement Equations

The key requirement for the conditionally-optimal particle filter that is developed in Section 5 is that the conditional mean function (given s_{t-1}) of the observables is piecewise-linear. This is guaranteed if the state-transition equation has the canonical form (30) and the measurement equation is linear in s_t as in

$$y_t^o = A_0 + A_s s_t + u_t, \quad u_t \sim N(0, \varsigma \Sigma_u), \quad (32)$$

where y_t^o is the vector of observable variables, u_t is a vector of measurement errors and the constant ς allows us to scale the measurement error covariance.

The small-scale New Keynesian DSGE model is typically estimated based on output growth $y_{gr,t}^o$, inflation π_t^o , and interest rates R_t^o . In addition, we will include a measure of the consumption-output ratio. Starting from the definition of s_t given in (31) we define the augmented vector $\tilde{s}_t = [s'_t, \hat{y}_{t-1}]'$ and add the trivial equation $\hat{y}_{t-1} = \hat{y}_{t-1}$ to the canonical form in (30). Because the \hat{y}_{t-1} identity is linear, the structure of the canonical form is preserved. Assuming that output growth is measured in quarter-on-quarter percentages, and inflation and interest rates are measured in annualized percentages, the system of measurement equations is

$$\begin{aligned} y_{gr,t}^o &= 100 \log(\gamma) + 100(\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t) + \sigma_{u,y} u_{y,t} \\ \pi_t^o &= 400 \log(\pi_*) + 400 \hat{\pi}_t + \sigma_{u,\pi} u_{\pi,t} \\ R_t^o &= 400 \log(R_*) + 400 \hat{R}_t + \sigma_{u,R} u_{R,t}. \end{aligned} \tag{33}$$

We use data on government spending G_t to construct a measure of the consumption output ratio: $C_t/Y_t = 1 - G_t/Y_t$. We define cy_t^o as linearly detrended $100 \cdot \log(1 - G_t/Y_t)$. Because in our model $1 - G_t/Y_t = 1/g_t$, we obtain the additional measurement equation

$$cy_t^o = -100 \log g_* - 100 \hat{g}_t + \sigma_{u,c} u_{c,t}. \tag{34}$$

Thus, we are treating the exogenous process \hat{g}_t as observed in the estimation. Because the law-of-motion of \hat{g}_t is linear, the PLC structure of the empirical model is maintained.

4 PLC Model Solution

In this section we describe how the free coefficients ϑ in the PLC decision rules, as defined in (23), are determined. In a generic global approximation, the decision rules are assumed to be linear combinations of some known, flexible functions called basis functions, which are in turn nonlinear functions of states. The decision rules are parameterized by some coefficients that are the weights attached to each of the basis functions. There are two popular ways to choose these coefficients. In the collocation approach these coefficients are chosen such that the residuals, the errors in equilibrium conditions, are set to exactly zero at some carefully selected points in the state space. These points typically come from a grid that is constructed using a tensor product of grids for each state variable, which in turn are constructed using the roots of a set of complete polynomials.

It is well known that tensor product grids used to approximate the solution of nonlinear models suffer from the curse of dimensionality. Maliar and Maliar (2014, 2015) propose a series of techniques based on stochastic simulations to construct lower dimensional grids that represents the ergodic distribution of the model. However, these simulation-based methods require a time-consuming iterative procedure and, in general, there does not seem to be a guarantee for the convergence of the grid and the approximate solution.

For our application, where we need to solve the model with different parameters tens of thousands of times, neither the tensor grids that are used in the collocation approach, nor the iterative approach that uses the ergodic distribution seem feasible. Coleman, Lyon, Maliar, and Maliar (2018) propose the use of random and quasi-random grids on a fixed hypercube because they are easier and faster to construct but lack the dimensionality reduction. Smolyak grids (Krueger and Kubler (2004), Malin, Krueger, and Kubler (2011), Judd, Maliar, Maliar, and Valero (2014)) offer a balance in this trade-off, combining the advantages of a fixed and predetermined domain and the dimensionality reduction of sparse grid methods. We follow this approach and build a sparse Smolyak grid and minimize the sum of squared residuals over this grid to find the decision rule coefficients.⁵

Equilibrium Conditions. More formally, let us denote the generic equilibrium conditions as

$$\mathcal{H}[f_0(\cdot), \mathbb{X}] = 0, \quad \forall \mathbb{X} \in \mathcal{X}, \quad (35)$$

where $f_0(\mathbb{X})$ corresponds to the optimal decision rules. To simplify the notation, we dropped the vector of DSGE model parameters θ from the conditioning set. For instance, for the New Keynesian DSGE model (35) becomes

$$\mathbb{E}_t \left[\begin{array}{c} \mathcal{R}_c \left(\hat{y}_0(\mathbb{X}_t), \hat{c}_0(\mathbb{X}_t), \hat{\pi}_0(\mathbb{X}_t), \hat{R}_0(\mathbb{X}_t), \hat{y}_0(\mathbb{X}_{t+1}), \hat{c}_0(\mathbb{X}_t), \hat{\pi}_0(\mathbb{X}_{t+1}), \hat{R}_0(\mathbb{X}_{t+1}), \dots \right) \\ \mathcal{R}_\pi \left(\hat{y}_0(\mathbb{X}_t), \hat{c}_0(\mathbb{X}_t), \hat{\pi}_0(\mathbb{X}_t), \hat{R}_0(\mathbb{X}_t), \hat{y}_0(\mathbb{X}_{t+1}), \hat{c}_0(\mathbb{X}_t), \hat{\pi}_0(\mathbb{X}_{t+1}), \hat{R}_0(\mathbb{X}_{t+1}), \dots \right) \end{array} \right] = 0,$$

where we explained how we construct the decision rules $\hat{y}(\cdot)$, $\hat{c}(\cdot)$, $\hat{\pi}(\cdot)$ and $\hat{R}(\cdot)$ in Section 3.3. Expectations over \mathbb{X}_{t+1} can be evaluated by using the law of motion of the exogenous shocks in (16) and noting the the first three elements of \mathbb{X}_{t+1} , \hat{R}_t , 1, and \hat{y}_t , are known at time t .

⁵One interpretation of our approach is that we are using the sum squared residual over the Smolyak grid as a proxy for intergrating the squared residual function over the ergodic distribution. Heiss and Winschel (2008) show Monte Carlo results in the context of calculation of the likelihood function of a mixed logit model, which also involves evaluating an integral without a closed-form expression, that using a Smolyak grid provides a superior performance over simulation techniques.

Thus, the equilibrium conditions only depend on the two decision rules $y_0(\cdot)$ and $\pi_0(\cdot)$ and the current states \mathbb{X}_t , just like (35) requires.

We approximate $f_0(\mathbb{X})$ by PLC decision rules $g(\mathbb{X}; \vartheta) \in \mathcal{G}$, where ϑ contains the free coefficients that are necessary to characterize the PLC function, and \mathcal{G} is the set of all PLC functions. To determine ϑ , we minimize the norm of the vector-valued function $\mathcal{H}[g(\mathbb{X}; \vartheta), \mathbb{X}]$ over a set of M grid points \mathcal{S} obtained using a sparse Smolyak grid:

$$\vartheta = \arg \min_{\vartheta} \frac{1}{M} \sum_{\mathbb{X} \in \mathcal{S}} \|\mathcal{H}[g(\mathbb{X}; \vartheta), \mathbb{X}; \theta]\|^2.$$

Solution Grid. In constructing the grid \mathcal{S} we follow Judd, Maliar, Maliar, and Valero (2014). The Smolyak grid is a sparse grid defined on the interval $[-1, 1]$. To use it in an application, it has to be scaled so that it represents the space of \mathbb{X}_t . The scaling of the grid amounts to picking minimum and maximum values for each state variable. The extrema correspond to -1 and 1 in the original domain of the Smolyak grid, respectively. One of the properties of the Smolyak grid is it places grid points at the edges of the domain – at -1 and 1 . Thus, we recommend picking values for the scaling that are not too extreme in order to have some mass on both sides of the grid point.

In the context of the New Keynesian DSGE model, we proceed as follows. For the exogenous state variables in \mathbb{X}_t we linearly scale the grid so that it starts from the 10th percentile to the 90th percentile of the distribution of each state variable. For the endogenous state variable \hat{y}_{t-1} we use the same scaling as the exogenous state \hat{z}_t since we have verified that they have similar dispersion when simulating the model. Finally, for \hat{R}_{t-1} we use the the observed nominal interest rate data. Because we want to analyze the ZLB, the grid is scaled so that it's minimum value matches the ZLB with $\hat{R}_{t-1} = -\log(r_*\pi_*)$, which happens to be the 10th percentile. The maximum value is matched to the 90th percentile of R_t in the data. For this variable in particular we scale the grid so that the middle of the Smolyak grid coincides with the steady state at $\hat{R}_{t-1} = 0$.

Integration and Minimization. Expectations in the residual functions as in (11) are computed using the monomial integration rule M2 as in Judd, Maliar, and Maliar (2010). For a generic expression $\mathbb{E}_t[v(x_{t+1})]$ our implementation with four random variables requires computing $v(\cdot)$ at 33 nodes and taking a weighted average. In our experience this method produces results that are very similar to using Gauss-Hermite integration for each random variable. As an example, with 5 nodes per random variable, the latter approach would make it necessary to evaluate $v(\cdot)$ at 625 nodes and increase running time considerably.

To minimize the objective function, we utilize a gradient-based nonlinear solver with Jacobians evaluated analytically. As an initial guess for the solver, we use the log-linear approximation. Because the log-linear decision rules are a special case of the PLC decision rules – recall that we defined the model variables in log-deviations from the steady state – we can denote them by $g^{(0)}(\mathbb{X})$. We find the free coefficients ϑ_0 (with $\alpha_{11}^i = \alpha_{21}^i$ and $\alpha_{12}^i = \alpha_{22}^i$ for all $i = 1, \dots, k$) that generate the same decision rules and use this to initialize the minimization algorithm.

Interpretation of PLC Decision Rules. We offer two interpretations of the PLC decision rules. First, they can be viewed as an approximation to the optimal decision rules $f_0(\mathbb{X})$. In fact, our motivation for constructing PLC rules was that the decision rules computed in Aruoba, Cuba-Borda, and Schorfheide (2018) with Chebyshev polynomials for a New Keynesian DSGE model that is essentially identical to the model in Section 2, appeared to be almost piecewise-linear. While in any given model the PLC decision rules may or may not provide accurate approximations of the optimal decision rules, there is no sense in which the PLC rules become more accurate “asymptotically.”

Second, the PLC rules can be viewed as describing the behavior of boundedly rational agents. In principle, bounded rationality can take many forms. The basic notion is that decision making is constrained by agents’ ability to gather, retain, and process decision-relevant information. Boundedly rational agents may also be unable to solve complicated mathematical problems. Under this interpretation, the PLC rules can be viewed as more easily computable decisions that have the additional benefit of being linear, except when the constraint in the model becomes binding.

5 Particle Filters for PLC Models

Econometric inference is based on the state-space representation that comprises the nonlinear transition equation (30) and the linear measurement equation (32). In the formulas below we write y_t instead of y_t^o and represent the transition and measurement equations through the densities $p(s_t|s_{t-1}, \theta)$ and $p(y_t|s_t, \theta)$, respectively. We will subsequently use $Y_{t_1:t_2}$ and $S_{t_1:t_2}$ to denote the sequences y_{t_1}, \dots, y_{t_2} and s_{t_1}, \dots, s_{t_2} . Moreover, we use θ to denote the vector of model parameters.

The state-space model provides a joint density for the states s_t and the observations y_t :

$$p(Y_{1:T}, S_{1:T} | \theta) = \prod_{t=1}^T p(y_t | s_t, \theta) p(s_t | s_{t-1}, \theta). \quad (36)$$

Of particular interest are the sequence of estimates $p(s_t | Y_{1:t})$ of the state vector and the likelihood function, which is defined as

$$p(Y_{1:T} | \theta) = \prod_{t=1}^T p(y_t | Y_{1:t-1}, \theta) = \prod_{t=1}^T \int \int p(y_t | s_t, \theta) p(s_t | s_{t-1}, \theta) p(s_{t-1} | Y_{1:t-1}, \theta) ds_t ds_{t-1}. \quad (37)$$

These objects can be obtained from a nonlinear filter. We will use an algorithm that belongs to the class of sequential Monte Carlo filters to approximate $p(s_t | Y_{1:t}, \theta)$ and $p(y_t | Y_{1:t-1}, \theta)$, also known as particle filters.

In Section 5.1 we present a generic particle filtering algorithm that can be configured in many different ways by choosing a proposal density that mutates particles representing the state s_{t-1} into particles representing s_t . If the mutation is based on forward simulation of the state-transition equation, one obtains the bootstrap particle filter (BSPF) described in Section 5.2. Our key contribution is to derive the proposal density that is tailored toward the piecewise-linear canonical form of the state-space model and leads to the conditionally-optimal particle filter discussed in Section 5.3.

5.1 Generic Particle Filter

There exists a large literature on particle filters. Surveys and tutorials can be found, for instance, in Arulampalam, Maskell, Gordon, and Clapp (2002), Cappé, Godsill, and Moulines (2007), Doucet and Johansen (2011), and Creal (2012). Kantas, Doucet, Singh, Maciejowski, and Chopin (2014) discuss using particle filters in the context of estimating the parameters of state-space models. A particle filter represents the density $p(s_t | Y_{1:t}, \theta)$ through a swarm of particles $\{s_t^j, W_t^j\}_{j=1}^M$ with the property that posterior expectations $\mathbb{E}[h(s_t) | Y_{1:t}, \theta]$ can be approximated by Monte Carlo averages of the form

$$\frac{1}{M} \sum_{j=1}^M h(s_t^j) W_t^j.$$

The approximation typically holds in form of a Law of Large Numbers or a Central Limit Theorem. Textbook treatments of the statistical theory underlying particle filters can be found in Cappé, Moulines, and Ryden (2005), Liu (2001), and Del Moral (2013). Following the exposition in Herbst and Schorfheide (2015) the particle filter can be implemented using the following algorithm:

Algorithm 1 (Generic Particle Filter)

1. **Initialization.** Draw the initial particles from the distribution $s_0^j \stackrel{iid}{\sim} p(s_0|\theta)$ and set $W_0^j = 1$, $j = 1, \dots, M$.

