

DISCUSSION PAPER SERIES

DP15388

**PIECEWISE-LINEAR APPROXIMATIONS
AND FILTERING FOR DSGE MODELS
WITH OCCASIONALLY BINDING
CONSTRAINTS**

Boragan Aruoba, Pablo Cuba-Borda, Kenji Hilga-
Flores, Frank Schorfheide and Sergio Villalvazo

MONETARY ECONOMICS AND FLUCTUATIONS



PIECEWISE-LINEAR APPROXIMATIONS AND FILTERING FOR DSGE MODELS WITH OCCASIONALLY BINDING CONSTRAINTS

Boragan Aruoba, Pablo Cuba-Borda, Kenji Hilga-Flores, Frank Schorfheide and Sergio Villalvazo

Discussion Paper DP15388
Published 19 October 2020
Submitted 15 October 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Monetary Economics and Fluctuations

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Boragan Aruoba, Pablo Cuba-Borda, Kenji Hilga-Flores, Frank Schorfheide and Sergio Villalvazo

PIECEWISE-LINEAR APPROXIMATIONS AND FILTERING FOR DSGE MODELS WITH OCCASIONALLY BINDING CONSTRAINTS

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 with 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

Keywords: Bayesian estimation, Effective Lower Bound on Nominal Interest Rates, Nonlinear Filtering, Nonlinear Solution Methods, Particle MCMC

Boragan Aruoba - aruoba@umd.edu
University of Maryland

Pablo Cuba-Borda - pablo.a.cubaborda@frb.gov
Federal Reserve Board, DC

Kenji Hilga-Flores - kenjihf@umd.edu
University of Maryland

Frank Schorfheide - schorf@ssc.upenn.edu
University of Pennsylvania and CEPR

Sergio Villalvazo - vsergio@sas.upenn.edu
University of Pennsylvania

Acknowledgements

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. Much of this paper was

written while Aruoba and Schorfheide visited the Federal Reserve Bank of Philadelphia, whose hospitality they are thankful for. Higa-Flores and Villalvazo gratefully acknowledge financial support from the Becker Friedman Institute under the Macro Financial Modeling Project. 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 any other person associated with the Federal Reserve System.

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

S. Borağan 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: October 9, 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 with 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: C5, E4, E5

KEY WORDS: Bayesian Estimation, Effective Lower Bound on Nominal Interest Rates, Nonlinear Filtering, Nonlinear Solution Methods, Particle MCMC

* 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 (vsergio@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. Much of this paper was written while Aruoba and Schorfheide visited the Federal Reserve Bank of Philadelphia, whose hospitality they are thankful for. Higa-Flores and Villalvazo gratefully acknowledge financial support from the Becker Friedman Institute under the Macro Financial Modeling Project. 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 any other person 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 effective lower bound (ELB) on nominal interest rates. In order for these models to be usable 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—for example, 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. Examples of this approach include Christiano and Fisher (2000), Adam and Billi (2007), Fernández-Villaverde, Gordon, Guerrón-Quintana, and Rubio-Ramírez (2015), Maliar and Maliar (2015), Nakata (2016), Gust, Herbst, Lopez-Salido, and Smith (2017), Aruoba, Cuba-Borda, and Schorfheide (2018), Mendoza and Villalvazo (2020), and Atkinson, Richter, and Throckmorton (2020).

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. With 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 can then compare the initial guess about the duration of the binding regime to the backward solution and iterate 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. Variants of this approach have been used in Eggertsson and Woodford (2003), Christiano, Eichenbaum, and Trabandt (2015), Guerrieri and Iacoviello (2015), Kulish, Morley, and Robinson (2017), Holden (2019), and Boehl (2019). The Guerrieri and Iacoviello (2015) paper is accompanied by a popular model

solution toolbox called OccBin that implements a variant of the EFPF approach. We will refer to OccBin in various instances throughout our paper.

Given the model solution, one then constructs 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 an ELB constraint, Gust et al. (2017) and Aruoba et al. (2018) use a particle filter in combination with a global solution to construct likelihood functions. Guerrieri and Iacoviello (2017) use an EFPF 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.

Against this backdrop, 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. 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. For instance, the small-scale New Keynesian model in our empirical application is estimated using U.S. data from 1984 to 2018 in about 13.5⁶ 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 along the locus of the state space in which a constraint becomes binding. The coefficients of the decision rules are determined to ensure that the model's equilibrium conditions are (approximately) satisfied. The equilibrium conditions typically take the form of nonlinear expectational difference equations. We require that the (potentially transformed) state variables enter the occasionally-binding constraint linearly. For the remaining equilibrium conditions a (log)linearization is optional. In determining the decision rule coefficients, we take account of the fact that, in the next period, the constraint could either be binding or non-binding. Thus, we are capturing precautionary behavior. Importantly, the decision rule coefficients only have to be computed once (as opposed to for each period t separately as in the EFPF approach). Compared with 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 simplify computations.

The motivation for the piecewise-linear functional form is twofold. First, we have observed in a number of models that more densely parameterized nonlinear decision rules look approximately piecewise linear. For instance, in Aruoba et al. (2018) we considered a New Keynesian DSGE model and stitched together higher-order Chebyshev polynomials along the locus in the state-space where the ELB constraint becomes binding. We found that the decision rules on both sides of the kink are approximately linear. In the Online Appendix we solve a consumption-savings model with an occasionally-binding borrowing constraint and demonstrate that a global solution technique produces approximately piecewise linear decision rules. Second, in Section 3 we solve a simplified version of the New Keynesian DSGE model with an ELB constraint and show that the piecewise-linear structure is exact.

To solve the nonlinear filtering problem, we develop a conditionally optimal particle filter (COPF). A particle filter is a stochastic algorithm that 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 $\{s_t^j, W_t^j\}_{j=1}^M$. Because of its stochastic structure, repeated runs of the filter generate a distribution of likelihood values. An important property of the particle filter is that the average likelihood across repeated runs is equal to the exact likelihood (unbiasedness). The tuning of the particle filter determines the precision of the approximation. A key step in the specification of the algorithm is the mutation of time $t-1$ particle values into time t particle values. We show how, in the case of a piecewise linear DSGE model solution, the mutation step can be executed optimally, conditional on the stage $t-1$ particle values.

In a sequence of numerical illustrations based on a small-scale New Keynesian DSGE model with an ELB 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 mutates particle values by simulating the model solution forward, our 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 for COPF versus 1,000 for BSPF in our experiments), 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.

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. The ones most closely related to our work are Cuba-Borda, Guerrieri, Iacoviello, and Zhong (2019) and Atkinson et al. (2020).¹ Cuba-Borda et al. (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 data generating process (DGP). In their setting, measurement error and solution approximation error make 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 allows us to exactly identify the regime, and we modify the COPF to capture this feature.

Atkinson et al. (2020) compare the performance of a fully nonlinear solution and a variant of the BSPF for estimation, with the approximated solution using OccBin and the inversion filter. They simulate data from a DSGE model that includes more frictions and shocks than the model used for estimation, and the latter is close to the model we use in this paper in terms of size. As such, their estimated model is misspecified with respect to the DGP. 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, and allows for more general model structures and state-space representations than the inversion filter.

Based on 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

¹Boehl (2019) combines his model solution, which is a variant of the EPFP, with a variant of an ensemble Kalman filter. His paper presents an application but does not focus on accuracy comparisons of solution and estimation methods.

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 multiplier during the Great Recession, when the United States was at the ELB, was larger than in normal times when interest rates were positive, albeit with a mean estimate of 0.7 still small in absolute terms. 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 ELB constraint used in the subsequent analysis. In Section 3 we solve a simplified version of the New Keynesian model and show that the resulting decision rules are piecewise linear and continuous. We also provide a comparison to the OccBin solution. In Section 4, we describe how to impose continuity on piecewise-linear decision rules and derive a canonical form for the DSGE model solution. Section 5 discusses how the decision rule coefficients are determined to approximately satisfy the model’s equilibrium conditions. The COPF is derived in Section 6. Section 7 presents some numerical experiments to document the accuracy of the likelihood approximation through the COPF, and Section 8 contains the empirical analysis. Finally, Section 9 concludes. Derivations and further implementation details are provided in the Online Appendix. The Appendix also contains a section that shows how to solve a consumption-savings model with an occasionally-binding borrowing constraint using the techniques proposed in this paper and compares our PLC solution to an “exact” solution and a solution constructed with OccBin.

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 et al. (2018). Variants of this model have been widely studied in the literature, and its properties are discussed in detail in Woodford (2003). To make this paper self-contained and introduce some important notation, we briefly 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 .² The households maximize

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

subject to the budget constraint

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

Here β is the discount factor, d_t is an exogenous shock to the discount factor, $1/\tau$ is the intertemporal elasticity of substitution, and η is the Frisch labor supply elasticity. P_t is the price of the final good. The households receive the real wage W_t in exchange for labor services. B_t is the quantity of nominal bonds, which pay gross interest R_t . Furthermore, the households receive profits D_t from the firms and pay lump-sum taxes T_t . $S C_t$ is the net cash inflow from trading a full set of state-contingent securities.

Firms. The final-goods producers generate aggregate output Y_t aggregating intermediate goods $Y_t(j)$, $j \in [0, 1]$. Under the assumption of perfect competition and free entry, the demand for the intermediate inputs and the price of the aggregate final good are given by

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-1/\nu} Y_t \quad \text{and} \quad P_t = \left(\int_0^1 P_t(j)^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}}, \quad (2)$$

respectively. 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 adjust-

²The habit stock is proxied by the level of technology A_t , which ensures that the economy evolves along a balanced growth path. Since we will not focus on it in the subsequent analysis, we do not make a money-holding motive, such as valuing transaction services from real money balances, explicit in the description of the environment. Such a motive is necessary to make the ELB a relevant constraint in a model like this.

ment 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 ELB constraint is an important part of our analysis we introduce it explicitly as follows:

$$R_t = \max \{1, R_t^* e^{\sigma_R \epsilon_{R,t}}\}, \quad R_t^* = \left[r \pi_* \left(\frac{\pi_t}{\pi_*} \right)^{\psi_1} \left(\frac{Y_t}{\gamma Y_{t-1}} \right)^{\psi_2} \right]^{1-\rho_R} R_{t-1}^{\rho_R}, \quad (5)$$

Here R_t^* is the systematic part of monetary policy which reacts to an inflation gap and an output growth gap, r is the steady-state real interest rate, π_* is the target-inflation rate, γ is the growth rate of the economy, and $\epsilon_{R,t}$ is a monetary policy shock.

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. Aggre-

gate productivity evolves according to

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

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

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

The shock to the discount factor evolves according to

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

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. In one of the equilibria, the so-called targeted-inflation equilibrium, the endogenous variables fluctuate around the steady state in which inflation equals the value targeted by the central bank. Another important equilibrium is the so-called deflation equilibrium where the economy fluctuates around the so-called deflation steady state in which nominal interest rates are zero.³ In the remainder of the paper we mostly focus on the targeted-inflation equilibrium, though we also discuss a deflation equilibrium in Section 3. The former is essentially the equilibrium that arises in linearized New Keynesian DSGE models, adjusted for the presence of the ELB constraint. The corresponding steady state is given by

$$\pi_*, \quad r_* = \frac{\gamma}{\beta}, \quad R_* = r_* \pi_*, \quad y_* = [(1 - \nu)g_*^\tau]^{-\frac{1}{\tau+1/\eta}}, \quad c_* = \frac{y_*}{g_*}. \quad (11)$$

³See, for instance, Benhabib et al. (2001), Aruoba and Schorfheide (2016) and Aruoba et al. (2018) for a discussion of multiplicity of equilibria in this model.