2. **Recursion.** For $t = 1, \dots, T$:

(a) **Forecasting s_t .** Draw \tilde{s}_t^j from density $g_t(\tilde{s}_t^j | s_{t-1}^j, \theta)$ and define the importance weights

$$\omega_t^j = \frac{p(\tilde{s}_t^j | s_{t-1}^j, \theta)}{g_t(\tilde{s}_t^j | s_{t-1}^j, \theta)}. \quad (38)$$

(b) **Forecasting y_t .** Define the incremental weights

$$\tilde{w}_t^j = p(y_t | \tilde{s}_t^j, \theta) \omega_t^j. \quad (39)$$

The predictive density $p(y_t | Y_{1:t-1}, \theta)$ can be approximated by

$$\hat{p}(y_t | Y_{1:t-1}, \theta) = \frac{1}{M} \sum_{j=1}^M \tilde{w}_t^j W_{t-1}^j. \quad (40)$$

(c) Define the normalized weights

$$\tilde{W}_t^j = \frac{\tilde{w}_t^j W_{t-1}^j}{\frac{1}{M} \sum_{j=1}^M \tilde{w}_t^j W_{t-1}^j}. \quad (41)$$

(d) **Selection.** Resample the particles, for instance, via multinomial resampling. Let $\{s_t^j\}_{j=1}^M$ denote M iid draws from a multinomial distribution characterized by support points and weights $\{\tilde{s}_t^j, \tilde{W}_t^j\}$ and set $W_t^j = 1$ for $j = 1, \dots, M$. An approximation of $\mathbb{E}[h(s_t) | Y_{1:t}, \theta]$ is given by

$$\bar{h}_{t,M} = \frac{1}{M} \sum_{j=1}^M h(s_t^j) W_t^j. \quad (42)$$

3. Likelihood Approximation. *The approximation of the log-likelihood function is given by*

$$\ln \hat{p}(Y_{1:T}|\theta) = \sum_{t=1}^T \ln \left(\frac{1}{M} \sum_{j=1}^M \tilde{w}_t^j W_{t-1}^j \right). \quad (43)$$

The most important choice in the configuration of the algorithm is the proposal density $g_t(\tilde{s}_t|s_{t-1}^j, \theta)$. Different choices of the proposal density lead to different versions of the particle filter.

5.2 Bootstrap Particle Filter

The BSPF was originally proposed by Gordon and Salmond (1993). It uses the state-transition equation as the proposal density, that is, $g_t(\tilde{s}_t|s_{t-1}^j, \theta) = p(\tilde{s}_t|s_{t-1}^j, \theta)$. This choice is attractive because it is straightforward to implement the forecasting step by forward simulation of the transition equation and the importance weights simplify to $\omega_t^j = 1$. A well-known disadvantage is that the proposal distribution is blind and hence ignores information about s_t contained in the current observation y_t . This can lead to a large variance of the incremental weights $\tilde{\omega}_t^j$. This problem is exacerbated if the measurement error variance is small and $p(y_t|\tilde{s}_t^j)$ has thin tails or if the model is inappropriately parameterized or misspecified and therefore has difficulties predicting y_t one-step ahead. Because the BSPF has been used in the DSGE model literature (see Fernández-Villaverde and Rubio-Ramírez (2007), An and Schorfheide (2007) and Herbst and Schorfheide (2015)), we will include it as a benchmark.

5.3 Conditionally-Optimal Particle Filter

The proposal density for the COPF utilizes information in y_t with the goal of minimizing the variance of the incremental weights $\tilde{\omega}_t^j$. It is given by

$$g_t^*(\tilde{s}_t|s_{t-1}^j, \theta) = p(\tilde{s}_t|y_t, s_{t-1}^j, \theta) \propto p(y_t|\tilde{s}_t, \theta)p(\tilde{s}_t|s_{t-1}^j, \theta). \quad (44)$$

Combining the formula for $g_t^*(\tilde{s}_t|s_{t-1}^j, \theta)$ with the expressions for the importance weights ω_t^j in (38) and the incremental weights $\tilde{\omega}_t^j$ in (39), we obtain

$$\tilde{\omega}_t^j = \frac{p(y_t|\tilde{s}_t^j, \theta)p(\tilde{s}_t^j|s_{t-1}^j, \theta)}{p(\tilde{s}_t^j|y_t, s_{t-1}^j, \theta)} = p(y_t|s_{t-1}^j). \quad (45)$$

The second equality follows from Bayes Theorem. It can be shown that conditional on $\{s_{t-1}^j\}$ the proposal density $g_t^*(\tilde{s}_t|s_{t-1}^j, \theta)$ minimizes the variance of the incremental weight \tilde{w}_t^j in (39).

While direct sampling from the conditionally-optimal proposal density is elusive for most nonlinear state-space models, we can derive a convenient formula for the piecewise-linear state-transition equation (30). Note that conditional on s_{t-1} the current state s_t is determined by η_t . It turns out, that it is more convenient to derive a conditionally-optimal proposal density for η_t , denoted by $g_t^*(\eta_t|s_{t-1}^j, \theta)$.

In order to state the result, we have to define the following objects:

$$\begin{aligned}\nu_t^j(\cdot) &= y_t - A_0 - A_s(\Phi_0(\cdot) - \Phi_1(\cdot)s_{t-1}^j) \\ \bar{\eta}_t^j(\cdot) &= (\varsigma I + \Phi_\eta'(\cdot)A_s'\Sigma_u^{-1}A_s\Phi_\eta(\cdot))^{-1}\Phi_\eta'(\cdot)A_s'\Sigma_u^{-1}\nu_t^j(\cdot) \\ \bar{\Omega}(\cdot) &= \varsigma(\varsigma I + \Phi_\eta'(\cdot)A_s'\Sigma_u^{-1}A_s\Phi_\eta(\cdot))^{-1}.\end{aligned}\tag{46}$$

We use the argument (\cdot) to indicate that the expressions are obtained either based on $(\Phi_0(n), \Phi_1(n), \Phi_\eta(n))$ or $(\Phi_0(b), \Phi_1(b), \Phi_\eta(b))$. Here $\nu_t^j(\cdot)$ is the error made in forecasting y_t based on s_{t-1}^j . $\bar{\eta}_t^j(\cdot)$ and $\bar{\Omega}(\cdot)$ are the posterior mean vector and covariance matrix of $\eta_t|y_t, s_{t-1}^j$ absent any truncation, i.e., for $\zeta(s_{t-1}^j)$ being $+\infty$ or $-\infty$. Moreover, let

$$\begin{aligned}D_t^j(n) &= (2\pi)^{-n_y/2}|\Sigma_u|^{-1/2}|\varsigma I + \Phi_\eta(n)'A_s'\Sigma_u^{-1}A_s\Phi_\eta(n)|^{1/2} \\ &\quad \times \exp\left\{-\frac{1}{2}\nu_t^j(n)'(\varsigma\Sigma_u + A_s\Phi_\eta(n)\Phi_\eta'(n)A_s')^{-1}\nu_t^j(n)\right\} \\ &\quad \times \Phi_N((\zeta(s_{t-1}) - \bar{\eta}_{1,t}^j(n))/\sqrt{\bar{\Omega}_{11}(n)}), \\ D_t^j(b) &= (2\pi)^{-n_y/2}|\Sigma_u|^{-1/2}|\varsigma I + \Phi_\eta(b)'A_s'\Sigma_u^{-1}A_s\Phi_\eta(b)|^{1/2} \\ &\quad \times \exp\left\{-\frac{1}{2}\nu_t^j(b)'(\varsigma\Sigma_u + A_s\Phi_\eta(b)\Phi_\eta'(b)A_s')^{-1}\nu_t^j(b)\right\} \\ &\quad \left(1 - \Phi_N((\zeta(s_{t-1}) - \bar{\eta}_{1,t}^j(b))/\sqrt{\bar{\Omega}_{11}(b)})\right).\end{aligned}\tag{47}$$

It can be shown that $p(y_t|s_{t-1}^j) = D_t^j(n) + D_t^j(b)$.

The characterization of the conditionally-optimal proposal density is summarized in Proposition 1. A proof of the proposition is provided in the Online Appendix.

Proposition 1 *Suppose the state-transition equation is given by (30), $\eta_t \sim N(0, I)$, $\eta_{1,t}$ is a scalar, and the measurement equation is given by (32), then draws from the conditionally-*

optimal proposal densities $g_t^*(\tilde{s}_t|s_{t-1}^j, \theta)$, $j = 1, \dots, M$, defined in (44) can be generated as follows:

1. Let

$$\xi_t^j = \begin{cases} \text{'n'} & \text{with prob. } \lambda_t^j \\ \text{'b'} & \text{with prob. } 1 - \lambda_t^j \end{cases}, \quad \text{where } \lambda_t^j = \frac{D_t^j(n)}{D_t^j(n) + D_t^j(b)}.$$

2. If $\xi_t^j = \text{'n'}$ then generate η_t from the distribution

$$\eta_{1,t}^j \sim N(\bar{\eta}_{1,t}^j(n), \bar{\Omega}_{11}(n)) \mathbb{I}\{\eta_{1,t}^j \leq \zeta(s_{t-1}^j)\}, \quad \eta_{2,t}^j | \eta_{1,t}^j \sim N(\bar{\eta}_{2|1}^j(n, \eta_{1,t}^j), \bar{\Omega}_{2|1}(n)) \quad (48)$$

and let

$$\tilde{s}_t^j = \Phi_0(n) + \Phi_1(n)s_{t-1}^j + \eta_t^j.$$

If $\xi_t^j = \text{'b'}$ then generate η_t^j from the distribution

$$\eta_{1,t}^j \sim N(\bar{\eta}_1^j(b), \bar{\Omega}_{11}(b)) \mathbb{I}\{\eta_{1,t}^j > \zeta(s_{t-1}^j)\}, \quad \eta_{2,t}^j | \eta_{1,t}^j \sim N(\bar{\eta}_{2|1}^j(b, \eta_{1,t}^j), \bar{\Omega}_{2|1}(b)) \quad (49)$$

and

$$\tilde{s}_t^j = \Phi_0(b) + \Phi_1(b)s_{t-1}^j + \eta_t^j.$$

3. The incremental particle weight is $\tilde{\omega}_t^j = D(n) + D(b)$. ■

Vanishing Measurement Errors. It is instructive to examine what happens as $\varsigma \rightarrow \infty$. For the conditional density of $y_t|s_{t-1}$ to be nonsingular in the limit, it has to be the case that the number of rotated structural innovations is at least n_y . Formally, the covariance matrices $A_s \Phi_\eta(\cdot) \Phi'_\eta(\cdot) A'_s$ have to be non-singular. For expositional purposes, we consider the case in which $n_y = n_\eta$ and $(A_s \Phi_\eta(\cdot))$ are invertible $n_y \times n_y$ matrices. This means, ignoring the truncation, under the invertibility assumption we can solve for the innovations η_t as a function of (y_t, s_{t-1}^j) :

$$\eta_{t*}^j(\cdot) = (A_s \Phi_\eta(\cdot))^{-1}(y_t - A_s(\Phi_0(\cdot) + \Phi_1(\cdot)s_{t-1}^j)).$$

Now consider what happens if we let the measurement error variance converge to zero. First, the expressions in (46) remain well defined in the limit:

$$\bar{\eta}_t^j \rightarrow \eta_{t*}^j(\cdot), \quad \bar{\Omega}(\cdot) \rightarrow 0.$$

The posterior variance converges to zero and the posterior mean converges to the innovation $\eta_{t*}^j(\cdot)$ that generates the observed y_t conditional on s_{t-1}^j . For the limit behavior on $D_t^j(n)$ the crucial term is

$$\lim_{\varsigma \rightarrow 0} \Phi_N((\zeta(s_{t-1}^j) - \bar{\eta}_{1,t}^j(n))/\sqrt{\bar{\Omega}_{11}(n)}) = \begin{cases} 1 & \text{if } \zeta(s_{t-1}^j) - \bar{\eta}_{1,t}^j(n) \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

This term measures whether it is possible to explain y_t using the (n) coefficients, accounting for the fact that the n regime is only active if $\bar{\eta}_{1,t}^j(n) \leq \zeta(s_{t-1}^j)$. A similar analysis can be conducted for the term $D_t^j(b)$. Thus, for each particle j , there are four possible cases (ignoring equalities):

$$\text{Case 1 : } \bar{\eta}_{1,t}^j(n) < \zeta(s_{t-1}^j), \quad \bar{\eta}_{1,t}^j(b) < \zeta(s_{t-1}^j)$$

$$\text{Case 2 : } \bar{\eta}_{1,t}^j(n) \geq \zeta(s_{t-1}^j), \quad \bar{\eta}_{1,t}^j(b) \geq \zeta(s_{t-1}^j)$$

$$\text{Case 3 : } \bar{\eta}_{1,t}^j(n) < \zeta(s_{t-1}^j), \quad \bar{\eta}_{1,t}^j(b) \geq \zeta(s_{t-1}^j)$$

$$\text{Case 4 : } \bar{\eta}_{1,t}^j(n) \geq \zeta(s_{t-1}^j), \quad \bar{\eta}_{1,t}^j(b) < \zeta(s_{t-1}^j).$$

In Case 1 and $D_t^j(b) = 0$ and $\lambda_t^j = 1$. Here only the (n) decision rules can rationalize the data conditional on s_{t-1}^j . Case 2 is the opposite: $D_t^j(n) = 0$, $\lambda_t^j = 0$, and only the (b) decision rules can rationalize the data. Under Case (3) both $D_t^j(n)$ and $D_t^j(b)$ are strictly positive, $0 < \lambda_t^j < 1$, and both decision rules could explain the data. Finally, in Case 4 y_t is inconsistent with s_{t-1}^j and none of the decision rules can explain the data. If each $j = 1, \dots, M$ falls into Case 4, then the particle-filter based likelihood approximation for this particular parameterization of the DSGE model will be zero. Note that if the measurement error variance is strictly greater than zero, (potentially very large) measurement errors could also rationalize the data under Case 4.

The previous calculations highlight that, unlike for the BSPF, the weights of the COPF do not degenerate if one decreases the measurement error variance. In this case, if $A_s \Phi_\eta(\cdot)$ is a square matrix, the COPF specializes to the inversion filter that solves for the innovations as a function of y_t and s_{t-1}^j . Because our model is piecewise-linear, this inversion may have one, two, or no solution.

Perfectly Observed Regimes. Our ZLB application has the special feature that the observation y_t identifies the regime. Let $y_t = [y'_{1,t}, y_{2,t}]'$ where $y_{2,t}$ corresponds to the nominal interest rate. Suppose the ZLB is binding in the b regime and non-binding in the n regime. Then, the ex-post regime probability λ^j is independent of s_{t-1}^j and given by the indicator

function.

$$\lambda^j = \mathbb{I}\{y_{2,t} > c\}. \quad (50)$$

While the distribution of y_t is continuous in the n regime, for the binding regime the continuous part of the distribution concentrates in the lower-dimensional subspace defined by $y_{2,t} = c$. Thus, the formulas for $D_t^j(b)$ in (47) and the moments of the distribution of η_t^j in Proposition 1 in the b regime have to be adjusted to account for the reduced dimensionality of the continuous part of the y_t distribution. Further details are provided in the Online Appendix.

6 Numerical Illustrations

We now illustrate the proposed filtering method based on data simulated from the DSGE model of Section 2. In Section 6.1 we compare the accuracy of log-likelihood approximations obtained with the proposed COPF and a naive bootstrap particle filter (BSPF). In Section 6.2 we embed the particle-filter approximations of the DSGE model likelihood function into an otherwise standard random walk Metropolis-Hastings (RWMH) algorithm and assess the improvements attainable with the COPF.

6.1 Accuracy of Likelihood Approximation

We simulate a sample of 1,000 observations from the DSGE model loosely parameterized based on the empirical estimates reported in Section 7 and select a subsample of $T = 140$ observations which is the sample size in the empirical application in Section 7. In order to increase the likelihood of hitting the ZLB in the simulation, we lower the target inflation rate π_* to an annualized rate of 0.5%. In the selected subsample the ZLB binds 22% of the periods which is roughly consistent with our actual sample. The parameter values are summarized in the second column of Table 1.

We begin by comparing the accuracy of the particle-filter-based likelihood approximation at the “true” parameter values that were used to generate the data. For the COPF we set $M = 250$ and for the BSPF we consider $M = 1,000$ and $M = 10,000$. The average run times for a likelihood evaluation on a single core are 2.6, 1.8, and 18.3 seconds, respectively. While the data are generated without measurement errors, the likelihood evaluation imposes non-zero measurement errors to facilitate the application of the BSPF. The measurement error

Table 1: DGP and Prior

Parameter	DGP	Prior Distribution				
		Density	P(1)	P(2)	HPD Low	HPD High
τ	2.00	\mathcal{G}	2.00	0.20	1.67	2.32
κ	0.13	\mathcal{B}	0.10	0.05	0.02	0.17
ψ_1	2.60			fixed at 2.60		
ψ_2	0.98			fixed at 0.98		
ρ_R	0.80	\mathcal{B}	0.80	0.10	0.65	0.96
ρ_g	0.97	\mathcal{B}	0.80	0.10	0.65	0.96
ρ_d	0.91	\mathcal{B}	0.80	0.10	0.65	0.96
ρ_z	0.37	\mathcal{B}	0.40	0.20	0.08	0.73
σ_R	0.0019	$\sqrt{\mathcal{IG}}$	0.005	4.00	0.003	0.010
σ_g	0.0025	$\sqrt{\mathcal{IG}}$	0.005	4.00	0.003	0.010
σ_d	0.017	$\sqrt{\mathcal{IG}}$	0.01	4.00	0.005	0.020
σ_z	0.0058	$\sqrt{\mathcal{IG}}$	0.01	4.00	0.005	0.020
η	0.72			fixed at 0.72		
ν	0.10			fixed at 0.10		
χ_H	1.00			fixed at 1.00		
g^*	1.27	\mathcal{G}	1.20	0.20	0.88	1.63
rAnet	0.22	\mathcal{G}	1.00	0.40	0.39	1.64
gamQnet	0.33	\mathcal{N}	0.50	0.25	0.09	0.91
piAnet	0.50	\mathcal{N}	2.50	1.00	0.87	4.14

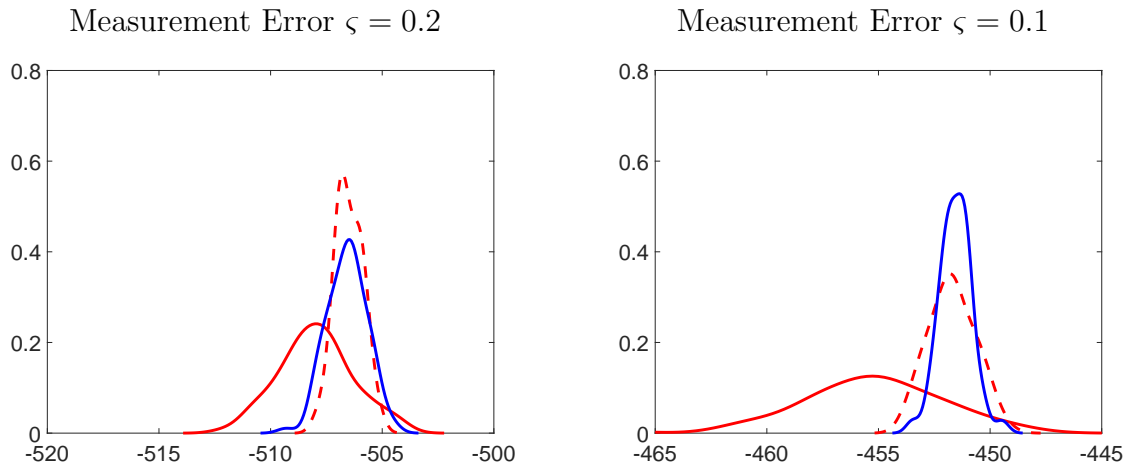
Notes: \mathcal{G} is Gamma distribution; \mathcal{B} is Beta distribution; \mathcal{IG} is Inverse Gamma distribution; and \mathcal{N} is Normal distribution. P(1) and P(2) are mean and standard deviations for Beta, Gamma, and Normal distributions. The \mathcal{IG} distribution is parameterized as scaled inverse χ^2 distribution with density $p(\sigma^2|s^2, \nu) \propto (\sigma^2)^{-\nu/2-1} \exp[-\nu s^2/(2\sigma^2)]$, where P(1) is $\sqrt{s^2}$ and P(2) is ν . The density of σ is obtained by the change of variables $\sigma = \sqrt{\sigma^2}$. HPD(Low,High) refers to the boundaries of 90% highest prior density intervals. We use the following parameter transformations: $\beta = \exp\{-\text{rAnet}/400\}$, $\gamma = \exp\{\text{gamQnet}/100\}$, and $\pi_* = \exp\{\text{piAnet}/400\}$.

captures the discrepancy between model and data. The smaller the measurement errors, the larger the penalty for deviations between model-predicted and actual observations. Thus, small measurement errors can be provided as a stress test for the filter in case the model (or its parameterization) is at odds with the data.

The likelihood evaluation is conditional on a vector of initial states which we take to be the “true” initial states associated with the simulated observations. When we subsequently estimate the model, we treat the initial states as unknown parameters and include them into the vector θ . The prior for the initial states captures the uncertainty about the level of the states in period $t = 0$.

As discussed in detailed in Section 5, particle filters are stochastic algorithms. We run

Figure 1: Density of Log-Likelihood Approximations



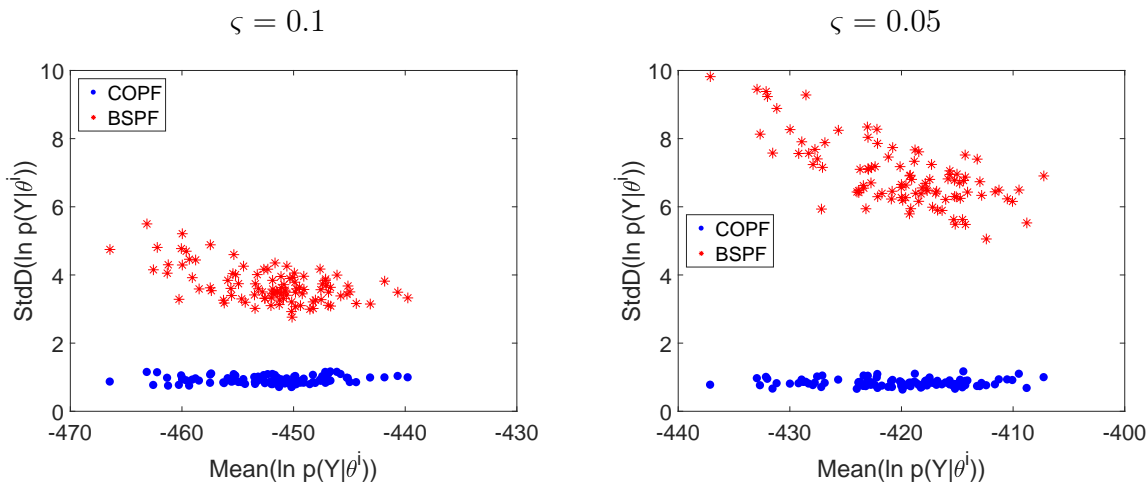
Notes: Density plots are based on $N_{run} = 100$ runs of the BSPF ($M = 1,000$ is red solid and $M = 10,000$ is red dashed) and COPF ($M = 250$, blue solid), respectively.

the COPF and the BSPF $N_{run} = 100$ times, respectively. Holding the parameter value θ constant, each time we run the filter, we obtain a slightly different log-likelihood approximation $\log \hat{p}(Y|\theta)$. From repeated runs of the filter we construct kernel estimates of the likelihood, which are depicted in Figure 1. We consider two values for the scaling parameter for the measurement errors, ς : 0.2 and 0.1.

Because the BSPF ignores the information in y_t when generating the proposal draws \tilde{s}_t^j , the variance of the particle weights increases as ς falls. This, in turn, translates into an increase in the variance of the log-likelihood approximation, which is clearly visible by comparing the solid red densities across panels for the two different values of ς . The precision of the COPF, on the other hand, increases as ς falls. When $\varsigma \rightarrow 0$, it is possible to uniquely determine the innovations η_t^j conditional on s_{t-1}^j during non-ZLB periods because the number of observables equals the number of shocks. In order to achieve a similar accuracy as the COPF, the number of particles for the BSPF has to be increased to $M = 10,000$, which implies that the BSPF likelihood evaluation takes roughly seven times as long as the COPF likelihood evaluation.

Figure 2 depicts standard deviations of log-likelihood approximations as a function of the mean log-likelihood value across $N_{run} = 100$ runs of the filter for $\varsigma = 0.1$ and $\varsigma = 0.05$. Each dot (or asterisk) in the two scatter plots corresponds to a parameter value θ^i generated from the particle Markov chain Monte Carlo (PMCMC) algorithm in Section 6.2 below.

Figure 2: Comparison of Log-Likelihood Approximations



Notes: Standard deviations of log-likelihood approximations are based on $N_{run} = 100$ runs of the two filters. Each dot (or asterisk) corresponds to a particular θ^i . We are using $M = 1,000$ particles for the BSPF (red asterisks) and $M = 250$ particles for the COPF (blue dots).

Two important findings emerge. First, as we have seen already from Figure 1, the COPF likelihood approximation is less dispersed than the BSPF approximation. The accuracy gain from the conditionally-optimal proposal density increases as the measurement error variance decreases, because the BSPF performance deteriorates. Second, while the accuracy of the COPF is independent of the log-likelihood value associated with the posterior draw θ^i , the accuracy of the BSPF approximation deteriorates the further the θ^i draw is in the tails of the posterior distribution.