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 $x_*e^{\hat{x}_t}$.⁴ Our goal is to write the equilibrium conditions as a system of expectational difference equations of the form

$$\mathbb{E}_t[\mathcal{R}(\hat{y}_t, \hat{c}_t, \hat{\pi}_t, \hat{R}_t, \hat{y}_{t+1}, \hat{c}_{t+1}, \hat{\pi}_{t+1}, \hat{R}_{t+1}, \dots)] = 0, \quad (12)$$

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

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

$$\mathcal{R}_c(\cdot) = \hat{d}_{t+1} - \hat{d}_t - \tau(\hat{c}_{t+1} - \hat{c}_t) + \hat{R}_t - \hat{\pi}_{t+1} - \hat{z}_{t+1}. \quad (13)$$

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

$$\begin{aligned} \mathcal{R}_\pi(\cdot) = & \ln \left[\frac{1}{\nu} \left(\frac{1}{\nu} + \left(1 - \frac{1}{\nu}\right) e^{\tau\hat{c}_t + \hat{y}_t/\eta} \right) - \phi\pi_*^2 (e^{\hat{\pi}_t} - 1) \left[\left(1 - \frac{1}{2\nu}\right) e^{\hat{\pi}_t} + \frac{1}{2\nu} \right] \right. \\ & \left. + \phi\beta\pi_*^2 \left(e^{\hat{d}_{t+1} - \hat{d}_t} \right) \left(e^{-\tau(\hat{c}_{t+1} - \hat{c}_t)} \right) \left(e^{\hat{y}_{t+1} - \hat{y}_t} \right) \left(e^{\hat{\pi}_{t+1}} - 1 \right) e^{\hat{\pi}_{t+1}} \right]. \end{aligned} \quad (14)$$

The aggregate resource constraint leads to

$$\mathcal{R}_y(\cdot) = \hat{y}_t - \hat{c}_t + \ln \left[\frac{1}{e^{\hat{g}_t}} - \frac{\phi}{2} g_* (\pi_* e^{\hat{\pi}_t} - \bar{\pi})^2 \right]. \quad (15)$$

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

$$\mathcal{R}_R(\cdot) = \hat{R}_t - \max \left\{ (1 - \rho_R) [\psi_1 \hat{\pi}_t + \psi_2 (\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t)] + \rho_R \hat{R}_{t-1} + \sigma_R \epsilon_{R,t}, -\ln(r\pi_*) \right\}. \quad (16)$$

We stack the residual functions for the exogenous shocks as follows:

$$\mathcal{R}_{exo}(\cdot) = \begin{bmatrix} \hat{z}_t - \rho_z \hat{z}_{t-1} - \sigma_z \epsilon_{z,t} \\ \hat{d}_t - \rho_d \hat{d}_{t-1} - \sigma_d \epsilon_{d,t} \\ \hat{g}_t - \rho_g \hat{g}_{t-1} - \sigma_g \epsilon_{g,t} \\ \epsilon_{R,t} - \sigma_R \epsilon_{R,t} \end{bmatrix}. \quad (17)$$

⁴Introducing \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.

3 Solving a Simplified Version of the DSGE Model

In order to highlight some important features of the proposed solution method, we first consider a highly simplified version of the DSGE model introduced in Section 2 that can be (almost) solved analytically.⁵ In particular, we will show that the PLC form emerges as the exact solution to the simplified DSGE model. We also discuss the multiplicity of solutions and provide a comparison to OccBin.

On the model described in the previous section we impose the parameter restrictions $\tau = 1$, $\gamma = 1$, $\eta = \infty$, $g_* = 1$, $\bar{\pi} = \pi_*$, $\psi_1 = \psi$, $\psi_2 = 0$, $\rho_R = 0$, $\sigma_z = 0$, $\sigma_g = 0$, and $\rho_d = 0$. We log-linearize the equilibrium conditions (except for the ELB constraint) around the targeted inflation steady state in (11) and regard the resulting equations as *the model* to be solved. Some details of the calculations are relegated to the Online Appendix.

Equilibrium Conditions. The residual functions (13), (14), and (16) simplify to

$$\begin{aligned}\mathcal{R}_c(\cdot) &= \hat{d}_{t+1} - \hat{d}_t - (\hat{c}_{t+1} - \hat{c}_t) + \hat{R}_t - \hat{\pi}_{t+1} \\ \mathcal{R}_\pi(\cdot) &= \hat{\pi}_t - \beta\hat{\pi}_{t+1} - \kappa\hat{c}_t \\ \mathcal{R}_R(\cdot) &= \hat{R}_t - \max \left\{ \psi\hat{\pi}_t + \sigma_R\epsilon_{R,t}, -\ln(r\pi_*) \right\}\end{aligned}\tag{18}$$

Under parameterizations in which monetary policy is active, i.e., $\psi > 1$, the model has typically two stationary solution in which $(\hat{R}_t, \hat{c}_t, \hat{\pi}_t)$ are independently and identically distributed (iid) over time. These correspond to the targeted-inflation and the deflation equilibria defined in the previous section.

In both equilibria expected consumption and inflation are time invariant and can be replaced by $\mu_c = \mathbb{E}_t[\hat{c}_{t+1}]$ and $\mu_\pi = \mathbb{E}_t[\hat{\pi}_{t+1}]$. Thus, setting the expected value of the residual function to zero, and conducting a few basic algebraic manipulations, we obtain:

$$\begin{aligned}\hat{R}(\epsilon_{d,t}, \epsilon_{R,t}) &= \max \left\{ \frac{1}{1 + \psi\kappa} \left[\psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi + \psi\kappa\sigma_d\epsilon_{d,t} + \sigma_R\epsilon_{R,t} \right], -\ln(r\pi_*) \right\} \\ \hat{c}(\epsilon_{d,t}, \epsilon_{R,t}) &= -\hat{R}(\epsilon_{d,t}, \epsilon_{R,t}) + \mu_c + \mu_\pi + \sigma_d\epsilon_{d,t} \\ \hat{\pi}(\epsilon_{d,t}, \epsilon_{R,t}) &= -\kappa\hat{R}(\epsilon_{d,t}, \epsilon_{R,t}) + \kappa\mu_c + (\kappa + \beta)\mu_\pi + \kappa\sigma_d\epsilon_{d,t}\end{aligned}\tag{19}$$

Here we replaced \hat{R}_t , \hat{c}_t , and $\hat{\pi}_t$ by decision rules that are time-invariant functions of the state variables $(\epsilon_{d,t}, \epsilon_{R,t})$.

⁵A similar model was solved in Mendes (2011).

Constructing a Solution. In order to solve the system (19) we need to find constants μ_c and μ_π such that

$$\mu_c = \mathbb{E}[c(\epsilon_{d,t}, \epsilon_{R,t})] \quad \text{and} \quad \mu_\pi = \mathbb{E}[\pi(\epsilon_{d,t}, \epsilon_{R,t})].$$

Because of the max operator in the monetary policy rule, this requires the computation of the mean of a truncated Normal random variable. We rotate the vector of innovations $[\epsilon_{d,t}, \epsilon_{R,t}]'$ to separate the component that enters the monetary policy rule from a second component that is orthogonal:

$$\eta_{1,t} = \frac{1}{\sigma_\eta}(\psi\kappa\sigma_d\epsilon_{d,t} + \sigma_R\epsilon_{R,t}), \quad \eta_{2,t} = \frac{1}{\sigma_\eta}(\sigma_R\epsilon_{d,t} - \psi\kappa\sigma_d\epsilon_{R,t}), \quad \sigma_\eta = \sqrt{(\psi\kappa\sigma_d)^2 + \sigma_R^2}. \quad (20)$$

By construction, the innovations $\eta_{1,t}$ and $\eta_{2,t}$ are also $N(0, 1)$.

Using the expression for $\eta_{1,t}$ in (20) we can rewrite the interest rate rule as

$$R(\eta_{1,t}) = \max \left\{ \frac{1}{1 + \psi\kappa} \left[\psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi + \sigma_\eta\eta_{1,t} \right], -\ln(r\pi_*) \right\}. \quad (21)$$

Define the cutoff value

$$\bar{\eta}_1 = -\frac{1}{\sigma_\eta} \left[(1 + \psi\kappa) \ln(r\pi_*) + \psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi \right] \quad (22)$$

such that $R(\eta_{1,t}) = -\ln(r\pi_*)$ whenever $\eta_{1,t} \leq \bar{\eta}_1$. Using the formula for the mean of a truncated standard normal random variable, we obtain

$$\begin{aligned} \mathbb{E}[R(\eta_{1,t})] &= -\Phi_N(\bar{\eta}_1) \ln(r\pi_*) \\ &\quad + \frac{1}{1 + \psi\kappa} \left[(1 - \Phi_N(\bar{\eta}_1)) (\psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi) + \sigma_\eta\phi_N(\bar{\eta}_1) \right]. \end{aligned} \quad (23)$$

Here $\Phi_N(\cdot)$ and $\phi_N(\cdot)$ are the cumulative density function (cdf) and the probability density function (pdf) of a standard Normal random variable. Taking expectations of the second and third equation in (19) and substituting out $\mathbb{E}[\hat{R}(\epsilon_{d,t}, \epsilon_{R,t})] = \mathbb{E}[R(\eta_{1,t})]$ using (23) leads to the following nonlinear system of equations:

$$\begin{aligned} \mu_c &= \frac{1 - \Phi_N(\bar{\eta}_1)}{1 + \psi\kappa} \left[\mu_c + (1 - \psi\beta)\mu_\pi \right] - \frac{\sigma_\eta\phi_N(\bar{\eta}_1)}{1 + \psi\kappa} + \Phi_N(\bar{\eta}_1) \left[\ln(r\pi_*) + \mu_c + \mu_\pi \right]. \\ \mu_\pi &= \frac{1 - \Phi_N(\bar{\eta}_1)}{1 + \psi\kappa} \left[\kappa\mu_c + (\kappa + \beta)\mu_\pi \right] - \frac{\kappa\sigma_\eta\phi_N(\bar{\eta}_1)}{1 + \psi\kappa} + \Phi_N(\bar{\eta}_1) \left[\kappa \ln(r\pi_*) + \kappa\mu_c + (\kappa + \beta)\mu_\pi \right], \end{aligned} \quad (24)$$

where $\bar{\eta}_1$ is given in (22). Conditional on $\bar{\eta}_1$ the system is linear in (μ_c, μ_π) which means that it can be reduced to a single nonlinear equation in terms of $\bar{\eta}_1$ that needs to be solved numerically. This equation typically has two solutions, which generate the targeted-inflation and deflation equilibrium, respectively.

Properties of the Solution. (i) *The decision rules are piecewise-linear and continuous for each (μ_c, μ_π) .* Plugging the expression for $\hat{R}(\epsilon_{d,t}, \epsilon_{R,t}) = R(\eta_{1,t})$ in (23) into the consumption and inflation decision rules in (19) and using the relationship between the ϵ_t 's and η_t 's in (20), we can rewrite the decision rules for consumption and inflation as follows:

$$\begin{aligned} c(\eta_{1,t}, \eta_{2,t}) &= \begin{cases} \frac{1}{1+\psi\kappa} \left[\mu_c + (1-\psi\beta)\mu_\pi \right] + (\gamma_{d,1} - \frac{1}{1+\psi\kappa})\sigma_\eta\eta_{1,t} + \gamma_{d,2}\sigma_\eta\eta_{2,t} & \text{if } \eta_{1,t} > \bar{\eta}_1 \\ \ln(r\pi_*) + \mu_c + \mu_\pi + \gamma_{d,1}\sigma_\eta\eta_{1,t} + \gamma_{d,2}\sigma_\eta\eta_{2,t} & \text{otherwise} \end{cases} \\ \pi(\eta_{1,t}, \eta_{2,t}) &= \begin{cases} \frac{1}{1+\psi\kappa} \left[\kappa\mu_c + (\kappa + \beta)\mu_\pi \right] + \kappa(\gamma_{d,1} - \frac{1}{1+\psi\kappa})\sigma_\eta\eta_{1,t} + \kappa\gamma_{d,2}\sigma_\eta\eta_{2,t} & \text{if } \eta_{1,t} > \bar{\eta}_1 \\ \kappa \ln(r\pi_*) + \kappa\mu_c + (\kappa + \beta)\mu_\pi + \kappa\gamma_{d,1}\sigma_\eta\eta_{1,t} + \kappa\gamma_{d,2}\sigma_\eta\eta_{2,t} & \text{otherwise} \end{cases} \end{aligned} \quad (25)$$

Here, the γ constants are functions of the structural parameters obtained from (20) by rearranging the equations to express the ϵ_t 's as a function of the η_t 's. The ELB becomes binding at the locus in the state space defined by

$$\bar{\eta}_1 = \eta_{1,t} = \frac{1}{\sigma_\eta} (\psi\kappa\sigma_d\epsilon_{d,t} + \sigma_R\epsilon_{R,t}). \quad (26)$$

The second innovation, $\eta_{2,t}$, does not enter the policy rule and therefore cannot push the economy toward the ELB. The decision rules are piecewise linear functions of the innovations $\eta_{1,t}$ and $\eta_{2,t}$. The slope coefficients associated with $\eta_{1,t}$ change when the economy hits the ELB at $\eta_{1,t} = \bar{\eta}_1$, whereas the slope coefficients for $\eta_{2,t}$ do not change. We verify in the Online Appendix that the decision rules are also continuous at $\eta_{1,t} = \bar{\eta}_1$, meaning that for each $\eta_{2,t}$ and for each sequence $\eta_{1,t}^n \rightarrow \bar{\eta}_1$ as $n \rightarrow \infty$:

$$|c(\eta_{1,t}^n, \eta_{2,t}) - c(\bar{\eta}_1, \eta_{2,t})| \rightarrow 0 \quad \text{and} \quad |\pi(\eta_{1,t}^n, \eta_{2,t}) - \pi(\bar{\eta}_1, \eta_{2,t})| \rightarrow 0. \quad (27)$$

Because the transformation between $(\epsilon_{d,t}, \epsilon_{R,t})$ and $(\eta_{1,t}, \eta_{2,t})$ is continuous, the continuity result also holds for the decision rules $\hat{c}(\epsilon_{d,t}, \epsilon_{R,t})$ and $\hat{\pi}(\epsilon_{d,t}, \epsilon_{R,t})$, expressed in terms of the original state variable, as $(\psi\kappa\sigma_d\epsilon_{d,t}^n + \sigma_R\epsilon_{R,t}^n)/\sigma_\eta \rightarrow \bar{\eta}_1$.

(ii) *The law of motion given by the interest rate rule (21) and the consumption and inflation decision rules (25) is coherent and complete for each (μ_c, μ_π) and each realization*

of the innovations $(\eta_{1,t}, \eta_{2,t})$. The system takes the form of a linear simultaneous equations model with regime switches. The concepts of coherency and completeness were introduced by Gourieroux et al. (1980) and more recently studied in the context of ELB applications by Mavroeidis (2020) and Ascari and Mavroeidis (2020). Coherency requires that given an innovation $(\eta_{1,t}, \eta_{2,t})$ there exists a solution to the system of equations. Completeness refers to the uniqueness of that solution. Define

$$R_t^{(1)} = \frac{1}{1 + \psi\kappa} \left[\psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi + \sigma_\eta\eta_{1,t} \right] \quad \text{and} \quad R_t^{(2)} = -\ln(r\pi_*).$$

Coherency and completeness require that: (a) $\eta_{1,t} > \bar{\eta}_1$ implies $R_t = R_t^{(1)}$ is a solution to the maximization on the right-hand side of (23), whereas $R_t = R_t^{(2)}$ is not. Likewise, it is required that (b) $\eta_{1,t} < \bar{\eta}_1$ implies $R_t = R_t^{(2)}$ is a solution to the maximization on the right-hand side of (23), whereas $R_t = R_t^{(1)}$ is not. Coherency and completeness follows directly from the linearity of $R_t^{(1)}$ with respect to $\eta_{1,t}$ and the definition of $\bar{\eta}_1$ in (22) and require no further restrictions on the domain of the innovations.

(iii) *Solutions for μ_c and μ_π .* The means μ_c and μ_π together with the cutoff value $\bar{\eta}_1$ are determined by the nonlinear system of equations (22) and (24). To understand the properties of the nonlinear system, assume that $\psi > 1$, i.e., monetary policy is active, and $\sigma_\eta = 0$, i.e., there is no uncertainty. First, suppose we start with the conjecture that $\bar{\eta}_1 = -\infty$. Then (24) simplifies to

$$\mu_c = \frac{1}{1 + \psi\kappa} \left[\mu_c + (1 - \psi\beta)\mu_\pi \right], \quad \mu_\pi = \frac{1}{1 + \psi\kappa} \left[\kappa\mu_c + (\kappa + \beta)\mu_\pi \right]$$

which is solved by

$$\mu_c = 0, \quad \mu_\pi = 0. \tag{28}$$

Recall that the system was expressed in deviations from the targeted-inflation steady state. Thus if $\mu_c = \mu_\pi = 0$, then the means of consumption and inflation are equal to the steady state. The analysis is completed by noting that (22) implies that indeed $\bar{\eta}_1 = -\infty$ as initially assumed.

Second, suppose we start from the conjecture that $\bar{\eta}_1 = +\infty$. Then (24) simplifies to

$$\mu_c = \ln(r\pi_*) + \mu_c + \mu_\pi, \quad \mu_\pi = \kappa \ln(r\pi_*) + \kappa\mu_c + (\kappa + \beta)\mu_\pi,$$

which implies

$$\mu_c = -\frac{1}{\kappa}(1 - \beta) \ln(r\pi_*), \quad \mu_\pi = -\ln(r\pi_*). \quad (29)$$

Substituting the means into (22), we obtain $\bar{\eta}_1 = (\psi - 1) \ln(r\pi_*)/\sigma_\eta > 0$. Thus, as $\sigma_\eta \rightarrow 0$, $\bar{\eta}_1 \rightarrow +\infty$ as required. In this case, the system is in the so-called deflation steady state and the ELB constraint is always binding.

Once we allow for uncertainty, $\sigma_\eta > 0$, then the means in (28) and (29) no longer solve the system of equations (22) and (24). However, for values of σ_η that are not “too large,” one can obtain solutions that are “close” to the ones derived above. Due to the nonlinearity of the $\Phi_N(\cdot)$ and $\phi_N(\cdot)$ functions, these solutions can only be computed numerically.

Comparison to OccBin Solution. It is instructive to compare the above solution to the one generated by OccBin. OccBin requires the choice of a reference regime. Because subsequently we focus on the targeted inflation equilibrium in which the ELB is non-binding with high probability, we impose that the ELB is non-binding in the reference regime. Consider a generic period t . The OccBin solution is based on the assumption that for $\tau \geq T$ the economy will be in the reference regime. Given the lack of dynamics in the simple model, we can choose $T = t + 1$ and assume that the regime will remain in the targeted-inflation steady state: $\hat{R}_{t+1} = 0$, $\hat{c}_\tau = 0$, and $\hat{\pi}_\tau = 0$. For period t , the algorithm draws the shocks $(\epsilon_{d,t}, \epsilon_{R,t})$ and solves the system

$$\begin{aligned} \hat{R}_t &= \max \{ \psi \hat{\pi}_t + \epsilon_{R,t}, -\ln(r\pi_*) \} \\ \hat{c}_t &= \mathbb{E}_t[\hat{c}_{t+1}] - \hat{R}_t + \mathbb{E}_t[\hat{\pi}_{t+1}] \\ \hat{\pi}_t &= \beta \mathbb{E}_t[\hat{\pi}_{t+1}] + \kappa \hat{c}_t \end{aligned}$$

under the restriction that $\mathbb{E}_t[\hat{c}_{t+1}] = 0$ and $\mathbb{E}_t[\hat{\pi}_{t+1}] = 0$. The solution is identical to (21) and (25) with $\mu_c = \mu_\pi = 0$ imposed. Thus, in the context of our stylized model the difference between our solution, which happens to be exact, and the OccBin solution is that the latter does not take into account the uncertainty about the regime in period $t + 1$.

Summary. We draw the following conclusions from the analysis of the simplified DSGE model. First, the log-linearized model has static solutions that are exact and in which the decision rules for consumption and inflation are piecewise-linear and continuous. We use this as a motivation for subsequently considering a class of approximate solutions with PLC decision rules for richer nonlinear DSGE models with occasionally-binding constraints. Second, although we approximated the equilibrium conditions around the targeted-inflation

steady state, because of the nonlinearity generated by the occasionally-binding constraint, for $\sigma_\eta > 0$ the decision rules do not pass through the steady state around which the model was initially approximated. In fact, because the decision rules are flexible enough to have unrestricted intercepts, we can also generate the deflation equilibrium which is far away from the targeted-inflation steady state. Third, an important difference between our proposed PLC solution and the OccBin solution is that our decision rule coefficients capture uncertainty about the future.

4 PLC Decision Rules and the Canonical Form

In the analysis of the simplified model in the previous section, the PLC decision rules emerged from the analytical solution of the model. For more elaborate DSGE models, we will parameterize a family of piecewise-linear decision rules and then impose coefficient restrictions that guarantee that the decision rules are continuous at the kink, where the constraint changes from being slack to being binding. The remaining free coefficients of the decision rules can then be used to (approximately) satisfy the equilibrium conditions of the model by setting the residual functions (close) to zero. A discussion of how to do this numerically is deferred to Section 5.

4.1 PLC Decision Rules

Let $\mathbb{X} = [x_1, X_2'] \in \mathcal{X}$ be an $n \times 1$ vector of non-redundant state variables. We assume that \mathbb{X} also contains a constant. Here x_1 is one particular element of \mathbb{X} that enters the characterization of the locus of points in the state space at which the constraint becomes binding and the decision rules have their kink. The reason for separating out one of the \mathbb{X} elements will become clear below. Let Y denote a $k \times 1$ vector of control variables. As we make explicit below, we assume Y depends on x_1 and X_2 linearly where the coefficients may depend on whether the constraint is binding or not.

We assume that there is a linear(ized) scalar-valued function $h(x_1, X_2, Y)$ that determines whether the constraint is binding:

$$h(x_1, X_2, Y) = \begin{cases} > 0 & \text{if constraint is non-binding (n)} \\ \leq 0 & \text{if constraint is binding (b)} \end{cases} . \quad (30)$$

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

$$h(x_1, X_2, y) = \gamma_1 x_1 + \gamma_2' X_2 + \gamma_Y' Y. \quad (31)$$

The γ 's are not free coefficients. They are obtained from the equilibrium conditions of the DSGE model. In the simplified model of Section 3 we set $x_1 = 1$, $X_2 = [\epsilon_d, \epsilon_R]'$, and $y = [\hat{c}, \hat{\pi}]'$ such that $h(x_1, X_2, y) = \ln(r\pi_*) + \epsilon_R + \psi\hat{c}$ with $\gamma_1 = \ln(r\pi_*)$, $\gamma_2' = [0, 1]'$, and $\gamma_Y' = [0, \psi]$.

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 $h(\ell(X_2), X_2, Y(\ell(X_2), X_2)) = 0$, that is, the constraint is just binding. Here $Y(\cdot)$ denotes the assumed piecewise linear decision rules for the control variables. The linearity of $h(\cdot)$ in (31) 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

$$\ell(X_2) = \delta' X_2, \quad (32)$$

where δ is a $(n-1) \times 1$. The δ coefficients will be determined as functions of the decision rule coefficients and the coefficients of the constraint function $h(\cdot)$. So far we have not yet made a determination whether the constraint is slack if $x_1 < \delta' X_2$. In the simplified model the equilibrium kink function is given by (26).

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

$$Y^i(x_1, X_2) = \begin{cases} \alpha_{1,1}^i x_1 + \alpha_{1,2}^i X_2 & \text{if } x_1 \geq \ell(X_2) \\ \alpha_{2,1}^i x_1 + \alpha_{2,2}^i X_2 & \text{if } x_1 < \ell(X_2) \end{cases} \quad i = 1, \dots, k, \quad (33)$$

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 (33) makes the benefit of using the kink function $\ell(\cdot)$ clear: given the state variables x_1 and X_2 , we can easily determine on which side of the constraint we need to be, even when the constraint contains some control variables. In the simplified model the equilibrium decision rules are given by (25). Once we replace the η_t 's by the ϵ_t 's we obtain the same form as (33), where the α coefficients were determined such that the decision rules satisfy the equilibrium conditions of the model.