6.2 Particle MCMC

We now embed the particle filter likelihood approximations in a standard single-block RWMH algorithm. A textbook treatment of this algorithm can be found in Herbst and Schorfheide (2015). The particle RWMH algorithm operates on an enlarged probability space that includes all the random variables that are generated when running the particle filter. Write the particle filter approximation of the likelihood function as

$$\hat{p}(Y|\theta) = g(Y|\theta, U),$$

where U represents the random variables associated with the particle filter implementation. The sampler operates on an enlarged probability space that includes U and generates draws

from the joint posterior distribution

$$p(\theta, U|Y) \propto g(Y|\theta, U)p(\theta)p(U).$$

Because particle filter approximations of likelihood functions are unbiased we deduce that the marginal posterior of θ is identical to the posterior that obtains if the exact likelihood is available:

$$p(\theta|Y) = \int p(\theta, U|Y)dU \propto \int g(Y|\theta, U)p(\theta)p(U)dU = p(Y|\theta)p(\theta).$$

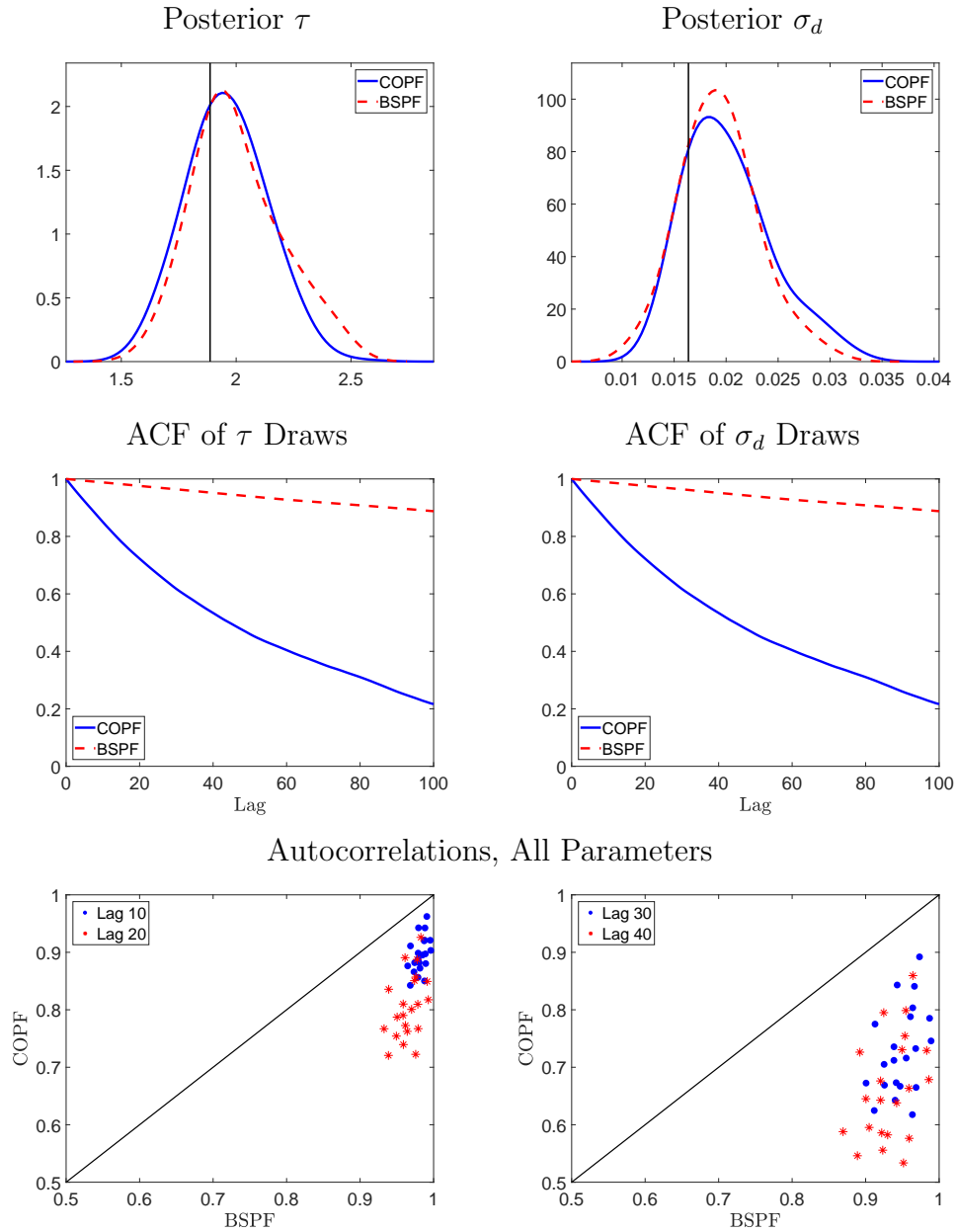
A formal analysis of PMCMC algorithms is provided in Andrieu, Doucet, and Holenstein (2010). Unfortunately, there is no free lunch. The use of an enlarged probability space leads to an increase in the persistence of the Markov chain generated by the posterior sampler. This in turn, leads to less precise Monte Carlo approximations of posterior moments. In a nutshell, the higher the variance of the likelihood approximation, the larger the persistence of the resulting Markov chain. If the variance is too high, sampling from the posterior distribution will fail. The use of an accurate particle filter such as the COPF can alleviate this problem.

The RWMH algorithm requires a covariance matrix for the proposal distribution (we use a multivariate normal distribution) that is constructed as follows. We start from a log-linearized version of the DSGE model that ignores the ZLB constraint. For this version, the likelihood function can be evaluated with the Kalman filter (KF). We conduct a preliminary MCMC run based on a diagonal covariance matrix with scaled prior variances on the diagonal. Based on the output of the preliminary run, we compute the posterior variance covariance matrix for the linearized model that ignores the ZLB constraint, denoted by \hat{V}_θ . A version of this matrix that is scaled by c is used for all subsequent calculations.

We set the scale of the measurement error variance to $\varsigma = 0.1$. Thus, according to the left panel of Figure 2, the standard deviation of the log-likelihood approximation of the COPF is around 1 whereas the standard deviation for the COPF ranges from 3 to 5. We keep the number of particles for the BSPF at $M = 1,000$ and reduce the number of particles to $M = 150$ for the COPF to lower the runtime. The scale factor for the proposal covariance matrix of the RWMH algorithm is set to $c = 0.1$. This leads to an overall run time of 7.7 hours for the COPF-RWMH algorithm on a single core with an acceptance rate of 30%.⁶

⁶Herbst and Schorfheide (2015) documented that for the estimation of small-scale linearized DSGE models an acceptance rate between 15% and 30% is associated with the most accurate Monte Carlo approx-

Figure 3: Posterior Draws: Density and Autocorrelation



Notes: BSPF configuration: number of particles $M = 1,000$, ME scale $\varsigma = 0.1$, proposal covariance scale $c = 0.1$, run time is 10:10:27 (hh:mm:ss). COPF configuration: number of particles $M = 150$, ME scale $\varsigma = 0.1$, proposal covariance scale $c = 0.1$, run time is 7:39:20 (hh:mm:ss). Top row: kernel density estimates of posterior distributions based on MCMC output. Vertical lines indicate true parameter values. Center row: autocorrelation functions of posterior draws based on COPF and BSPF. Bottom row: scatter plots of autocorrelations BSPF vs. COPF for various lags. Solid line is 45 degree line.

Replacing the COPF with the commonly-used BSPF increases the run time of the RWMH

iminations of posterior means.

algorithm to 10.2 hours and lowers the acceptance rate to 4%, drastically increasing the persistence of the Markov chain. Under some regularity conditions, the sequence of posterior draws generated from a RWMH algorithm satisfies a Central Limit Theorem (CLT) for dependent processes. The numerical accuracy of the Monte Carlo approximation of posterior means depends on the long-run covariance matrix of the sequence of parameter draws θ^i , $i = 1, \dots, N$. The larger the autocorrelation of these draws, the less precise the Monte Carlo approximation.

The first row of Figure 3 compares posterior densities constructed from the output of the COPF- and BSPF-RWHM algorithms for two representative parameters: τ and σ_d . The posterior densities look very similar and both peak near the “true” parameter values depicted by the solid vertical lines. The second row shows the autocorrelation functions for the τ^i and σ_d^i sequences. Here stark differences emerge. While under the BSPF-based sampler the autocorrelation at lag 100 is still around 0.9, it is approximately 0.2 for draws from the COPF-RWHM.

The last row of the figure compares autocorrelations for lags 10, 20, 30, and 40 for all estimated parameters. The solid lines are 45-degree lines. The two panels show that the COPF is able to reduce the autocorrelation for all estimated parameters and hence it drastically improves the performance of the MCMC algorithm. Because the COPF can deliver accurate likelihood approximations with a relatively small number of particles, $M = 150$ in our numerical illustration, it is possible to accurately estimate a DSGE model with occasionally-binding constraint without supercomputing capabilities in a relatively short amount of time.

7 Empirical Application

We now estimate the small-scale New Keynesian DSGE model based on quarterly U.S. data and conduct a fiscal policy experiment. The estimation results are summarized in Section 7.1 and the fiscal policy analysis appears in Section 7.2.

7.1 Estimation

The DSGE model is estimated based on data on GDP growth (q-o-q %), the log consumption-GDP ratio (scaled by 100), GDP deflator inflation (annualized %), and nominal interest rates

(annualized %) with data from 1984:Q1 to 2018:Q4. The data for the estimation was extracted from the FRB St. Louis FRED database (vintage 2019-10-30). Output growth is defined as real gross domestic product (GDPC1) growth converted into per capita terms. Our measure of population is Civilian Noninstitutional Population (CNP16OV). We compute population growth rates as log differences and apply an eight-quarter backward-looking moving average filter to the growth rates to smooth out abrupt changes in the population growth series. In constructing a measure of the consumption-GDP ratio, we define consumption as the difference between output and government spending. Thus, our measure of consumption includes investment and net exports. Government spending is constructed as real government consumption expenditures and gross investment (GCEC1). We remove a linear trend from the log consumption-GDP ratio to correct for different time trends in the price deflators of the GDP components. Inflation is defined as the log difference in the GDP deflator (GDPDEF) and the interest rate is the average effective federal funds rate (FEDFUNDS) within each quarter. During the period 2009:Q1 to 2015:Q4 when the effective federal funds rate was between 0 and 25 basis points, we set the interest rate exactly equal to zero and regard the ZLB as binding.

The prior distribution used for the estimation is identical to the one in Table 1. We absorb the initial values of the latent state variables into the parameter vector and specify prior distributions over the initial states; see Table A-1 in the Online Appendix.⁷ We fix a number of parameters prior to estimation. Because our sample does not include observations on labor market variables we fix the Frisch labor supply elasticity. Based on Ríos-Rull, Schorfheide, Fuentes-Albero, Kryshko, and Santaaulalia-Llopis (2012), who provide a detailed discussion of parameter values that are appropriate for DSGE models of U.S. data, we set $\eta = 0.72$. The parameter ν , which captures the elasticity of substitution between intermediate goods, is not separately identifiable from the slope of the Phillips curve κ which in turn determines the adjustment cost parameter ϕ . We set $\nu = 0.1$ which generates a markup of 10%. We fix the preference parameter at $\chi_H = 1$. It determines steady state hours worked and is neither relevant for the model dynamics nor identifiable based on our observables. We also fix the monetary policy coefficients ψ_1 and ψ_2 at 2.60 and 0.98, respectively, which are values estimated in Aruoba and Schorfheide (2016).

We start out by estimating the log-linearized version of the DSGE model that ignores the ZLB constraint, setting the measurement error variances of the state-space model to zero.

⁷This approach has the advantage that uncertainty about the initial state does not add to the variability of the particle-filter-based likelihood approximation conditional on a parameter θ^i .

Table 2: Posterior Distribution (PLC / COPF)

Parameter	Mean	MAP	HPD Low	HPD High
τ	2.08	2.20	1.75	2.39
κ	0.12	0.17	0.08	0.17
ρ_R	0.82	0.82	0.79	0.85
ρ_g	0.982	0.987	0.967	0.997
ρ_d	0.95	0.97	0.93	0.97
ρ_z	0.32	0.30	0.18	0.47
σ_R	0.0018	0.0019	0.0016	0.0021
σ_g	0.0026	0.0024	0.0023	0.0028
σ_d	0.0270	0.0361	0.0170	0.0370
σ_z	0.0060	0.0063	0.0052	0.0068
g^*	1.27	1.28	1.25	1.29
rAnet	0.64	0.83	0.27	1.04
gamQnet	0.36	0.32	0.23	0.48
piAnet	2.94	3.34	2.48	3.36

Notes: The estimation period is 1984:Q1 to 2018:Q4. The following parameter are fixed during the estimation: $\psi_1 = 2.6$, $\psi_2 = 0.98$, $\eta = 0.72$, $\nu = 0.10$, and $\chi_H = 1.00$. We use the following parameter transformations: $\beta = \exp\{-\text{rAnet}/400\}$, $\gamma = \exp\{\text{gamQnet}/100\}$, and $\pi_* = \exp\{\text{piAnet}/400\}$. MAP refers to the maximum posterior probability estimate. HPD(Low,High) refers to the boundaries of 90% highest posterior density intervals. COPF configuration: number of particles $M = 150$, ME scale $\varsigma = 0.001$, proposal covariance scale $c = 0.2$, $N = 55,000$ draws (drop first 10%), acceptance rate is 25%, run time is 13:27:23 (hh:mm:ss).

Draws from the posterior are generated by a single-block RWMH algorithm. In an initial run, we use a diagonal matrix with the prior variances to configure the covariance matrix of the proposal distribution. In the main run, we use the estimated posterior covariance matrix \hat{V} from the initial run to construct a proposal covariance matrix Σ with the scaling factor $c = 0.2$. We generate $N = 110,000$ draws, discarding the first 10,000.

To estimate the DSGE model with PLC decision rules, we generate $N = 55,000$ draws from the posterior distribution of θ using the particle RWMH algorithm, discarding the first 5,000 draws. We use a scaling of $c = 0.2$ for the covariance matrix of the proposal distribution. The scale factor for the measurement error variances is set to $\varsigma = 0.001$. The likelihood function is approximated using the COPF with $M = 150$ particles. We assume that conditional on observing the nominal interest rate it is known whether the ZLB is binding; see (50). The resulting acceptance rate of the particle RWMH algorithm is 25% and the run time is 13.5 hours on a single core, which comes to about 0.9 seconds per draw.