4.2 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. 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 :

$$\alpha_{1,1}^i \delta' + \alpha_{1,2}^i = \alpha_{2,1}^i \delta' + \alpha_{2,2}^i. \quad (34)$$

Next, we impose restrictions that make the $\ell(\cdot)$ and $Y(\cdot)$ functions consistent with the constraint in (30). 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 set of $(n - 1)$ restrictions:

$$\gamma_1 \delta' + \gamma_2 + \sum_{i=1}^k \gamma_Y^i (\alpha_{1,1}^i \delta' + \alpha_{1,2}^i) = 0. \quad (35)$$

Counting all unknowns and restrictions, we have $k(n + 1)$ degrees of freedom.⁶ 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 $k(n + 1)$

$$\vartheta = [\alpha_{1,1}^1, \dots, \alpha_{1,1}^k, \alpha_{1,2}^1, \dots, \alpha_{1,2}^k, \alpha_{2,1}^1, \dots, \alpha_{2,1}^k]'. \quad (36)$$

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”

⁶There 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 (34) and the $(n - 1)$ restrictions in (35), we get $(k + 1)(n - 1)$ restrictions. Subtracting the number of restrictions from the number of unknowns, we get $k(n + 1)$.

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. In our application the choice of which coefficients go in to ϑ is driven by numerical considerations. The “1” regime corresponds to the ELB not being binding. In the targeted-inflation equilibrium, this is the more likely regime and a good starting value for the numerical procedure that is used to determine ϑ is given by the decision-rule coefficients of a log-linear approximation that ignores the ELB constraint.

Conditional on ϑ , we can rewrite (35) 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

$$\delta'(\vartheta) = \frac{1}{a(\vartheta)} B'(\vartheta), \quad (37)$$

where $a(\vartheta)$ is a scalar and $B(\vartheta)$ is a $(n - 1)$ -dimensional vector. By combining (37) with (34) we obtain an expression for the constrained decision rule coefficients $\alpha_{2,2}^i$:

$$\alpha_{2,2}^{i'}(\vartheta) = (\alpha_{1,1}^i - \alpha_{2,1}^i) \left(\frac{1}{a(\vartheta)} B(\vartheta) \right) + \alpha_{1,2}^{i'}. \quad (38)$$

The last step is to determine which part of the decision rule in (33) 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

$$\begin{aligned} h[x_1, X_2, Y(x_1, X_2)] &= \gamma_1 x_1 + \gamma_2' X_2 + \sum_{i=1}^k \gamma_Y^i (\alpha_{1,1}^i x_1 + \alpha_{1,2}^{i'} X_2) \\ &= \underbrace{\left(\gamma_1 + \sum_{i=1}^k \gamma_Y^i \alpha_{1,1}^i \right)}_{c(\vartheta)} x_1 + \underbrace{\left(\gamma_2' + \sum_{i=1}^k \gamma_Y^i \alpha_{1,2}^{i'} \right)}_{D'(\vartheta)} X_2, \end{aligned} \quad (39)$$

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. Collecting the δ terms, we notice that the restriction in (35) implies $c(\vartheta)\delta'(\vartheta) + D'(\vartheta) = 0$, or $D'(\vartheta) = -c(\vartheta)\delta'(\vartheta)$. Using this

on (39), we get

$$h[x_1, X_2, Y(x_1, X_2)] = c(\vartheta)[x_1 - \delta'(\vartheta)X_2]. \quad (40)$$

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 (33) corresponds to the constraint being slack.

4.3 Example: The Full New Keynesian Model

We now adapt the generic notation so far to the New Keynesian model with ELB constraint described 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 $n = 7$.⁷ 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 $\hat{\pi}_t = \hat{\pi}(x_1, X_2)$, $\hat{c}_t(x_1, X_2)$ follows from solving for \hat{c}_t in $\mathcal{R}_y(\cdot)$ in (15) and $\hat{R}_t(x_1, X_2)$ follows from solving for \hat{R}_t in $\mathcal{R}_R(\cdot)$ in (16). Thus, we set $Y(\cdot) = [\hat{y}(\cdot), \hat{\pi}(\cdot)]'$ and $k = 2$.

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

$$h(x_{1,t}, X_{2,t}, Y_t(\cdot)) = (1 - \rho)[\psi_1\hat{\pi}(\cdot) + \psi_2(\hat{y}(\cdot) - \hat{y}_{t-1} + \hat{z}_t)] + \rho_R\hat{R}_{t-1} + e_{R,t} + \ln(r_*\pi_*). \quad (41)$$

Thus, the γ coefficients in (31) are

$$\gamma_1 = \rho_R, \quad \gamma'_2 = [\ln(r_*\pi_*), -(1 - \rho_R)\psi_2, (1 - \rho_R)\psi_2, 0, 0], \quad \gamma'_Y = [(1 - \rho_R)\psi_1, (1 - \rho_R)\psi_2],$$

and we can write $c(\vartheta)$ in (39) as

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

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”

⁷In 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 being equal, higher lagged interest rates move the economy away from the ELB constraint.

regime as the regime where the ELB is slack, “n” (non-binding), and the “2” regime as the “b” (binding) regime. We check that $c(\vartheta)$ is indeed positive every time we solve the model.

4.4 Canonical Form

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

$$s_t = \begin{cases} \Phi_0(n) + \Phi_1(n)s_{t-1} + \Phi_\eta(n)\eta_t & \text{if } \eta_{1,t} < \zeta(s_{t-1}) \\ \Phi_0(b) + \Phi_1(b)s_{t-1} + \Phi_\eta(b)\eta_t & \text{otherwise,} \end{cases} \quad (42)$$

which is a VAR for s_t with endogenous regime switching. The innovations η_t are a function of the structural innovations ϵ_t . This transformation is done so that the first element of η_t , $\eta_{1,t}$, is a linear combination of structural shocks that determines whether the constraint is binding in period t . Precise definitions of η_t , the impact matrices $\Phi_\eta(\cdot)$ and the threshold function $\zeta(\cdot)$ will be provided below. Recall that we used the rotated η_t shocks also in the construction of the solution to the simplified DSGE model in Section 3 and its decision rules (25) were written in the canonical form.

Equation (42) 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.⁸ Whether the coherency and completeness conditions are satisfied – recall that we showed in Section 3 that they are satisfied for the simplified New Keynesian model – depends on the structure of the $\Phi(\cdot)$ matrices and is model specific. The construction of the canonical form for the New Keynesian DSGE model of Section 2 is outlined in the Online Appendix.

⁸Aruoba et al. (2020) estimate a structural VAR that takes the form of (42). Chen (2017) and Bianchi and Melosi (2017) use an exogenous regime-switching process to characterize the ELB dynamics. Such models can be solved using the tools proposed by Farmer, Waggoner, and Zha (2011). Benigno, Foerster, Otrok, and Rebucci (2016) endogenize the regime-switching probability in a model of financial crisis, but, unlike in our paper, the transition from one to the other regime remains partly decoupled from the realization of the fundamental shocks.

4.5 Measurement Equations

The key requirement for the conditionally optimal particle filter that is developed in Section 6 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 (42) and the measurement equation is linear in s_t as in

$$y_t^o = A_0 + A_s s_t + u_t, \quad u_t \sim N(0, \varsigma \Sigma_u), \quad (43)$$

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 using 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 (A.6), 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 (42). 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

$$\begin{aligned} y_{gr,t}^o &= 100 \ln(\gamma) + 100(\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t) + \sigma_{u,y} u_{y,t} \\ \pi_t^o &= 400 \ln(\pi_*) + 400 \hat{\pi}_t + \sigma_{u,\pi} u_{\pi,t} \\ R_t^o &= 400 \ln(R_*) + 400 \hat{R}_t + \sigma_{u,R} u_{R,t}. \end{aligned} \quad (44)$$

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 $100 \cdot \ln(1 - G_t/Y_t)$. Because in our model $1 - G_t/Y_t = 1/g_t$, we obtain the additional measurement equation

$$cy_t^o = -100 \ln g_* - 100 \hat{g}_t + \sigma_{u,c} u_{c,t}. \quad (45)$$

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.

5 PLC Model Solution

In this section, we describe how the free coefficients ϑ in the PLC decision rules, as defined in (36), are determined. We first discuss the equilibrium conditions that define the objective function that we will minimize to find the optimal ϑ coefficients. We then describe the choice of the solution grid, the integration method and the optimization.

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

$$\mathcal{H}[f_0(\cdot), \mathbb{X}] = 0, \quad \forall \mathbb{X} \in \mathcal{X}, \quad (46)$$

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 (46) becomes

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

where we explained how we construct the decision rules $\hat{y}(\cdot)$, $\hat{c}(\cdot)$, $\hat{\pi}(\cdot)$, and $\hat{R}(\cdot)$ in Section 4.3. Expectations over \mathbb{X}_{t+1} can be evaluated by using the law of motion of the exogenous shocks in (17) and noting that the first three elements of \mathbb{X}_{t+1} , \hat{R}_t , 1, and \hat{y}_t , are known at time t . 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 (46) 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:

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

In the simplified DSGE model in Section 3 we were able to find two sets of decision rule coefficients ϑ that set $\mathcal{H}[\cdot]$ exactly equal to zero for all \mathbb{X} and generate what we called the targeted-inflation equilibrium and the deflation equilibrium. In the numerical illustrations of Section 7 and the empirical application of Section 8 we will focus on the targeted-inflation equilibrium by choosing starting values for the ϑ optimization that generate the decision

rules for a linearized version of the model without ELB constraint.⁹

Solution Grid. There are two popular ways to choose the solution grid. In the collocation approach, the grid 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 represent 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 collocation approach that uses tensor grids, 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.

In constructing the grid \mathcal{S} , we follow Judd et al. (2014) and build a sparse Smolyak grid.¹⁰ 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.

⁹While some progress has been made in Ascari and Mavroeidis (2020) studying parameter and innovation domain restrictions that guarantee coherency and completeness, formal results for general DSGE models with endogenous state variables and continuously distributed innovations remain elusive.

¹⁰One interpretation of our approach is that we are using the sum squared residual over the Smolyak grid as a proxy for integrating the squared residual function over the ergodic distribution. Monte Carlo experiments in Heiss and Winschel (2008) in the context of the calculation of the likelihood function of a mixed logit model, which also involves evaluating an integral without a closed-form expression, show that using a Smolyak grid provides superior performance over simulation techniques.

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 and goes 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 observed nominal interest rate data. Because we want to analyze the ELB, the grid is scaled so that its minimum value matches the ELB with $\hat{R}_{t-1} = -\ln(r_*\pi_*)$, which happens to be the 10th percentile in the data. 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 (12) 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 a 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 decision rules from a 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 et al. (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' abilities to gather, retain, and process decision-relevant information. Boundedly-rational agents may also be unable to solve a complicated mathematical problem. 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.