The parameter estimates are summarized in Table 2. The table reports lower and upper

endpoints of highest posterior density (HPD) sets as well as posterior means.⁸ The parameter estimates are similar to the ones reported elsewhere in the literature for variants of the small-scale New Keynesian DSGE model. The estimated slope of the Phillips is $\hat{\kappa} = 0.12$. The government spending shock is close to a unit-root process ($\hat{\rho}_g = 0.982$) and the estimated autoregressive parameter of the discount factor shock is $\hat{\rho}_d = 0.95$. Thus, innovations to these processes will have a long-lasting effect. One of the outputs of the estimation is the set of filtered values for the exogenous variables. We use these explicitly in our policy experiment and comment on them in the next section.

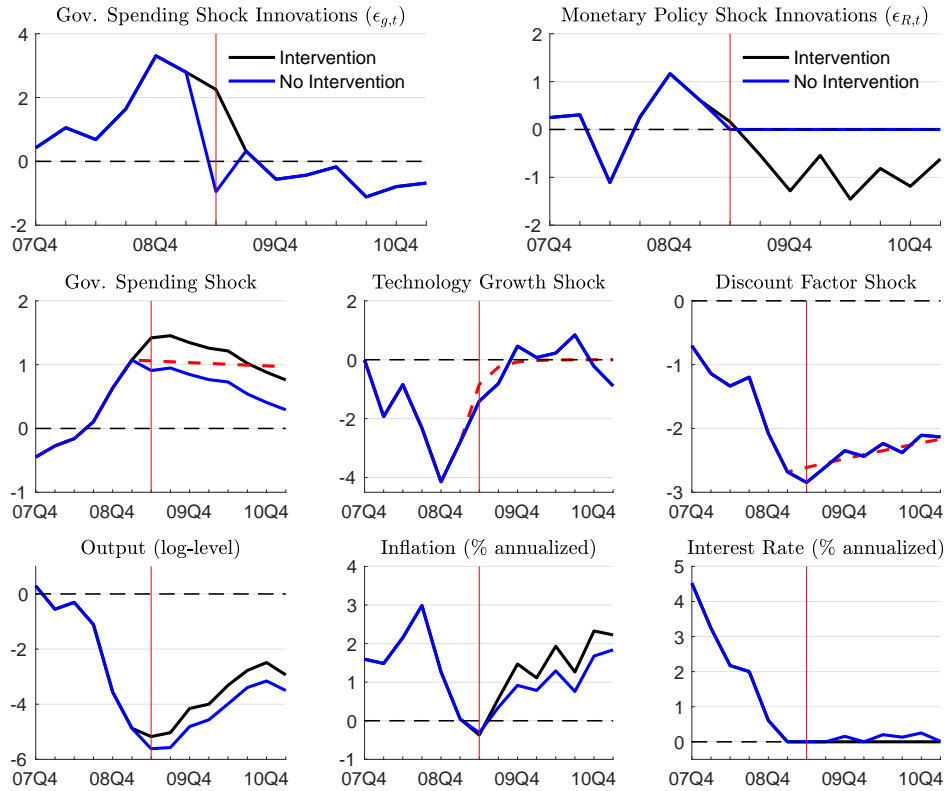
7.2 Fiscal Policy Analysis

The recent literature has emphasized that the effects of expansionary fiscal policies on output may be larger if the economy is at or near the ZLB. In the absence of the ZLB, a typical interest rate feedback rule implies that the central bank raises nominal interest rates in response to rising inflation and output caused by an increase in government spending. This monetary contraction raises the real interest rate, reduces private consumption, and overall dampens the stimulating effect of the fiscal expansion. If, however, the economy remains at the ZLB despite the expansionary fiscal policy, then the increase in inflation that results from the fiscal expansion reduces the real rate. In turn, current-period demand is stimulated, amplifying the positive effect on output. We will use our model to provide a quantitative assessment of this effect.

ARRA Government Spending. Because our model solution is nonlinear, the effect of a fiscal intervention depends on the initial condition and the size of the intervention. We use the Great Recession and the subsequent period in the U.S. as our laboratory and consider a fiscal intervention that is calibrated to a portion of the ARRA of February 2009 as we explain below. Our analysis is conducted from an ex-post perspective, where we extract the historical shocks that make our model match the realized U.S. data, which include both a fiscal and monetary intervention, and ask what would have happened if one or both of the policy interventions were not implemented.

⁸We compare these estimates with those obtained from a linearized model using the Kalman Filter and data that excludes the ZLB episode. The most noteworthy differences are in ρ_d and σ_d , both of which need to be larger when the ZLB episode is used in order to deliver large (negative) and persistent shocks that take the economy to the ZLB. We also find that rAnet and gamQnet estimates are somewhat smaller in the full sample, both of which are consistent with related results in the literature.

Figure 4: Ex-Post Policy Analysis



Notes: Vertical red line corresponds to 2009:Q2, which is the date of the ARRA intervention. Intervention (black) versus no-intervention paths (blue). Along the no-intervention path we set monetary policy shocks to zero and lower the innovation to the government spending shock in 2009:Q2 by the size of the ARRA intervention. Red-dashed lines represent paths in the absence of exogenous shock innovations from 2009:Q2 onwards. The level processes in the second row of the figure are standardized by the unconditional standard deviations of the corresponding AR(1) processes. Inflation and the interest rate are expressed in terms of annualized percentage rates.

ARRA of February 2009 consisted of a combination of tax cuts and benefits; entitlement programs; and funding for federal contracts, grants, and loans. We focus on the third component because it can be interpreted as an increase in g_t . We model the ARRA spending as a one-period positive shock of δ^{ARRA} to the demand shock process, where we calibrated $\delta^{ARRA} = 0.0077$ using data on the disbursement of ARRA funds, as we explain in the Online Appendix. This one-time shock is roughly $3.2\sigma_g$ and since \hat{g}_t is highly serially correlated, the effect of the shock will slowly decay over time. We assume that the ARRA innovation to government spending took place in the second quarter of 2009.

Ex-Post Policy Analysis. We use the COPF to obtain estimates of the exogenous shock

processes for the period 2009:Q2-2011:Q1. The subsequent results are based on the maximum posterior probability (MAP) estimator of the DSGE model parameters. The panels in the first two rows of Figure 4 show the filtered monetary policy and government spending innovations and the levels of the government spending, technology growth, and discount factor shock processes. Recall that in the model $\epsilon_{R,t}$ and $\epsilon_{g,t}$ are $N(0, 1)$ random variables. The level processes in the second row of the figure are standardized by the unconditional standard deviations of the corresponding AR(1) processes. For the government spending and the monetary policy shocks we distinguish between intervention and no-intervention paths.

According to the estimated model, the drop in output during the Great Recession is generated by drastic falls in the technology growth and discount factor shock processes. Because of the stylized structure of the DSGE model, these two shocks also absorb the contribution of financial shocks and financial accelerator effects. While the technology shock is not very persistent ($\hat{\rho}_z = 0.32$) and reverts back to zero by the end of 2009, the mean reversion of the discount factor shock is very slow and it remains below 2 standard deviations until the end of 2010. Meanwhile the government spending process is positive, indicating that fiscal policy started to become expansionary (relative to the historical average) in 2008. The filtered monetary policy innovations $\hat{\epsilon}_{R,t|t}$ turn out to be negative past 2009:Q2, which captures an effort by the Federal Reserve to keep the policy rate lower than what the policy rule implies. This is how our model that abstracts from explicitly modeling unconventional monetary policies implemented in this period (quantitative easing, forward guidance) handles the existence of these policies in the data.

Because the actual path of the government spending shock already contains the effect of fiscal expansion due to ARRA, we compute the counterfactual path by subtracting the effect of ARRA from the filtered demand shock $\hat{g}_{t|t}$ using

$$\hat{g}_{t|t}^C = \hat{g}_{t|t} - \rho_g^{t-T_*} \delta^{ARRA} \text{ for } t = T_*, T_* + 1, \dots, T_* + 7. \quad (51)$$

where T_* corresponds to 2009:Q2, the period the ARRA intervention is implemented in. The magnitude of the ARRA intervention is reflected in the difference between the intervention (black) and no-intervention (blue) government spending innovation depicted in the top left panel of Figure 4. The ARRA intervention shifts \hat{g}_t persistently downward as shown in the center left panel of the figure. To measure the effect of the combined fiscal and monetary policy, we set the counterfactual monetary policy shocks to zero.

The main finding is depicted in the bottom panels. The ex-post effect of the intervention

Table 3: Ex-Post Multipliers

Intervention	1Q	4Q	8Q
Fiscal Only	0.79	0.76	0.73
Fiscal + Monetary	0.76	0.81	0.86

Notes: Ex-post analysis is conditional on filtered 2009:Q2 - 2011:Q1 shocks. The pure fiscal multipliers are obtained by setting the monetary policy shocks to zero, whereas the combined fiscal and monetary multipliers are based on leaving the monetary policy shocks at their filtered values.

is defined as $X^o - X^C$ where X^o is the observed value of a generic variable and X^C is the counterfactual path along which the policy intervention is removed.⁹ In the figure, the ex-post effect is given by the gap between the no intervention and the intervention path. In the absence of the ARRA and monetary interventions, output and inflation would have been persistently lower than they actually were. In particular, annualized inflation would have been 40 basis points lower and annualized output growth would have been 28 basis points lower, on average over these eight quarters.

Based on the output and government spending paths we can also compute cumulative dollar-for-dollar multipliers

$$\mu_H = \frac{\sum_{\tau=1}^H (Y_{T^*-1+\tau}^o - Y_{T^*-1+\tau}^C)}{\sum_{\tau=1}^H (G_{T^*-1+\tau}^o - G_{T^*-1+\tau}^C)} \text{ where } H = 1, \dots, 8,$$

which are reported in Table 3. The pure fiscal multipliers are obtained by setting the monetary policy shocks to zero, whereas the combined fiscal and monetary multipliers are based on leaving the monetary policy shocks at their filtered values. The ex-post multipliers are around 0.8 according to our estimated model. We started out this section providing an explanation for why fiscal multipliers tend to be higher when the economy is at the ZLB. If we compute fiscal multipliers conditional on the economy being in a state in which the central bank would respond to rising output and inflation with an increase in interest rate, then the fiscal multipliers would be around 0.65, which is indeed lower than our ex-post multiplier. However, the presence of the ZLB only generates a modest increase in the multiplier.

The difference between the two types of multipliers reported in Table 3 is small, because as the no-intervention path of the nominal interest rate indicates, the adverse discount factor shock kept the economy close to the ZLB, leaving very little room for conventional monetary

⁹Details on the algorithm to compute the effects of the policy interventions are reported in the Online Appendix.

policy interventions ex post.

The red-dashed lines in the second-row panels of Figure 4 represent the post-2009:Q1 path of the exogenous shock processes in the absence of further innovations, which is the expected path conditional on 2009:Q1 innovation. Subsequent technology and discount factor innovations, had only small effects on the path of the respective exogenous processes, keeping them roughly in line with expectations. Thus, the low level of the discount factor shock also implied very low interest rates from an ex-ante perspective, leaving little scope for the Fed to boost the effects of a fiscal expansion through a zero-interest-rate policy.

8 Conclusion

Likelihood-based estimation of nonlinear DSGE models is computationally challenging. While it is becoming easier for economists to access powerful computer clusters that enable massive parallel computation, the ability to solve and estimate models on a desktop computer remains useful and desirable. Computations can often be simplified and accelerated considerably by taking shortcuts in regard to model solution and/or estimation techniques. The goal of this paper has been to develop a new solution method that captures important aspects of the nonlinearity generated by occasionally-binding constraints and, at the same time, allows for efficient filtering and likelihood-based estimation. The piecewise-linearity of the decision rules allows us to solve the model faster and to derive a conditionally-optimal proposal distribution for a particle filter. This filter delivers a much more accurate likelihood approximation than a standard bootstrap particle filter and enables us to estimate a nonlinear New Keynesian DSGE model with a ZLB constraint in a relatively short amount of time on a single core processor.

References

- ADAM, K., AND R. BILLI (2007): “Discretionary Monetary Policy at the Zero Lower Bound on Nominal Interest Rates,” *Journal of Monetary Economics*, 54, 728–752.
- AN, S., AND F. SCHORFHEIDE (2007): “Bayesian Analysis of DSGE Models,” *Econometric Reviews*, 26(2-4), 113–172.

- ANDRIEU, C., A. DOUCET, AND R. HOLENSTEIN (2010): “Particle Markov Chain Monte Carlo Methods,” *Journal of the Royal Statistical Society Series B*, 72(3), 269–342.
- ARULAMPALAM, S., S. MASKELL, N. GORDON, AND T. CLAPP (2002): “A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking,” *IEEE Transactions on Signal Processing*, 50(2), 174–188.
- ARUOBA, S. B., P. CUBA-BORDA, AND F. SCHORFHEIDE (2018): “Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries,” *Review of Economic Studies*, 85(1), 87–118.
- ARUOBA, S. B., AND F. SCHORFHEIDE (2016): “Inflation Dynamics During and After The Zero Lower Bound,” in *Inflation Dynamics and Monetary Policy*. 2015 Jackson Hole Symposium Volume Published by the Federal Reserve Bank of Kansas City.
- ATKINSON, T., A. RICHTER, AND N. A. THROCKMORTON (2020): “The Zero Lower Bound and Estimation Accuracy,” *Journal of Monetary Economics*, forthcoming.
- BATINI, N., L. EYRAUD, L. FORNI, AND A. WEBER (2014): “Fiscal Multipliers; Size, Determinants, and Use in Macroeconomic Projections,” IMF Technical Notes and Manuals 14/04, International Monetary Fund.
- BENIGNO, G., A. FOERSTER, C. OTROK, AND A. REBUCCI (2016): “Estimating Macroeconomic Models of Financial Crisis: An Endogenous Regime Switching Approach,” *Manuscript, University of Missouri*.
- BIANCHI, F., AND L. MELOSI (2017): “Escaping the Great Recession,” *American Economic Review*, 107(4), 1030–58.
- BOEHL, G. (2019): “Efficient Solution, Filtering and Estimation of Models with OBCs,” *Working Paper*.
- CAPPÉ, O., S. J. GODSILL, AND E. MOULINES (2007): “An Overview of Existing Methods and Recent Advances in Sequential Monte Carlo,” *Proceedings of the IEEE*, 95(5), 899–924.
- CAPPÉ, O., E. MOULINES, AND T. RYDEN (2005): *Inference in Hidden Markov Models*. Springer Verlag.