6 Particle Filters for PLC Models

The state-space representation associated with the PLC solution comprises the nonlinear transition equation (42) and the linear measurement equation (43). The state-space representation provides a joint density for the states s_t and the observations y_t (omitting the o superscript):

$$p(Y_{1:T}, S_{1:T}|\theta) = \prod_{t=1}^T p(y_t|s_t, \theta)p(s_t|s_{t-1}, \theta), \quad (47)$$

where $Y_{t_1:t_2}$ and $S_{t_1:t_2}$ denote the sequences y_{t_1}, \dots, y_{t_2} and s_{t_1}, \dots, s_{t_2} and θ is the vector of model parameters. 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

$$p(Y_{1:T}|\theta) = \prod_{t=1}^T p(y_t|Y_{1:t-1}, \theta) = \prod_{t=1}^T \int \int p(y_t|s_t, \theta)p(s_t|s_{t-1}, \theta)p(s_{t-1}|Y_{1:t-1}, \theta)ds_t ds_{t-1}. \quad (48)$$

These objects can be obtained from a nonlinear filter. We describe below how $p(s_t|Y_{1:t}, \theta)$ and $p(y_t|Y_{1:t-1}, \theta)$ can be efficiently approximated by a particle filter.¹¹

6.1 Generic Particle Filter

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. Following the notation in Herbst and Schorfheide (2015) we now use $h(\cdot)$ to denote a function of s_t for which a posterior expectation is supposed to be evaluated. $\mathbb{E}[h(s_t)|Y_{1:t}, \theta]$ can be approximated by Monte Carlo averages of the

¹¹Surveys 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. Textbook treatments of the statistical theory underlying particle filters can be found in Cappé, Moulines, and Ryden (2005), Liu (2001), and Del Moral (2013).

form $\frac{1}{M} \sum_{j=1}^M h(s_t^j) W_t^j$. The approximation typically holds in the form of a Law of Large Numbers and a Central Limit Theorem. The particle filter can be implemented using the following algorithm:¹²

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

$$\omega_t^j = p(\tilde{s}_t^j | s_{t-1}^j, \theta) / g_t(\tilde{s}_t^j | s_{t-1}^j, \theta). \quad (49)$$

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

$$\tilde{w}_t^j = p(y_t | \tilde{s}_t^j, \theta) \omega_t^j. \quad (50)$$

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

$$\hat{p}(y_t | Y_{1:t-1}, \theta) = \frac{1}{M} \sum_{j=1}^M \tilde{w}_t^j W_{t-1}^j. \quad (51)$$

(c) Define the normalized weights

$$\tilde{W}_t^j = \tilde{w}_t^j W_{t-1}^j / \frac{1}{M} \sum_{j=1}^M \tilde{w}_t^j W_{t-1}^j. \quad (52)$$

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

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

$$\ln \hat{p}(Y_{1:T} | \theta) = \sum_{t=1}^T \ln \left(\frac{1}{M} \sum_{j=1}^M \tilde{w}_t^j W_{t-1}^j \right). \quad (53)$$

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

¹²Exposition and notation are based on Herbst and Schorfheide (2015).

6.2 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.

6.3 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

$$g_t^*(\tilde{s}_t|s_{t-1}^j, \theta) = p(\tilde{s}_t|y_t, s_{t-1}^j, \theta) \propto p(y_t|\tilde{s}_t, \theta)p(\tilde{s}_t|s_{t-1}^j, \theta). \quad (54)$$

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 (49) and the incremental weights $\tilde{\omega}_t^j$ in (50), we obtain

$$\tilde{\omega}_t^j = \frac{p(y_t|\tilde{s}_t^j, \theta)p(\tilde{s}_t^j|s_{t-1}^j, \theta)}{p(\tilde{s}_t^j|y_t, s_{t-1}^j, \theta)} = p(y_t|s_{t-1}^j). \quad (55)$$

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 $\tilde{\omega}_t^j$ in (50).

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 (42). 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:

$$\begin{aligned}
\nu_t^j(\cdot) &= y_t - A_0 - A_s(\Phi_0(\cdot) - \Phi_1(\cdot)s_{t-1}^j) \\
\bar{\eta}_t^j(\cdot) &= (\varsigma I + \Phi'_\eta(\cdot)A'_s\Sigma_u^{-1}A_s\Phi_\eta(\cdot))^{-1}\Phi'_\eta(\cdot)A'_s\Sigma_u^{-1}\nu_t^j(\cdot) \\
\bar{\Omega}(\cdot) &= \varsigma(\varsigma I + \Phi'_\eta(\cdot)A'_s\Sigma_u^{-1}A_s\Phi_\eta(\cdot))^{-1}.
\end{aligned} \tag{56}$$

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—that is, for $\zeta(s_{t-1}^j)$ being $+\infty$ or $-\infty$. Moreover, let

$$\begin{aligned}
D_t^j(n) &= (2\pi)^{-n_y/2}|\Sigma_u|^{-1/2}|\varsigma I + \Phi_\eta(n)'A'_s\Sigma_u^{-1}A_s\Phi_\eta(n)|^{1/2} \\
&\quad \times \exp\left\{-\frac{1}{2}\nu_t^j(n)'(\varsigma\Sigma_u + A_s\Phi_\eta(n)\Phi'_\eta(n)A'_s)^{-1}\nu_t^j(n)\right\} \\
&\quad \times \Phi_N\left(\left(\zeta(s_{t-1}) - \bar{\eta}_{1,t}^j(n)\right)/\sqrt{\bar{\Omega}_{11}(n)}\right), \\
D_t^j(b) &= (2\pi)^{-n_y/2}|\Sigma_u|^{-1/2}|\varsigma I + \Phi_\eta(b)'A'_s\Sigma_u^{-1}A_s\Phi_\eta(b)|^{1/2} \\
&\quad \times \exp\left\{-\frac{1}{2}\nu_t^j(b)'(\varsigma\Sigma_u + A_s\Phi_\eta(b)\Phi'_\eta(b)A'_s)^{-1}\nu_t^j(b)\right\} \\
&\quad \left(1 - \Phi_N\left(\left(\zeta(s_{t-1}) - \bar{\eta}_{1,t}^j(b)\right)/\sqrt{\bar{\Omega}_{11}(b)}\right)\right).
\end{aligned} \tag{57}$$

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 (42), $\eta_t \sim N(0, I)$, $\eta_{1,t}$ is a scalar, and the measurement equation is given by (43). Draws from the conditionally optimal proposal densities $g_t^*(\tilde{s}_t|s_{t-1}^j, \theta)$, $j = 1, \dots, M$, defined in (54) 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

$$\eta_{1,t}^j \sim N(\bar{\eta}_{1,t}^j(n), \bar{\Omega}_{11}(n))\mathbb{I}\{\eta_{1,t}^j \leq \zeta(s_{t-1}^j)\}, \quad \eta_{2,t}^j|\eta_{1,t}^j \sim N(\bar{\eta}_{2|1}^j(n, \eta_{1,t}^j), \bar{\Omega}_{2|1}(n)) \tag{58}$$

and let

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

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

$$\eta_{1,t}^j \sim N(\bar{\eta}_1^j(b), \bar{\Omega}_{11}(b)) \mathbb{I}\{\eta_{1,t}^j > \zeta(s_{t-1}^j)\}, \quad \eta_{2,t}^j | \eta_{1,t}^j \sim N(\bar{\eta}_{2|1}^j(b, \eta_{1,t}^j), \bar{\Omega}_{2|1}(b)) \quad (59)$$

and

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

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. Suppose that $n_y = n_\eta$ 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 (56) 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):

$$\begin{aligned}
\text{Case 1} & : \bar{\eta}_{1,t}^j(n) < \zeta(s_{t-1}^j), \quad \bar{\eta}_{1,t}^j(b) < \zeta(s_{t-1}^j) \\
\text{Case 2} & : \bar{\eta}_{1,t}^j(n) > \zeta(s_{t-1}^j), \quad \bar{\eta}_{1,t}^j(b) > \zeta(s_{t-1}^j) \\
\text{Case 3} & : \bar{\eta}_{1,t}^j(n) < \zeta(s_{t-1}^j), \quad \bar{\eta}_{1,t}^j(b) > \zeta(s_{t-1}^j) \\
\text{Case 4} & : \bar{\eta}_{1,t}^j(n) > \zeta(s_{t-1}^j), \quad \bar{\eta}_{1,t}^j(b) < \zeta(s_{t-1}^j).
\end{aligned}$$

In Case 1 $D_t^j(b) = 0$ 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 4.¹³

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(s).

Perfectly Observed Regimes. Our ELB 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 ELB 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.

$$\lambda^j = \mathbb{I}\{y_{2,t} > c\}. \quad (60)$$

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 (57) and the moments of the distribution of η_t^j in

¹³The four cases distinguished here are closely connected to the coherency and completeness conditions discussed in Section 3. Previously, we asked whether conditional on a realization of η_t the vector s_t , using the notation of the canonical form (42), is uniquely determined. If the condition is not satisfied, then $p(s_t | s_{t-1}, \theta)$ is not well defined. In addition, for the filtering it matters whether in the absence of measurement errors, there exist one or more η_t 's that can rationalize the data.

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.

7 Numerical Illustrations

We now illustrate the proposed filtering method based on data simulated from the DSGE model of Section 2. First, we compare the distribution of the stochastic likelihood approximation conditional on a particular parameter θ between the proposed COPF and the basic BSPF. It has been shown in the literature, that likelihood approximations of particle filters are unbiased; see Herbst and Schorfheide (2015). In view of the unbiasedness, a low-variance approximation is preferable to a high-variance approximation. Jensen’s inequality implies that \log likelihood approximations are downward biased. The magnitude of the bias is connected to the variability of the likelihood approximation: the larger the variance, the larger also the downward bias.

Second, we embed the particle filter approximation of the likelihood function into a Metropolis-Hastings algorithm. It has also been shown in the literature that if the exact likelihood function is replaced by a noisy but unbiased estimate, MCMC algorithms still converge to the exact posterior distribution. However, there is no free lunch: the larger the variability of the likelihood approximation, the slower the convergence. In the numerical illustrations below, we will compare the speed of convergence as measured by the degree of serial correlation in the parameter draws, across different versions of the particle filter.

Throughout this section, we assume that all variables, including interest rates, are measured with error and that therefore regime (ELB binding versus non-binding) is not perfectly observed. Thus, conditional on the states s_t , y_t has a continuous distribution; see (43) and (44). All computations reported below are executed on a single core of PC with an Intel Xeon CPU E5-2687Wv3 at 3.10 GHz running Windows 10 (64bit) and JuliaPro 1.0.3.1.

7.1 Accuracy of Likelihood Approximation

We simulate a sample of $T = 140$ observations from the DSGE model, loosely parameterized based on the empirical estimates reported in Section 8. The parameter values are summarized in the second column of Table 1. The sample size matches the one used in the empirical

Table 1: DGP and Prior

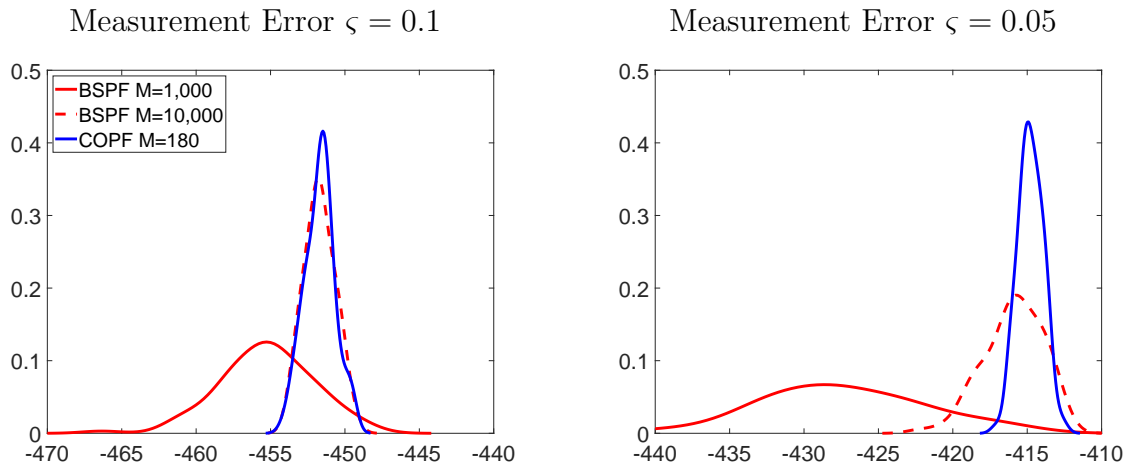
Parameter	DGP	Prior Distribution				
		Density	P(1)	P(2)	HPD Low	HPD High
τ	2.00	\mathcal{G}	2.00	0.20	1.67	2.32
κ	0.13	\mathcal{B}	0.10	0.05	0.02	0.17
ψ_1	2.60			fixed at 2.60		
ψ_2	0.98			fixed at 0.98		
ρ_R	0.80	\mathcal{B}	0.80	0.10	0.65	0.96
ρ_g	0.97	\mathcal{B}	0.80	0.10	0.65	0.96
ρ_d	0.91	\mathcal{B}	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 $\pi_* = \exp\{\text{piAnet}/400\}$.

application in Section 8. In order to increase the likelihood of hitting the ELB in the simulation, we lower the target inflation rate π_* to an annualized rate of 0.5%. In the selected subsample, the ELB binds 22% of the periods, which is roughly consistent with our actual sample.

We begin by comparing the accuracy of the particle-filter-based likelihood approximation using the same parameter values that were used to generate the data. The likelihood evaluation is conditional on a vector of initial states which are the true initial states associated with the simulated observations. We consider two values for the scaling parameter for the measurement errors, ς : 0.1 and 0.05. The smaller ς is, the larger is the penalty for deviations between model-predicted and actual observations. Thus, small measurement errors can be

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 ($M = 180$, blue solid), respectively.

provided as a stress test for the filter in case the model (or its parameterization) is at odds with the data.

The accuracy of the filter depends on the number of particles M . For the COPF, we set $M = 180$, and, for the BSPF, we consider $M = 1,000$ and $M = 10,000$. Starting from the benchmark of the BSPF with $M = 1,000$, the choice of $M = 180$ for the COPF equalizes the run times of the two filters, which is 1.54 seconds for $\zeta = 0.1$.¹⁴ Raising the number of particles from 1,000 to 10,000 increases the run time of the BSPF roughly tenfold, to 16.2 seconds. Moreover, for $\zeta = 0.1$ the BSPF is able to match the accuracy of the COPF with the increased number of particles.

We run the COPF and the BSPF $N_{run} = 100$ times, respectively, and construct kernel estimates of the likelihood from the repeated runs, which are depicted in Figure 1. The more concentrated the densities, the more accurate the likelihood approximation. Holding the run time fixed, the COPF is substantially more accurate than the BSPF. For $\zeta = 0.1$ the number of particles for the BSPF has to be increased from $M = 1,000$ to $M = 10,000$ to achieve similar accuracy to the COPF. This which implies that the BSPF likelihood evaluation takes roughly ten times as long as the COPF likelihood evaluation.

¹⁴The absolute run times depend on the programming language – Julia in our case – and the efficiency of the code. We are mainly interested in the relative run times. The code for the COPF and BSPF is almost identical, except that the COPF requires a few additional steps to evaluate the expressions in Proposition 1, which implies that for the same M , the COPF is slower than the BSPF.

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 (and by virtue of the concave transformation also the bias) 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-ELB periods, because the number of observables equals the number of shocks.

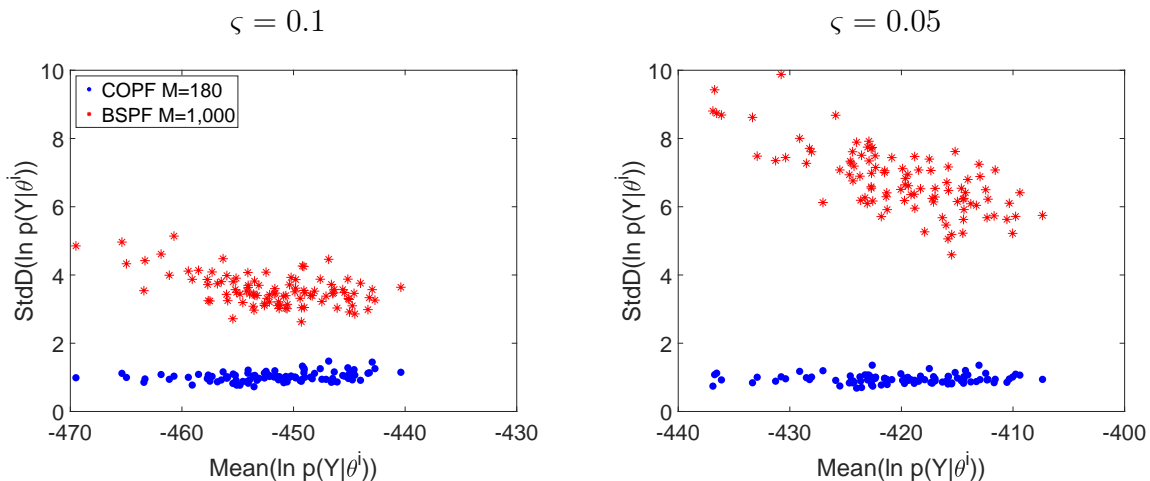
Figure 2 depicts standard deviations of log-likelihood approximations as a function of the mean log-likelihood value across $N_{run} = 100$ runs of the filter for $\varsigma = 0.1$ and $\varsigma = 0.05$. Each dot (or asterisk) in the two scatter plots corresponds to a different parameter value θ^i .¹⁵ 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.

7.2 Particle MCMC

We now embed the particle filter likelihood approximations in a standard single-block RWMH algorithm. The particle RWMH algorithm operates on an enlarged probability space that includes all the random variables that are generated when running the particle filter; see Herbst and Schorfheide (2015) for a textbook exposition and Andrieu et al. (2010) for a formal analysis. The use of an enlarged probability space leads to an increase in the persistence of the Markov chain generated by the posterior sampler. The noisier the likelihood approximation, the larger the persistence of the resulting Markov chain. A high degree of serial correlation is undesirable, because it leads to very noisy Monte Carlo approximations of posterior moments. We will show that the use of an accurate particle filter such as the COPF can alleviate this problem.

¹⁵The parameter values are obtained from the output of the particle Markov chain Monte Carlo (PMCMC) algorithm in Section 7.2.

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 $M = 180$ particles for the COPF (blue dots).

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 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 ELB constraint and sets the measurement error variance to zero ($\zeta = 0$). In this case the likelihood function can be evaluated with the Kalman filter (KF). We conduct two preliminary MCMC runs using the linearized model. The first run is based on a diagonal covariance matrix with scaled prior variances on the diagonal and the second run is based on the posterior covariance matrix of the first run. Finally, we compute the posterior covariance matrix from the second run and denote it by \hat{V}_θ . The particle MCMC runs using the nonlinear DSGE model are based on $c\hat{V}_\theta$ where $c = 0.1$.

For the estimation of the nonlinear model we set the scale of the measurement error variance to $\zeta = 0.1$ and keep the number of particles at $M = 180$ for the COPF and $M = 1,000$ for the BSPF. 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 BSPF ranges from 3 to 5. We generate 55,000 draws from the posterior distribution and discard the first 5,000 draws. The overall run time using the two filters is approximately the same: 24 hours and 37 minutes for the COPF-RWMH algorithm and 25 hours and 55 minutes for the BSPF-RWMH algorithm. Note that the codes for the COPF-RWMH and the BSPF-RWMH algorithms are identical, except for the more complicated way of generating draws from $g_t(\tilde{s}_t | s_{t-1}^j, \theta)$ and evaluating $\tilde{\omega}_t^j$ in the COPF likelihood approximation.

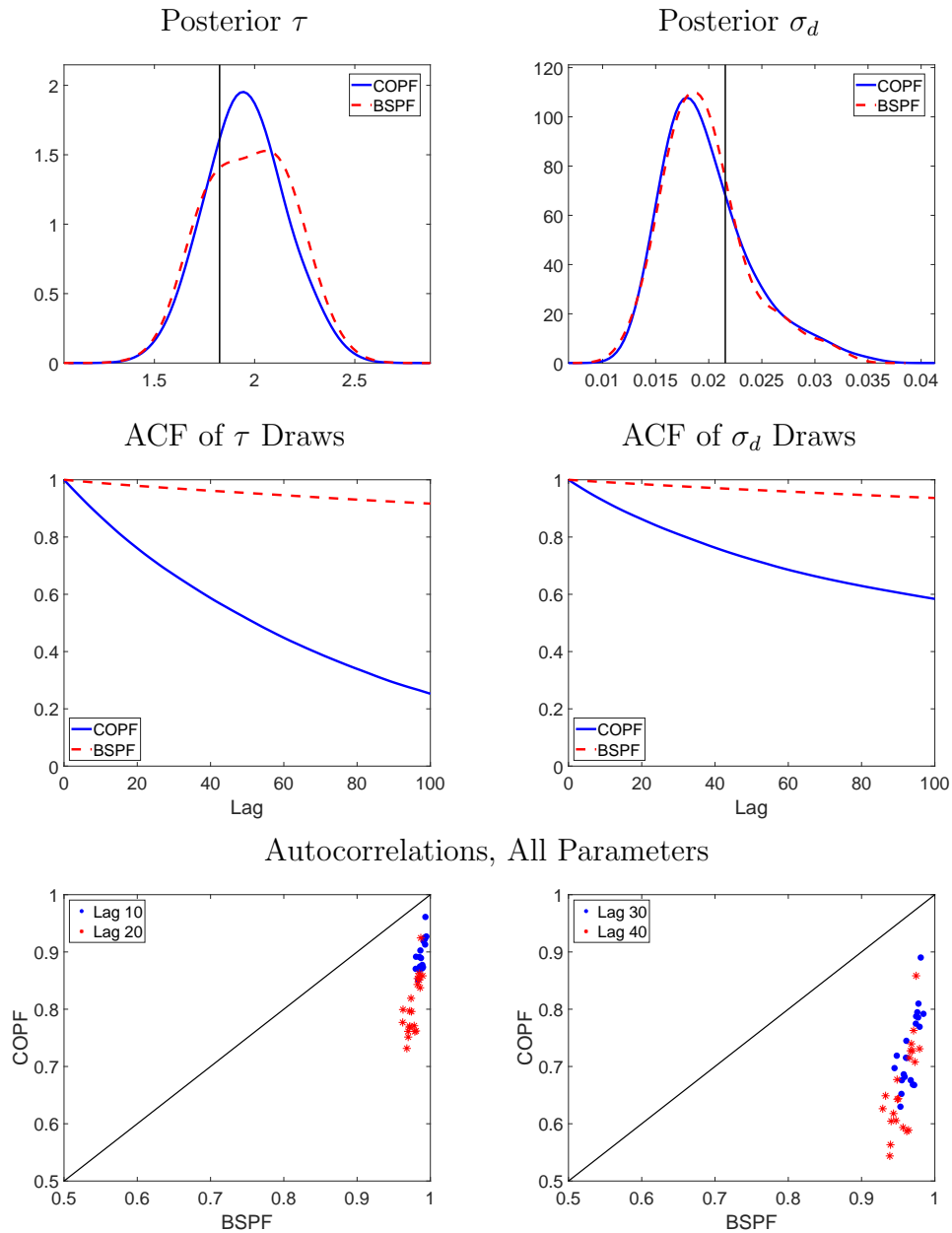
Using a scaling of $c = 0.1$, the acceptance rate for the proposed draws is 28% for the COPF-RWMH algorithm. Herbst and Schorfheide (2015) documented that for the estimation of small-scale linearized DSGE models, an acceptance rate between 15% and 30% is associated with the most accurate Monte Carlo approximations of posterior means. Replacing the COPF by the BSPF while keeping the scale factor $c = 0.1$, the acceptance rate drops to 2.6%. The lower acceptance rate is caused by the noisier likelihood approximation.¹⁶

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 for σ_d look very similar and both peak near the “true” parameter values depicted by the solid vertical line. For the parameter τ , the posterior density obtained from the COPF is less concentrated at the peak than the BSPF density. The reason is that BSPF chain moves very slowly – recall the 2.6% acceptance rate – and has not fully explored the high-density area of the posterior.

The second row of Figure 3 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 0.92 for the τ^i sequence, it is only 0.25 for draws from the COPF-RWMH. Suppose the autocorrelation function of the draws is given by ρ^j , where j is the temporal displacement among the draws. Then the variance of the mean of the draws is given by $\mathbb{V}_\pi[h(\theta)](1 + \rho)/(1 - \rho)$, where $\mathbb{V}_\pi[h(\theta)]$ is the posterior variance of $h(\theta)$ and $(1 + \rho)/(1 - \rho)$ can be viewed as inefficiency factor that arises due to the serial correlation of the draws. A reduction from $\rho = 0.92^{1/100} = 0.999$ to $\rho = 0.25^{1/100} = 0.986$ lowers the inefficiency factor from approximately 2,000 to 142. This implies only 7% of the draws are required to achieve the same accuracy of posterior mean approximations with the COPF-based sampler as with the BSPF-based sampler, or, holding the numbers of draws

¹⁶The acceptance rate could be increased by decreasing c , but it does not cure the persistence problem because the chain moves very slowly due to the smaller size of the accepted steps. In our case, setting $c = 0.005$ delivers an acceptance rate of 3.8% and an even more persistent sequence of parameter draws.