- CHEN, H. (2017): “The effects of the near-zero interest rate policy in a regime-switching dynamic stochastic general equilibrium model,” *Journal of Monetary Economics*, 90, 176 – 192.
- CHRISTIANO, L. J., AND M. EICHENBAUM (2012): “Notes on Linear Approximations, Equilibrium Multiplicity and E-learnability in the Analysis of the Zero Lower Bound,” *Manuscript, Northwestern University*.
- CHRISTIANO, L. J., M. EICHENBAUM, AND M. TRABANDT (2015): “Understanding the Great Recession,” *American Economic Journal: Macroeconomics*, 7(1), 110–67.
- CHRISTIANO, L. J., AND J. D. FISHER (2000): “Algorithms for Solving Dynamic Models with Occasionally Binding Constraints,” *Journal of Economic Dynamics & Control*, 24(8), 1179–1232.
- COLEMAN, C., S. LYON, L. MALIAR, AND S. MALIAR (2018): “Matlab, Python, Julia: What to Choose in Economics?,” CEPR Discussion Papers 13210, C.E.P.R. Discussion Papers.
- CREAL, D. (2012): “A Survey of Sequential Monte Carlo Methods for Economics and Finance,” *Econometric Reviews*, 31(3), 245–296.
- CUBA-BORDA, P., L. GUERRIERI, M. IACOVIELLO, AND M. ZHONG (2019): “Likelihood Evaluation of Models with Occasionally Binding Constraints,” *Journal of Applied Econometrics*, pp. 1–13.
- DE GROOT, O., B. DURDU, AND E. MENDOZA (2019): “Global v. Local Methods in the Analysis of Open-Economy Models with Incomplete Markets,” Discussion paper, National Bureau of Economic Research.
- DEL MORAL, P. (2013): *Mean Field Simulation for Monte Carlo Integration*. Chapman & Hall/CRC.
- DOUCET, A., AND A. M. JOHANSEN (2011): “A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later,” in *Handbook of Nonlinear Filtering*, ed. by D. Crisan, and B. Rozovsky. Oxford University Press.
- EGGERTSSON, G. B. (2011): “What fiscal policy is effective at zero interest rates?,” in *NBER Macroeconomics Annual 2010*, ed. by D. Acemoglu, and M. Woodford, vol. 25, pp. 59–112. University of Chicago Press.

- EGGERTSSON, G. B., AND M. WOODFORD (2003): “The Zero Bound on Interest Rates and Optimal Monetary Policy,” *Brookings Papers on Economic Activity*, 34(1), 139–235.
- FAIR, R. C., AND J. B. TAYLOR (1983): “Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models,” *Econometrica*, 51(4), 1169–1185.
- FARMER, R. E., D. F. WAGGONER, AND T. ZHA (2011): “Minimal state variable solutions to Markov-switching rational expectations models,” *Journal of Economic Dynamics and Control*, 35(12), 2150 – 2166.
- FERNÁNDEZ-VILLAYERDE, 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-VILLAYERDE, J., AND J. F. RUBIO-RAMÍREZ (2007): “Estimating Macroeconomic Models: A Likelihood Approach,” *Review of Economic Studies*, 74(4), 1059–1087.
- GORDON, N., AND D. SALMOND (1993): “A Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation,” *IEEE Proceedings-F*, 140, 107–113.
- 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.
- (2017): “Collateral constraints and macroeconomic asymmetries,” *Journal of Monetary Economics*, 90, 28 – 49.
- GUST, C., E. HERBST, D. LOPEZ-SALIDO, AND M. E. SMITH (2017): “The Empirical Implications of the Interest-Rate Lower Bound,” *American Economic Review*, 107(7), 1971–2006.
- HEISS, F., AND V. WINSCHER (2008): “Likelihood approximation by numerical integration on sparse grids,” *Journal of Econometrics*, 144(1), 62 – 80.
- HERBST, E., AND F. SCHORFHEIDE (2015): *Bayesian Estimation of DSGE Models*. Princeton University Press.
- HILLS, T. S., AND T. NAKATA (2018): “Fiscal Multipliers at the Zero Lower Bound: The Role of Policy Inertia,” *Journal of Money, Credit and Banking*, 50(1), 155–172.

- HOLDEN, T. D. (2019): “Existence and uniqueness of solutions to dynamic models with occasionally binding constraints,” *Manuscript, Deutsche Bundesbank*.
- JUDD, K. L., L. MALIAR, AND S. MALIAR (2010): “A Cluster-Grid Projection Method: Solving Problems with High Dimensionality,” *NBER Working Paper*, 15965.
- 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(C), 92–123.
- KANTAS, N., A. DOUCET, S. SINGH, J. MACIEJOWSKI, AND N. CHOPIN (2014): “On Particle Methods for Parameter Estimation in State-Space Models,” *arXiv Working Paper*, 1412.8659v1.
- KRUEGER, D., AND F. KUBLER (2004): “Computing Equilibrium in OLG Models with Stochastic Production,” *Journal of Economic Dynamics and Control*, 28(7), 1411–1436.
- KULISH, M., J. MORLEY, AND T. ROBINSON (2017): “Estimating DSGE Models with Zero Interest Rate Policy,” *Journal of Monetary Economics*, 88, 35–49.
- LIU, J. S. (2001): *Monte Carlo Strategies in Scientific Computing*. Springer Verlag.
- MALIAR, L., AND S. MALIAR (2014): “Numerical Methods for Large-Scale Dynamic Economic Models,” in *Handbook of Computational Economics Vol. 3*, ed. by K. Schmedders, and K. L. Judd, vol. 3, pp. 325–477. Elsevier.
- (2015): “Merging Simulation and Projection Approaches to Solve High-Dimensional Problems with an Application to a New Keynesian Model,” *Quantitative Economics*, 6(1), 1–47.
- MALIN, B. A., D. KRUEGER, AND F. KUBLER (2011): “Solving the multi-country real business cycle model using a Smolyak-collocation method,” *Journal of Economic Dynamics and Control*, 35(2), 229–239.
- MENDOZA, E. G., AND S. VILLALVAZO (2019): “FiPIt: A Simple, Fast Global Method for Solving Models with Two Endogenous States and Occasionally Binding Constraints,” *Manuscript, University of Pennsylvania*.
- NAKATA, T. (2016): “Optimal fiscal and monetary policy with occasionally binding zero bound constraints,” *Journal of Economic Dynamics and Control*, 73, 220 – 240.

RÍOS-RULL, J.-V., F. SCHORFHEIDE, C. FUENTES-ALBERO, M. KRYSHKO, AND R. SANTAEULALIA-LLOPIS (2012): “Methods versus Substance: Measuring the Effects of Technology Shocks,” *Journal of Monetary Economics*, 59(8), 826–846.

WOODFORD, M. (2003): *Interest and Prices*. Princeton University Press.

Online Appendix to “Piecewise-Linear Approximations and Filtering for DSGE Models with Occasionally-Binding Constraints”

**S. Boragan Aruoba, Pablo Cuba-Borda, Kenji Higa-Flores, Frank Schorfheide,
and Sergio Villalvazo**

This Appendix consists of the following sections:

- A. Equilibrium Conditions for the Model of Section 2
- B. Proofs and Derivations for Section 5
- C. Additional Details for the Empirical Application

A Equilibrium Conditions for the Model of Section 2

In this section we sketch the derivation of the equilibrium conditions presented in Section 2.

A.1 Households

The representative household solves

$$\max_{\{C_{t+s}, H_{t+s}, B_{t+s}, M_{t+s}\}} \mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s d_{t+s} \left(\frac{(C_{t+s}/A_{t+s})^{1-\tau} - 1}{1-\tau} - \chi_H \frac{H_{t+s}^{1+1/\eta}}{1+1/\eta} + \chi_M V \left(\frac{M_{t+s}}{P_{t+s} A_{t+s}} \right) \right) \right],$$

subject to:

$$P_t C_t + T_t + B_t + M_t = P_t W_t H_t + M_{t-1} + R_{t-1} B_{t-1} + P_t D_t + P_t S C_t,$$

Consumption and bond holdings. Let $\beta^s d_{t+s} \lambda_{t+s}$ be the Lagrange multiplier on the household budget constraint. Then the first-order condition with respect to consumption and bond holdings are given by:

$$\begin{aligned} P_t \lambda_t &= \left(\frac{C_t}{A_t} \right)^{-\tau} \frac{1}{A_t} \\ \lambda_t &= \beta \frac{d_{t+1}}{d_t} R_t \lambda_{t+1}. \end{aligned}$$

Combining the two equation of the bond holding first order condition, leads to the consumption Euler equation:

$$1 = \beta \mathbb{E}_t \left[\frac{d_{t+1}}{d_t} \left(\frac{C_{t+1}/A_{t+1}}{C_t/A_t} \right)^{-\tau} \frac{1}{\gamma z_{t+1}} \frac{R_t}{\pi_{t+1}} \right],$$

where $\gamma z_{t+1} = A_{t+1}/A_t$. We define the stochastic discount factor as:

$$Q_{t+1|t} = \frac{d_{t+1}}{d_t} \left(\frac{C_{t+1}/A_{t+1}}{C_t/A_t} \right)^{-\tau} \frac{1}{\gamma z_{t+1}}.$$

Labor-Leisure Choice. Taking first-order conditions with respect to H_t yields the standard

intratemporal optimality condition for the allocation of labor

$$\frac{W_t}{A_t} = \chi_H \left(\frac{C_t}{A_t} \right)^\tau H_t^{1/\eta}.$$

A.2 Intermediate Goods Firms

Each intermediate good producer buys labor services $H_t(j)$ at the real wage W_t . Firms face nominal rigidities in terms of price adjustment costs. The adjustment cost, expressed as a fraction of firms' real output, is given by the function $\Phi_p \left(\frac{P_t(j)}{P_{t-1}(j)} \right)$. We assume that the adjustment cost function is twice-continuously differentiable, weakly increasing and weakly convex, $\Phi'_p \geq 0$ and $\Phi''_p \geq 0$. The firm maximizes expected discounted real profits with respect to $H_t(j)$ and $P_t(j)$:

$$\mathbb{E}_t \sum_{s=0}^{\infty} \beta^s Q_{t+s|t} \left(\frac{P_{t+s}(j)}{P_{t+s}} A_{t+s} H_{t+s}(j) - \Phi_p \left(\frac{P_{t+s}(j)}{P_{t+s-1}(j)} \right) A_{t+s} H_{t+s}(j) - W_{t+s} H_{t+s}(j) \right),$$

subject to

$$A_t H_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-1/\nu} Y_t.$$

We use $\mu_{t+s}\beta^s Q_{t+s|t}$ to denote the Lagrange multiplier associated with this constraint. In equilibrium, the firms use the households' stochastic discount factor to discount future profits.

Price setting decision. Setting $Q_{t|t} = 1$, the first-order condition with respect to $P_t(j)$ is given by:

$$\begin{aligned} 0 = & \frac{A_t H_t(j)}{P_t} - \Phi'_p \left(\frac{P_t(j)}{P_{t-1}(j)} \right) \frac{A_t H_t(j)}{P_{t-1}(j)} - \frac{\mu_t}{\nu} \left(\frac{P_t(j)}{P_t} \right)^{-1/\nu-1} \frac{Y_t}{P_t} \\ & + \beta \mathbb{E}_t \left[Q_{t+1|t} \Phi'_p \left(\frac{P_{t+1}(j)}{P_t(j)} \right) A_{t+1} H_{t+1}(j) \frac{P_{t+1}(j)}{P_t^2(j)} \right]. \end{aligned}$$

Firms' labor demand. Taking first-order conditions with respect to $H_t(j)$ yields

$$W_t = \frac{P_t(j)}{P_t} A_t - \Phi_p \left(\frac{P_t(j)}{P_{t-1}(j)} \right) A_t - \mu_t A_t.$$

Symmetric equilibrium. We restrict attention to a symmetric equilibrium where all firms choose the same price $P_t(j) = P_t \forall j$. This assumption implies that in equilibrium all firms

face identical marginal costs and demand the same amount of labor input. Combining the firms' price setting and labor demand first order conditions and assuming that the price adjustment costs are quadratic, i.e.,

$$\Phi_p \left(\frac{P_t(j)}{P_{t-1}(j)} \right) = \frac{\phi}{2} \left(\frac{P_t(j)}{P_{t-1}(j)} - \bar{\pi} \right)^2,$$

we obtain:

$$(1 - \nu) - \chi_H \left(\frac{C_t}{A_t} \right)^\tau H_t^{1/\eta} - \frac{\phi}{2} \left(\frac{P_t}{P_{t-1}} - \bar{\pi} \right)^2 + \nu \phi \left(\frac{P_t}{P_{t-1}} - \bar{\pi} \right) \frac{P_t}{P_{t-1}} = \nu \beta \mathbb{E}_t \left[Q_{t+1|t} \frac{P_{t+1}}{P_t} \Phi'_p \left(\frac{P_{t+1}}{P_t} \right) \frac{Y_{t+1}}{Y_t} \right].$$

A.3 Equilibrium Conditions

Resource constraint. The derivation of the aggregate resource constraint is straightforward. In equilibrium real profits by intermediate producers is given by:

$$D_t = Y_t - \Phi_p(\pi_t) Y_t - W_t H_t$$

Substituting this into the household budget constraint we obtain:

$$C_t + \left[\frac{T_t}{P_t} + \frac{M_t}{P_t} + \frac{B_t}{P_t} - \frac{M_{t-1}}{P_t} - \frac{R_{t-1} B_{t-1}}{P_t} \right] = W_t H_t + Y_t - \Phi_p(\pi_t) Y_t - W_t H_t$$

From the government budget constraint in (7) we can see that the term in square brackets corresponds to real government expenditure G_t . Simplifying yields:

$$C_t + G_t = [1 - \Phi_p(\pi_t)] Y_t$$

The technology process introduces a long-run trend in the variables of the model. To make the model stationary we use the following transformations: $y_t = Y_t/A_t$, $c_t = C_t/A_t$, and note that $Y_t/Y_{t-1} = \frac{y_t}{y_{t-1}} \gamma z_t$. We also define the gross inflation rate $\pi_t = P_t/P_{t-1}$. The equilibrium conditions shown in Section 2.2 of the main text follow immediately.