Figure 3: Posterior Draws: Density and Autocorrelation



Notes: ME scale $\zeta = 0.1$, proposal covariance scale $c = 0.1$, $N = 55,000$ draws (drop first 10%). COPF: number of particles $M = 180$, run time is 24:37:33 (hh:mm:ss). BSPF: number of particles $M = 1,000$, run time is 25:55:14 (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.

fixed across samplers, the COPF-based sampler delivers Monte Carlo approximations that are six times as accurate (in terms of sampling variance).

The last row of Figure 3 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.

To summarize, when we calibrated the COPF and the BSPF to run at the same time, the former delivers a likelihood approximation that is as much as five times more accurate, and when used inside a RWMH algorithm we showed that the COPF yields a posterior mean that is substantially more accurate. Because the COPF needs a relatively small number of particles, $M = 180$ in our numerical illustration, it is possible to accurately estimate a DSGE model with an occasionally binding constraint without supercomputing capabilities in a relatively short amount of time.

8 Empirical Application

We now estimate the small-scale New Keynesian DSGE model based on quarterly U.S. data using the previously developed model solution and filtering techniques and conduct a fiscal policy experiment. The estimation results are summarized in Section 8.1, and the fiscal policy analysis appears in Section 8.2.

8.1 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 ELB 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.¹⁷ 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 et al. (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 ELB constraint, setting the measurement error variances of the state-space model to zero. 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.

For the estimation of the nonlinear version of the model we assume that interest rates are observed without error. Thus, conditional on R_t^o , it is known whether the ELB is binding or not; see (60). We maintain the assumption that the remaining variables are observed subject to a measurement error and set the scale factor for their measurement error variances to $\varsigma = 0.001$. 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

¹⁷This 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	2.05	1.75	2.40
κ	0.12	0.10	0.08	0.16
ρ_R	0.82	0.83	0.79	0.85
ρ_g	0.981	0.986	0.967	0.996
ρ_d	0.95	0.96	0.93	0.97
ρ_z	0.32	0.29	0.17	0.47
σ_R	0.0018	0.0018	0.0016	0.0020
σ_g	0.0026	0.0029	0.0023	0.0028
σ_d	0.0282	0.0290	0.0178	0.0399
σ_z	0.0060	0.0069	0.0053	0.0068
g^*	1.27	1.28	1.25	1.29
rAnet	0.67	0.23	0.27	1.08
gamQnet	0.36	0.47	0.24	0.48
piAnet	2.98	2.69	2.54	3.42

Notes: The estimation period is 1984:Q1 to 2018:Q4. The following parameters 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).

the covariance matrix of the proposal distribution. The likelihood function is approximated using the COPF with $M = 180$ particles. The resulting acceptance rate of the particle RWMH algorithm is 25%, and the run time is 16 hours and 14 minutes on a single core, which comes to about 0.9 seconds per draw. The assumption of a perfectly observed regime simplifies the COPF calculations and reduces the run time by 34%, compared to the results reported in Section 7.2.

The parameter estimates are summarized in Table 2. The table reports posterior means, the maximum posterior probability (MAP) estimates, and lower and upper endpoints of highest posterior density (HPD).¹⁸ 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 curve is $\hat{\kappa} = 0.12$. The government spending shock is close to

¹⁸We compared these estimates with those obtained from a linearized model using the Kalman Filter and data that exclude the ELB episode. The most noteworthy differences are in ρ_d and σ_d , both of which need to be larger when the ELB episode is used in order to deliver large (negative) and persistent shocks that take the economy to the ELB. 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.

a unit-root process ($\hat{\rho}_g = 0.981$), 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 discuss them in the next section.

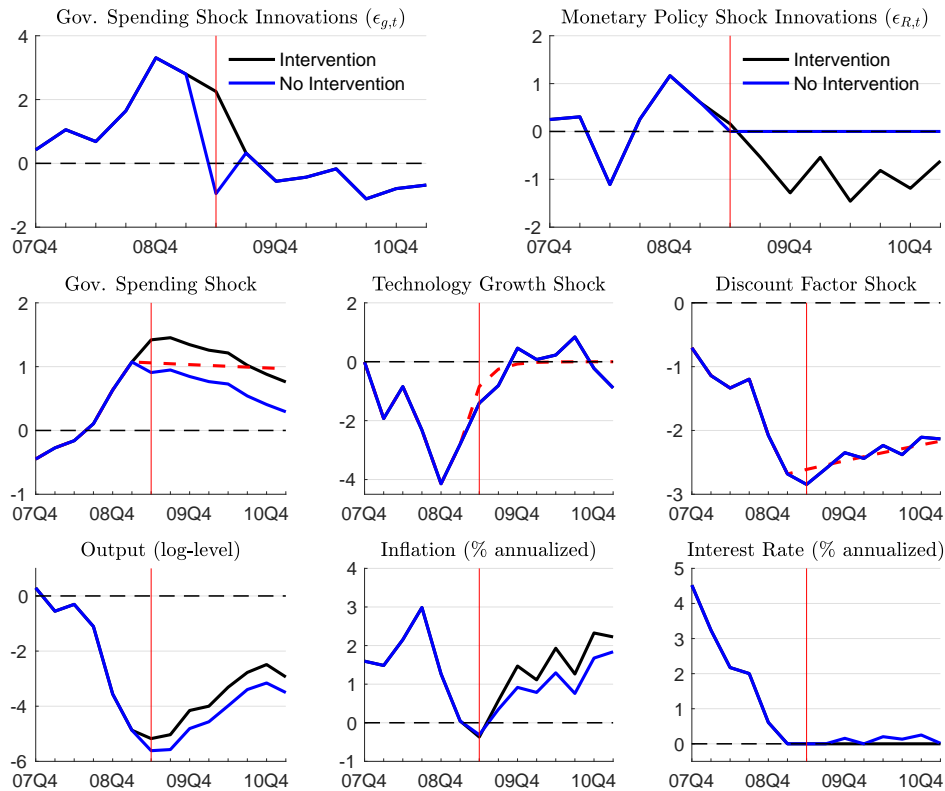
8.2 Fiscal Policy Analysis

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 ELB; see, for instance, Eggertsson (2011) and Christiano and Eichenbaum (2012). In the absence of the ELB, 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 ELB 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.

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.

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 $2.7\sigma_g$, 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.

Figure 4: Ex Post Policy Analysis



Notes: The 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.

We use the COPF to obtain estimates of the exogenous shock processes for the period 2009:Q2 through 2011:Q1. The subsequent results are based on the 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 ($\hat{\rho}_z = 0.29$) and reverts back to zero by the end of 2009, the mean reversion of the discount factor shock is very slow, and it remained 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

$$\hat{g}_{t|t}^C = \hat{g}_{t|t} - \rho_g^{t-T_*} \delta^{ARRA} \text{ for } t = T_*, T_* + 1, \dots, T_* + 7, \quad (61)$$

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 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.¹⁹ 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 41 basis points lower, and annualized output growth would have been 32 basis points lower, on average, over these eight quarters.

¹⁹Details on the algorithm to compute the effects of the policy interventions are reported in the Online Appendix.

Table 3: Ex Post Multipliers

Intervention	1Q	4Q	8Q
Fiscal Only	0.65	0.69	0.69
Fiscal + Monetary	0.63	0.77	0.82

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.

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 0.7 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 ELB. 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 (for example starting the exercise in 2007:Q1), then the fiscal multipliers would be around 0.6, which is indeed lower than our ex post multiplier. However, the presence of the ELB only generates a modest increase in the multiplier – on impact at the ELB the multiplier is 0.65 while away from the ELB the multiplier is 0.55.

Focusing on the two different policies at the ELB, 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 ELB, leaving very little room for conventional monetary 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.

9 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 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 ELB constraint in a relatively short amount of time on a single core processor.

References

- ADAM, K. AND R. BILLI (2007): “Discretionary Monetary Policy and 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, 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, 269–342.
- ARULAMPALAM, M. 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, 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, 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.
- ARUOBA, S. B., F. SCHORFHEIDE, AND S. VILLALVAZO (2020): “SVARs with Occasionally-Binding Constraints,” *Manuscript, University of Maryland and University of Pennsylvania*.
- ASCARI, G. AND S. MAVROEIDIS (2020): “The Unbearable Lightness of Equilibria in a Low Interest Rate Environment,” *Manuscript, Oxford University*.
- ATKINSON, T., A. RICHTER, AND N. A. THROCKMORTON (2020): “The Zero Lower Bound and Estimation Accuracy,” *Journal of Monetary Economics*, forthcoming.
- BENHABIB, J., S. SCHMITT-GROHÉ, AND M. URIBE (2001): “The Perils of Taylor Rules,” *Journal of Economic Theory*, 96, 40–69.
- 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, 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, 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, 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, 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, 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*, 1–13.
- 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, University of Chicago Press, vol. 25, 59–112.
- EGGERTSSON, G. B. AND M. WOODFORD (2003): “The Zero Bound on Interest Rates and Optimal Monetary Policy,” *Brookings Papers on Economic Activity*, 34, 139–235.
- FAIR, R. C. AND J. B. TAYLOR (1983): “Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models,” *Econometrica*, 51, 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, 2150 – 2166.

- FERNÁNDEZ-VILLAVERDE, J., G. GORDON, P. GUERRÓN-QUINTANA, AND J. F. RUBIO-RAMÍREZ (2015): “Nonlinear Adventures at the Zero Lower Bound,” *Journal of Economic Dynamics and Control*, 57, 182 – 204.
- FERNÁNDEZ-VILLAVERDE, J. AND J. F. RUBIO-RAMÍREZ (2007): “Estimating Macroeconomic Models: A Likelihood Approach,” *Review of Economic Studies*, 74, 1059–1087.
- GORDON, N. AND D. SALMOND (1993): “A Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation,” *IEEE Proceedings-F*, 140, 107–113.
- GOURIEROUX, C., J. J. LAFFONT, AND A. MONFORT (1980): “Coherency Conditions in Simultaneous Linear Equation Models with Endogeneous Switching Regimes,” *Econometrica*, 48, 675–695.
- 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, 1971–2006.
- HEISS, F. AND V. WINSCHERL (2008): “Likelihood approximation by numerical integration on sparse grids,” *Journal of Econometrics*, 144, 62 – 80.
- HERBST, E. AND F. SCHORFHEIDE (2015): *Bayesian Estimation of DSGE Models*, Princeton University Press.
- 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, 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.
- KOPECKY, K. A. AND R. M. SUEN (2010): “Finite state Markov-chain approximations to highly persistent processes,” *Review of Economic Dynamics*, 13, 701 – 714.
- KRUEGER, D. AND F. KUBLER (2004): “Computing Equilibrium in OLG Models with Stochastic Production,” *Journal of Economic Dynamics and Control*, 28, 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, Elsevier, vol. 3, 325–477.
- (2015): “Merging Simulation and Projection Approaches to Solve High-Dimensional Problems with an Application to a New Keynesian Model,” *Quantitative Economics*, 6, 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, 229–239.
- MAVROEIDIS, S. (2020): “Identification at the Zero Lower Bound,” *Manuscript, Oxford University*.
- MENDES, R. (2011): “Uncertainty at the Zero Lower Bound: A Theoretical Analysis,” *MRPA Working Paper No. 60103*.
- MENDOZA, E. G. AND S. VILLALVAZO (2020): “FiPIt: A simple, fast global method for solving models with two endogenous states & occasionally binding constraints,” *Review of Economic Dynamics*, 37, 81–102.
- 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, 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. Borağan 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. Derivations for Section 3
- C. Canonical Form for the Model of Section 2
- D. Proofs and Derivations for Section 6
- E. Additional Details for the Empirical Application
- F. Consumption-Savings Model with Borrowing Constraint

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 equations 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 goods 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 3