Steady States and Reparameterizations. Let

$$r = \gamma/\beta, \quad \phi = \frac{\tau(1-\nu)}{\nu\pi_*\kappa}, \quad b = \frac{1}{2\nu}.$$

The steady states are given by

$$\begin{aligned} c_* &= \left[\frac{1 - \nu + \phi\nu(1 - \beta)\pi_*(\pi_* - \bar{\pi}) - 0.5\phi * (\pi_* - \bar{\pi})^2}{\chi_H((1/(g_*)) - 0.5\phi(\pi_* - \bar{\pi})^2)^{-1/\eta}} \right]^{(1/(\tau+1/\eta))} \\ y_* &= \frac{c_*}{1/g_* - 0.5\phi(\pi_* - \bar{\pi})^2} \\ R_* &= \pi_* r \end{aligned}$$

B Derivations for Section 5

We provide a proof of Proposition 1 which contains the formulas for the terms λ_t , $\bar{\eta}_{1,1}(\cdot)$, $\bar{\Omega}_{11}(\cdot)$, $\bar{\eta}_{2|1,t}(\cdot)$, and $\bar{\Omega}_{2|1}(\cdot)$ that appear in the proposition. Throughout this section we set the intercept in the measurement equation $A_0 = 0$ and we drop the subscript from the matrix A_s .

Proof of Proposition 1. Conditional on s_{t-1} the current state s_t is determined by η_t . In order to derive $g_t^*(\tilde{s}_t|s_{t-1}^j)$ (we are omitting θ from the conditioning set), we will work in the (η_t, s_{t-1}) space and derive (also omitting tildes and j superscripts) the conditionally-optimal proposal distribution

$$g_t^*(\eta_t|y_t, s_{t-1}) = p(\eta_t|y_t, s_{t-1}) \propto p(y_t|\eta_t, s_{t-1})p(\eta_t)$$

and the incremental particle weights

$$\tilde{\omega}_t^j = p(y_t|s_{t-1}) = \int p(y_t|\eta_t, s_{t-1})p(\eta_t)d\eta_t.$$

Define

$$\hat{y}_{t|t-1}(\cdot) = A(\Phi_0(\cdot) + \Phi_1(\cdot)s_{t-1}) + A\Phi_\eta(\cdot)\eta_t, \quad \nu_t(\cdot) = y_t - \hat{y}_{t|t-1}(\cdot).$$

We will denote the density of a $N(\mu, \Sigma)$ random variable Y by $p_N(y; \mu, \Sigma)$. Using this notation, we write

$$\begin{aligned} & p(y_t|\eta_t, s_{t-1})p(\eta_t) \\ &= p_N(y_t; \hat{y}_{t|t-1}(n), \varsigma\Sigma_u)p_N(\eta_{1,t}; 0, 1)p_N(\eta_{2,t}; 0, I)\mathbb{I}\{\eta_{1,t} \leq \zeta(s_{t-1})\} \\ & \quad + p_N(y_t; \hat{y}_{t|t-1}(b), \varsigma\Sigma_u)p_N(\eta_{1,t}; 0, 1)p_N(\eta_{2,t}; 0, I)\mathbb{I}\{\eta_{1,t} > \zeta(s_{t-1})\} \\ &= I + II. \end{aligned} \tag{A.1}$$

We will begin by manipulating term $I(n)$. Omitting the (n) arguments we obtain:

$$\begin{aligned}
I &= \mathbb{I}\{\eta_{1,t} \leq \tilde{\zeta}(s_{t-1})\} (2\pi)^{-n_y/2} |\varsigma \Sigma_u|^{-1/2} (2\pi)^{-n_\eta/2} \\
&\quad \times \exp \left\{ -\frac{1}{2} (\nu_t - A\Phi_\eta \eta_t)' (\varsigma \Sigma_u)^{-1} (\nu_t - A\Phi_\eta \eta_t) \right\} \exp \left\{ -\frac{1}{2} (\eta_{1,t}^2 + \eta_{2,t}' \eta_{2,t}) \right\} \\
&= \mathbb{I}\{\eta_{1,t} \leq \tilde{\zeta}(s_{t-1})\} (2\pi)^{-n_y/2} |\varsigma \Sigma_u|^{-1/2} (2\pi)^{-n_\eta/2} \exp \left\{ -\frac{1}{2} \nu_t' \Sigma_u^{-1} \nu_t \right\} \\
&\quad \times \exp \left\{ -\frac{1}{2} \eta_t' \Phi_\eta' A' (\varsigma \Sigma_u)^{-1} A \Phi_\eta \eta_t + \nu_t' (\varsigma \Sigma_u)^{-1} A \Phi_\eta \eta_t \right\} \exp \left\{ -\frac{1}{2} (\eta_{1,t}^2 + \eta_{2,t}' \eta_{2,t}) \right\}.
\end{aligned}$$

Note that term I takes the form of a product between “likelihood function” and “prior.” The prior covariance matrix of η_t is $\underline{\Omega} = I$, and the negative Hessian and the maximum of the “log-likelihood” function are

$$\hat{\Omega}^{-1} = \Phi_\eta' A' (\varsigma \Sigma_u)^{-1} A \Phi_\eta, \quad \hat{\eta}_t = \hat{\Omega} \Phi_\eta' A' (\varsigma \Sigma_u)^{-1} \nu_t. \quad (\text{A.2})$$

With this notation we can write

$$\begin{aligned}
I &= \mathbb{I}\{\eta_{1,t} \leq \zeta(s_{t-1})\} (2\pi)^{-n_y/2} |\varsigma \Sigma_u|^{-1/2} (2\pi)^{-n_\eta/2} |\underline{\Omega}|^{-1/2} \exp \left\{ -\frac{1}{2} \nu_t' \Sigma_u^{-1} \nu_t \right\} \\
&\quad \times \exp \left\{ -\frac{1}{2} (\eta_t' \hat{\Omega}^{-1} \eta_t - 2 \hat{\eta}_t' \hat{\Omega}^{-1} \eta_t) \right\} \exp \left\{ -\frac{1}{2} \eta_t' \underline{\Omega} \eta_t \right\}.
\end{aligned}$$

Now define the quasi posterior mean and covariance matrices for $\eta_t | (y_t, s_{t-1})$

$$\bar{\Omega} = (\underline{\Omega}^{-1} + \hat{\Omega}^{-1})^{-1}, \quad \bar{\eta}_t = \bar{\Omega} \hat{\Omega}^{-1} \hat{\eta}_t. \quad (\text{A.3})$$

This leads to

$$\begin{aligned}
I &= (2\pi)^{-n_y/2} |\varsigma \Sigma_u|^{-1/2} |\underline{\Omega}|^{-1/2} |\bar{\Omega}|^{1/2} \exp \left\{ -\frac{1}{2} \nu_t' \Sigma_u^{-1} \nu_t \right\} \exp \left\{ \frac{1}{2} \bar{\eta}_t' \bar{\Omega}^{-1} \bar{\eta}_t \right\} \\
&\quad \times \mathbb{I}\{\eta_{1,t} \leq \zeta(s_{t-1})\} (2\pi)^{-n_\eta/2} |\bar{\Omega}|^{-1/2} \exp \left\{ -\frac{1}{2} (\eta_t - \bar{\eta}_t)' \bar{\Omega}^{-1} (\eta_t - \bar{\eta}_t) \right\}.
\end{aligned}$$

We now decompose the kernel of the “posterior” of $\eta_t | (y_t, s_{t-1})$ in the second line of the preceding equation into a conditional and a marginal distribution. We use the “1” subscript to indicate the marginal posterior of η_1 and the “2|1” subscript to indicate the conditional

mean and variance associated with the posterior of $\eta_{2,t}$ given $\eta_{1,t}$:

$$\bar{\eta}_{2|1,t}(\eta_{1,t}) = \bar{\eta}_{2,t} + \bar{\Omega}_{21}\bar{\Omega}_{11}^{-1}(\eta_{1,t} - \bar{\eta}_{1,t}), \quad \bar{\Omega}_{2|1,t} = \bar{\Omega}_{22} - \bar{\Omega}_{21}\bar{\Omega}_{11}^{-1}\bar{\Omega}_{12} \quad (\text{A.4})$$

Thus,

$$\begin{aligned} I &= (2\pi)^{-n_y/2} |\zeta \Sigma_u|^{-1/2} (2\pi)^{-n_\eta/2} |\underline{\Omega}|^{-1/2} |\bar{\Omega}|^{1/2} \exp \left\{ -\frac{1}{2} \nu_t' \Sigma_u^{-1} \nu_t \right\} \exp \left\{ \frac{1}{2} \bar{\eta}_t' \bar{\Omega}^{-1} \bar{\eta}_t \right\} \\ &\times (2\pi)^{-(n_\eta-1)/2} |\bar{\Omega}_{2|1}|^{-1/2} \exp \left\{ -\frac{1}{2} (\eta_{2,t} - \bar{\eta}_{2|1,t})' \bar{\Omega}_{2|1}^{-1} (\eta_{2,t} - \bar{\eta}_{2|1,t}) \right\} \\ &\times \mathbb{I}\{\eta_{1,t} \leq \zeta(s_{t-1})\} (2\pi)^{-1/2} |\bar{\Omega}_{11}|^{-1/2} \exp \left\{ -\frac{1}{2} (\eta_{1,t} - \bar{\eta}_{1,t})' \bar{\Omega}_{11}^{-1} (\eta_{1,t} - \bar{\eta}_{1,t}) \right\}. \end{aligned} \quad (\text{A.5})$$

This is the final form for term I in (A.1).

Integrating I in (A.5) with respect to $(\eta_{1,t}, \eta_{2,t})$ and re-introducing the (n) arguments yields

$$\begin{aligned} D(n) &= \int \int I(\eta_{1,t}, \eta_{2,t}) d\eta_{2,t} d\eta_{1,t} \\ &= (2\pi)^{-n_y/2} |\zeta \Sigma_u|^{-1/2} |\underline{\Omega}|^{-1/2} |\bar{\Omega}(n)|^{1/2} \Phi_N((\zeta(s_{t-1}) - \bar{\eta}_{1,t}(n)) / \sqrt{\bar{\Omega}_{11}(n)}) \\ &\times \exp \left\{ -\frac{1}{2} \nu_t(n)' \Sigma_u^{-1} \nu_t(n) \right\} \exp \left\{ \frac{1}{2} \bar{\eta}_t'(n) \bar{\Omega}^{-1}(n) \bar{\eta}_t(n) \right\}. \end{aligned} \quad (\text{A.6})$$

The analysis of term II proceeds in almost identical manner, with the understanding that term I depends on $\Phi_0(n)$, $\Phi_1(n)$, and $\Phi_\eta(n)$, whereas term II depends on $\Phi_0(b)$, $\Phi_1(b)$, and $\Phi_\eta(b)$. As a consequence the posterior coefficient matrices $\hat{\eta}$, $\hat{\Omega}$, $\bar{\eta}$, and $\bar{\Omega}$ should also be indexed by either (n) or (b) . Because for term II the inequality in the indicator function is reversed, we obtain

$$\begin{aligned} D(b) &= (2\pi)^{-n_y/2} |\zeta \Sigma_u|^{-1/2} |\underline{\Omega}|^{-1/2} |\bar{\Omega}(b)|^{1/2} \left(1 - \Phi_N((\zeta(s_{t-1}) - \bar{\eta}_{1,t}(b)) / \sqrt{\bar{\Omega}_{11}(b)}) \right) \\ &\times \exp \left\{ -\frac{1}{2} \nu_t(n)' \Sigma_u^{-1} \nu_t(n) \right\} \exp \left\{ \frac{1}{2} \bar{\eta}_t'(b) \bar{\Omega}^{-1}(b) \bar{\eta}_t(b) \right\} \end{aligned} \quad (\text{A.7})$$

Using the formulas for I , II , $D(n)$, and $D(b)$, we can write the posterior density of η_t as follows:

$$p(\eta_t | y_t, s_{t-1}) = \frac{p(y_t | \eta_t, s_{t-1}) p(s_{t-1})}{\int p(y_t | \eta_t, s_{t-1}) p(s_{t-1}) d\eta_t} = \frac{I(n) + II(b)}{D(n) + D(b)}. \quad (\text{A.8})$$

Thus, the resulting conditionally-optimal proposal is given by the following mixture. Define

$$\lambda = \frac{D(n)}{D(n) + D(b)}. \quad (\text{A.9})$$

Then, with probability λ

$$\eta_{1,t} \sim N(\bar{\eta}_{1,t}(n), \bar{\Omega}_{11}(n) \mathbb{I}\{\eta_{1,t} \leq \zeta(s_{t-1})\}), \quad \eta_{2,t} | \eta_{1,t} \sim N(\bar{\eta}_{2|1,t}(n, \eta_{1,t}), \bar{\Omega}_{2|1}(n)) \quad (\text{A.10})$$

and with probability $1 - \lambda$

$$\eta_{1,t} \sim N(\bar{\eta}_{1,t}(b), \bar{\Omega}_{11}(b) \mathbb{I}\{\eta_{1,t} > \zeta(s_{t-1})\}), \quad \eta_{2,t} | \eta_{1,t} \sim N(\bar{\eta}_{2|1,t}(b, \eta_{1,t}), \bar{\Omega}_{2|1}(b)). \quad (\text{A.11})$$

The incremental weight is constant and given by the following formula:

$$\tilde{\omega}_t^j = p(y_t | s_{t-1}^j) = D(n) + D(b). \quad (\text{A.12})$$

This completes the proof of the proposition. ■

In the remainder of this section we consider to special cases: (i) y_t identifies the regime without error. This is the case, for instance, for a DSGE model with ZLB constraint if the interest rate is observed without error, at least when it hits the ZLB. (ii) Measurement errors that are zero or very close to zero.

(i) Known Regime. Let $y_t = [y'_{1,t}, y_{2,t}]$ and partition $A' = [A'_1, A'_2]$ so that the partitions of A conform with the partitions of y_t . Assume that the n -regime is active if and only if $y_{2,t} > c$. In the b -regime $y_{2,t} = c$. Moreover, let $\delta(y_{2t}; c)$ denote the Dirac delta function with the property that $\delta(y_{2t}; c) = 0$ for $y_{2t} \neq c$ and $\int \delta(y_{2t}; c) dy_{2t} = 1$. Using the above notation, we can rewrite (A.1) as

$$\begin{aligned} & p(y_t | \eta_t, s_{t-1}) p(\eta_t) \\ &= p_N(y_t; \hat{y}_{t|t-1}(n), \varsigma \Sigma_u) p_N(\eta_{1,t}; 0, 1) p_N(\eta_{2,t}; 0, I) \mathbb{I}\{\eta_{1,t} \leq \zeta(s_{t-1})\} \\ & \quad + p_N(y_t; \hat{y}_{t|t-1}(b), \varsigma \Sigma_{u,11}) \delta(y_{2t}; c) p_N(\eta_{1,t}; 0, 1) p_N(\eta_{2,t}; 0, I) \mathbb{I}\{\eta_{1,t} > \zeta(s_{t-1})\} \\ &= I + II \end{aligned}$$

The formula for $D(n)$ in (A.6) remains unchanged. The formula for $D(b)$ in (A.7) changes to

$$\tilde{D}(b) = \delta(y_{2t}; c) D(b)$$

with the understanding that Σ_u in $D(b)$ needs to be replaced by $\Sigma_{u,11}$. After having observed y_t we know whether $y_{2,t} = c$ so that we can define

$$\lambda = \mathbb{I}\{y_{2,t} > c\}.$$

Conditional on λ , we can simulate $[\eta_{1,t}, \eta'_{2,t}]'$ from (A.10) and (A.11), respectively. Finally,

$$\tilde{\omega}_t^j = p(y_t | s_{t-1}^j) = \begin{cases} D(n) & \text{if } y_{2,t} > c \\ D(b) & \text{if } y_{2,t} = c \end{cases},$$

where here $D(b)$ corresponds to (A.7) and does not include the Dirac function $\delta(y_{2,t}; c)$.

(ii) Zero Measurement Errors. Consider a linear state-space model without regimes:

$$y_t = As_t + u_t, \quad s_t = \Phi_0 + \Phi_1 s_{t-1} + \Phi_\eta \eta_t.$$

Ignoring the censoring and dropping the regime indicator, note that $D = p(y_t | s_{t-1})$. We can write

$$y_t = A(\Phi_0 + \Phi_1 s_{t-1}) + A\Phi_\eta \eta_t + u_t.$$

Note that

$$A\Phi_\eta \eta_t + u_t \sim N(0, A\Phi_\eta \Phi'_\eta A' + \varsigma \Sigma_u).$$

Thus, we can deduce that the term $D(n)$ in (A.6) can be rewritten as

$$\begin{aligned} D(n) &= (2\pi)^{-n_u/2} |A\Phi_\eta \Phi'_\eta A' + \varsigma \Sigma_u|^{-1/2} \\ &\times \exp \left\{ -\frac{1}{2} (y_t - A(\Phi_0 + \Phi_1 s_{t-1}))' [A\Phi_\eta \Phi'_\eta A' + \varsigma \Sigma_u]^{-1} (y_t - A(\Phi_0 + \Phi_1 s_{t-1})) \right\} \\ &\times \Phi_N((\zeta(s_{t-1}) - \bar{\eta}_1) / \sqrt{\bar{\Omega}_{11}}) \end{aligned}$$

A similar adjustment can be made to the term $D(b)$ in (A.6). The advantage of this alternative expression is that we can take the limit $\varsigma \rightarrow 0$. The argument of the Gaussian cdf

behaves as follows:

$$\begin{aligned}
\bar{\eta} &= \bar{\Omega} \hat{\Omega}^{-1} \hat{\eta}_t \\
&= (\underline{\Omega}^{-1} + \hat{\Omega}^{-1})^{-1} \Phi'_\eta A' (\varsigma \Sigma_u)^{-1} \nu_t \\
&= \varsigma^{-1} (\underline{\Omega}^{-1} + \varsigma^{-1} \Phi'_\eta A' \Sigma_u^{-1} A \Phi_\eta)^{-1} \Phi'_\eta A' \Sigma_u^{-1} \nu_t \\
&= (\varsigma \underline{\Omega}^{-1} + \Phi'_\eta A' \Sigma_u^{-1} A \Phi_\eta)^{-1} \Phi'_\eta A' \Sigma_u^{-1} \nu_t, \\
&\longrightarrow (\Phi'_\eta A' \Sigma_u^{-1} A \Phi_\eta)^{-1} \Phi'_\eta A' \Sigma_u^{-1} \nu_t
\end{aligned}$$

which eliminates divisions by ς . Moreover,

$$\bar{\Omega} = \varsigma (\varsigma \underline{\Omega}^{-1} + \Phi'_\eta A' \Sigma_u^{-1} A \Phi_\eta)^{-1} \longrightarrow 0.$$

Thus,

$$\lim_{\varsigma \rightarrow 0} \Phi_N((\zeta(s_{t-1}) - \bar{\eta}_1) / \sqrt{\bar{\Omega}_{11}}) = \begin{cases} 1 & \text{if } \zeta(s_{t-1}) - \bar{\eta}_1 \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Because the posterior covariances matrices are zero in the limit, the sampling in (A.10) and (A.11) is replaced by setting $\eta_{1,t}$ and $\eta_{2,t}$ equal to their means.

C Additional Details for the Empirical Application

Prior. Table A-1 summarizes the prior distribution for the initial states used in the empirical analysis.

Table A-1: Prior Distributions for Initial States

Parameter	Density	P(1)	P(2)
$\epsilon_{R,0}$	\mathcal{N}	0.00	.002
\hat{g}_0	\mathcal{N}	0.00	.012
\hat{z}_0	\mathcal{N}	0.00	.008
\hat{d}_0	\mathcal{N}	0.00	.170
$\hat{\pi}_0$	\mathcal{N}	0.00	.030
\hat{c}_0	\mathcal{N}	0.00	.030
$\hat{R}_0 + 0.01$	\mathcal{G}	0.01	.008

Notes: \mathcal{N} is Normal distribution; \mathcal{G} is Gamma distribution. P(1) and P(2) are mean and standard deviations for Normal and Gamma distributions. We set $\hat{y}_0 = \hat{c}_0 + \hat{g}_0$ and $\hat{y}_{-1} = \hat{y}_0$

Calibration of ARRA. Table A-2 summarizes the award and disbursements of funds for federal contracts, grants, and loans. We translate the numbers in the table into a one-period location shift of the distribution of $\epsilon_{g,t}$ below.

Table A-2: ARRA Funds for Contracts, Grant, and Loans

	Awarded	Received	Nominal GDP
2009:2	158	36	3488
2009:3	17	18	3533
2009:4	26	8	3568
2010:1	16	24	3603
2010:2	33	26	3644
2010:3	9	21	3684
2010:4	4	19	3704
2011:1	4	20	3751
2011:2	8	17	3791
2011:3	0	12	3830
2011:4	3	9	3870
2012:1	0	8	3899

Notes: Data were obtained from www.recovery.gov.

Note that a one-time innovation $\epsilon_{g,t}^{ARRA}$ generates a response

$$\hat{g}_{t+h}^{ARRA} = \rho_g^h \sigma_g \epsilon_{g,t}^{ARRA}.$$

We use the log-linear approximation

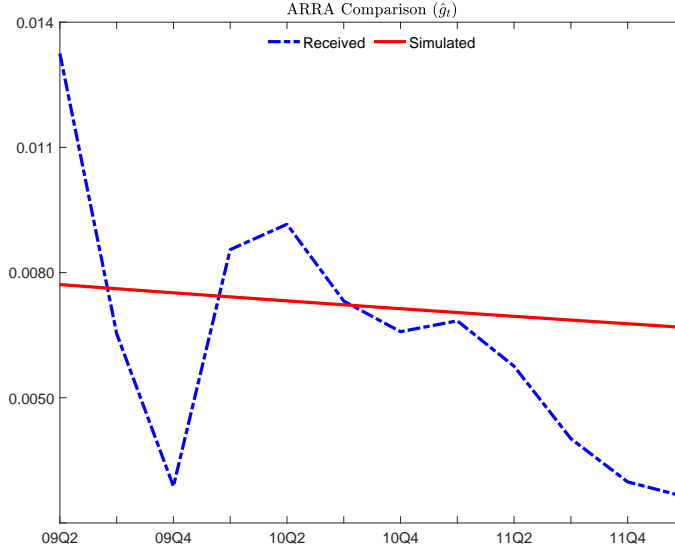
$$\hat{\zeta}_t^{ARRA} = \frac{1}{g_* - 1} \hat{g}_t^{ARRA},$$

where $\zeta_t = G_t/Y_t$. We connect $\hat{\zeta}_t^{ARRA}$ to the data in Table A-2 using the relationship

$$\hat{\zeta}_t^{ARRA} = \log \left(\frac{G_t^{ARRA}/Y_t}{G_*/Y_*} \right).$$

Figure A-2 compares the time path of \hat{g}_t^{ARRA} constructed from the impulse response to a $\epsilon_{g,t}^{ARRA} = 0.0077$ (the red solid line) and the time path constructed from the disbursements in Table A-2 (the blue dashed line).¹⁰

Figure A-1: Calibration of Fiscal Policy Intervention



Computational Details for Fiscal Policy Experiment. The following algorithm describes how we compute the effect of a combined fiscal and monetary intervention.

¹⁰Recall that $\sigma_g = 0.0024$, hence the ARRA impulse in our experiment is equal to $3.2 \times \sigma_g$.

Algorithm 2 (Effect of Combined Fiscal and Monetary Policy Intervention)

1. Initialize the simulation by setting $(R_0, y_0, z_0, g_0, d_0)$ equal to the mean estimate obtained with the particle filter
2. Generate a baseline trajectory that includes the intervention based on the sequence of innovations obtained from the COPF: $\{\epsilon_{z,T^*+s}, \epsilon_{g,T^*+s}, \epsilon_{d,T^*+s}, \epsilon_{R,T^*+s}\}_{s=0}^H$.
3. Generate the innovation sequence for the counterfactual trajectories without intervention according to

$$\begin{aligned}
\epsilon_{g,T^*} &= \epsilon_{g,T^*}^I - \delta^{ARRA}, & \epsilon_{g,T^*+s} &= \epsilon_{g,T^*+s}^I \quad \text{for } s = 1, \dots, H; \\
\epsilon_{z,T^*+s} &= \epsilon_{z,t}^I \quad \text{for } s = 0, \dots, H; \\
\epsilon_{d,T^*+s} &= \epsilon_{d,t}^I \quad \text{for } s = 0, \dots, H; \\
\epsilon_{R,T^*+s} &= 0 \quad \text{for } s = 0, \dots, H;
\end{aligned}$$

4. Conditional on $(R_0, y_0, z_0, g_0, d_0)$, compute $\{R_{T^*+s}, y_{T^*+s}, \pi_{T^*+s}\}_{s=0}^H$ and $\{R_{T^*+s}^I, y_{T^*+s}^I, \pi_{T^*+s}^I\}_{s=0}^H$ based on $\{\epsilon_{T^*+s}\}$ and $\{\epsilon_{T^*+s}^I\}$, respectively, and let

$$IRF(x_t | \epsilon_t^I, \epsilon_t) = (\ln x_t^I - \ln x_t). \quad (\text{A.13})$$

We report results for $\delta^{ARRA} = 0.0077$ and $H = 7$ in the main text. When we consider only a fiscal policy, we set $\epsilon_{R,t}^I = 0$ for $t = T^*, \dots, T^* + 7$ as well.