The log-linearized system (omitting hats) that characterizes the simplified model takes the form

$$\begin{aligned} R_t &= \max \left\{ \psi\pi_t + \sigma_R\epsilon_{R,t}, -\ln(r\pi_*) \right\} \\ c_t &= d_t + \mathbb{E}_t[c_{t+1} - d_{t+1}] - (R_t - \mathbb{E}_t[\pi_{t+1}]) \\ \pi_t &= \beta\mathbb{E}_t[\pi_{t+1}] + \kappa c_t. \end{aligned} \tag{A.1}$$

The assumption that shocks are *iid* implies that $\mathbb{E}_t[d_{t+1}] = 0$. Because we are focusing on solutions under which the endogenous variables are *iid*, we can set $\mathbb{E}_t[c_{t+1}] = \mu_c$ and $\mathbb{E}_t[\pi_{t+1}] = \mu_\pi$. This leads to

$$\begin{aligned} R_t &= \max \left\{ \psi\pi_t + \sigma_R\epsilon_{R,t}, -\ln(r\pi_*) \right\} \\ c_t &= -R_t + \mu_c + \mu_\pi + \sigma_d\epsilon_{d,t} \\ \pi_t &= \beta\mu_\pi + \kappa c_t. \end{aligned} \tag{A.2}$$

The consumption equation corresponds to the second equation in (19) in the main text. Combining the Euler equation and the Phillips curve leads to the following expression for inflation:

$$\pi_t = -\kappa R_t + \kappa\mu_c + (\kappa + \beta)\mu_\pi + \kappa\sigma_d\epsilon_{d,t}, \tag{A.3}$$

which is the third equation in (19) in the main text. We now can use (A.3) to eliminate inflation from the monetary policy rule:

$$R_t = \max \left\{ -\psi\kappa R_t + \psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi + \psi\kappa\sigma_d\epsilon_{d,t} + \sigma_R\epsilon_{R,t}, -\ln(r\pi_*) \right\} \quad (\text{A.4})$$

Note that if the ELB is non-binding the interest rate is given by

$$R_t = \frac{1}{1 + \psi\kappa} \left[\psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi + \psi\kappa\sigma_d\epsilon_{d,t} + \sigma_R\epsilon_{R,t} \right].$$

Thus, we can also write

$$R_t = \max \left\{ \frac{1}{1 + \psi\kappa} \left[\psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi + \psi\kappa\sigma_d\epsilon_{d,t} + \sigma_R\epsilon_{R,t} \right], -\ln(r\pi_*) \right\}, \quad (\text{A.5})$$

which is the first equation in (19).

The solution of the model requires the calculation of the mean of a truncated random variable. If $X \sim N(\mu, \sigma^2)$ and C is a truncation constant, then

$$\mathbb{E}[X|X \geq C] = \mu + \frac{\sigma\phi_N(\alpha)}{1 - \Phi_N(\alpha)},$$

where $\alpha = (C - \mu)/\sigma$, $\phi_N(x)$ and $\Phi_N(\alpha)$ are the probability density function (pdf) and the cumulative density function (cdf) of a $N(0, 1)$. To obtain $\mathbb{E}[R(\eta_{1,t})]$, we need to compute $\mathbb{E}[\eta_{1,t}|\eta_{1,t} \geq \bar{\eta}_1]$. Because $\eta_{1,t} \sim N(0, 1)$ we obtain

$$\mathbb{E}[\eta_{1,t}|\eta_{1,t} \geq \bar{\eta}_1] = \frac{\phi_N(\bar{\eta}_1)}{1 - \Phi_N(\alpha)},$$

which is used in (23) in the main text.

To show continuity of the consumption decision rule (see (25) in the main text), consider

the following limit from above:

$$\begin{aligned}
& \lim_{\eta_{1,t} \downarrow \bar{\eta}_1} c(\eta_{1,t}, \eta_{2,t}) \\
&= \frac{1}{1 + \psi\kappa} \left[\mu_c + (1 - \psi\beta)\mu_\pi \right] + \left(\gamma_{d,1} - \frac{1}{1 + \psi\kappa} \right) \sigma_\eta \bar{\eta}_1 + \gamma_{d,2} \sigma_\eta \eta_{2,t} \\
&= \frac{1}{1 + \psi\kappa} \left[\mu_c + (1 - \psi\beta)\mu_\pi + (1 + \psi\kappa) \ln(r\pi_*) + \psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi \right] \\
&\quad + \gamma_{d,1} \sigma_\eta \bar{\eta}_1 + \gamma_{d,2} \sigma_\eta \eta_{2,t} \\
&= \ln(r\pi_*) + \mu_c + \mu_\pi + \gamma_{d,1} \sigma_\eta \bar{\eta}_1 + \gamma_{d,2} \sigma_\eta \eta_{2,t} \\
&= c(\bar{\eta}_1, \eta_{2,t}).
\end{aligned}$$

To obtain the second equality, we use the formula for $\bar{\eta}_1$ from (22) in the main text:

$$\bar{\eta}_1 = -\frac{1}{\sigma_\eta} \left[(1 + \psi\kappa) \ln(r\pi_*) + \psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi \right].$$

The last two equalities establish the continuity. A similar calculation for the inflation decision rule yields

$$\begin{aligned}
& \lim_{\eta_{1,t} \downarrow \bar{\eta}_1} \pi(\eta_{1,t}, \eta_{2,t}) \\
&= \frac{1}{1 + \psi\kappa} \left[\kappa\mu_c + (\kappa + \beta)\mu_\pi + \kappa \left((1 + \psi\kappa) \ln(r\pi_*) + \psi\kappa\mu_c + \psi(\kappa + \beta)\mu_\pi \right) \right] \\
&\quad + \kappa\gamma_{d,1} \sigma_\eta \bar{\eta}_1 + \kappa\gamma_{d,2} \sigma_\eta \eta_{2,t} \\
&= \kappa \ln(r\pi_*) + \kappa\mu_c + (\kappa + \beta)\mu_\pi + \kappa\gamma_{d,1} \sigma_\eta \bar{\eta}_1 + \kappa\gamma_{d,2} \sigma_\eta \eta_{2,t} \\
&= \pi(\bar{\eta}_1, \eta_{2,t}).
\end{aligned}$$

Replacing the $\eta_{i,t}$'s by the $\epsilon_{i,t}$'s does not affect continuity because the transformation is linear.

C Canonical Form for the New Keynesian 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 ELB constraint.

Define the vectors s_t

$$s_t = [\hat{y}_t, \hat{\pi}_t, \hat{R}_t, \hat{z}_t, \hat{d}_t, \hat{g}_t, e_{R,t}]' \quad (\text{A.6})$$

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 (33) 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 (A.6). 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 (17).

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

$$\zeta(s_{t-1}) = \ln(r_*\pi_*) + \phi_0(n) + \phi_1'(n)s_{t-1}, \quad \eta_{1,t} = -\frac{1}{\|\phi_\epsilon(n)\|}\phi_\epsilon'(n)\epsilon_t$$

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

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

as in (42). 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 a 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.

D Derivations for Section 6

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

$$\begin{aligned} & p(y_t|\eta_t, s_{t-1})p(\eta_t) && \text{(A.7)} \\ &= p_N(y_t; \hat{y}_{t|t-1}(n), \varsigma\Sigma_u)p_N(\eta_{1,t}; 0, 1)p_N(\eta_{2,t}; 0, I)\mathbb{I}\{\eta_{1,t} \leq \zeta(s_{t-1})\} \\ &\quad + p_N(y_t; \hat{y}_{t|t-1}(b), \varsigma\Sigma_u)p_N(\eta_{1,t}; 0, 1)p_N(\eta_{2,t}; 0, I)\mathbb{I}\{\eta_{1,t} > \zeta(s_{t-1})\} \\ &= I + II. \end{aligned}$$

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

$$\hat{\Omega}^{-1} = \Phi_\eta' A' (\varsigma \Sigma_u)^{-1} A \Phi_\eta, \quad \hat{\eta}_t = \hat{\Omega} \Phi_\eta' A' (\varsigma \Sigma_u)^{-1} \nu_t. \quad (\text{A.8})$$

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 $\eta_t | (y_t, s_{t-1})$

$$\bar{\Omega} = (\underline{\Omega}^{-1} + \hat{\Omega}^{-1})^{-1}, \quad \bar{\eta}_t = \bar{\Omega} \hat{\Omega}^{-1} \hat{\eta}_t. \quad (\text{A.9})$$

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 $\eta_{1,t}$:

$$\bar{\eta}_{2|1,t}(\eta_{1,t}) = \bar{\eta}_{2,t} + \bar{\Omega}_{21}\bar{\Omega}_{11}^{-1}(\eta_{1,t} - \bar{\eta}_{1,t}), \quad \bar{\Omega}_{2|1,t} = \bar{\Omega}_{22} - \bar{\Omega}_{21}\bar{\Omega}_{11}^{-1}\bar{\Omega}_{12}. \quad (\text{A.10})$$

Thus,

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

This is the final form for term I in (A.7).

Integrating I in (A.11) with respect to $(\eta_{1,t}, \eta_{2,t})$ and re-introducing the (n) arguments yields

$$\begin{aligned} D(n) &= \int \int I(\eta_{1,t}, \eta_{2,t}) d\eta_{2,t} d\eta_{1,t} \\ &= (2\pi)^{-n_y/2} |\zeta \Sigma_u|^{-1/2} |\underline{\Omega}|^{-1/2} |\bar{\Omega}(n)|^{1/2} \Phi_N \left((\zeta(s_{t-1}) - \bar{\eta}_{1,t}(n)) / \sqrt{\bar{\Omega}_{11}(n)} \right) \\ &\times \exp \left\{ -\frac{1}{2} \nu_t(n)' \Sigma_u^{-1} \nu_t(n) \right\} \exp \left\{ \frac{1}{2} \bar{\eta}_t'(n) \bar{\Omega}^{-1}(n) \bar{\eta}_t(n) \right\}. \end{aligned} \quad (\text{A.12})$$

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

$$\begin{aligned} D(b) &= (2\pi)^{-n_y/2} |\zeta \Sigma_u|^{-1/2} |\underline{\Omega}|^{-1/2} |\bar{\Omega}(b)|^{1/2} \left(1 - \Phi_N \left((\zeta(s_{t-1}) - \bar{\eta}_{1,t}(b)) / \sqrt{\bar{\Omega}_{11}(b)} \right) \right) \\ &\times \exp \left\{ -\frac{1}{2} \nu_t(n)' \Sigma_u^{-1} \nu_t(n) \right\} \exp \left\{ \frac{1}{2} \bar{\eta}_t'(b) \bar{\Omega}^{-1}(b) \bar{\eta}_t(b) \right\}. \end{aligned} \quad (\text{A.13})$$

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

$$p(\eta_t | y_t, s_{t-1}) = \frac{p(y_t | \eta_t, s_{t-1}) p(s_{t-1})}{\int p(y_t | \eta_t, s_{t-1}) p(s_{t-1}) d\eta_t} = \frac{I(n) + II(b)}{D(n) + D(b)}. \quad (\text{A.14})$$

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

$$\lambda = \frac{D(n)}{D(n) + D(b)}. \quad (\text{A.15})$$

Then, with probability λ

$$\eta_{1,t} \sim N(\bar{\eta}_{1,t}(n), \bar{\Omega}_{11}(n)) \mathbb{I}\{\eta_{1,t} \leq \zeta(s_{t-1})\}, \quad \eta_{2,t} | \eta_{1,t} \sim N(\bar{\eta}_{2|1,t}(n, \eta_{1,t}), \bar{\Omega}_{2|1}(n)) \quad (\text{A.16})$$

and with probability $1 - \lambda$

$$\eta_{1,t} \sim N(\bar{\eta}_{1,t}(b), \bar{\Omega}_{11}(b)) \mathbb{I}\{\eta_{1,t} > \zeta(s_{t-1})\}, \quad \eta_{2,t} | \eta_{1,t} \sim N(\bar{\eta}_{2|1,t}(b, \eta_{1,t}), \bar{\Omega}_{2|1}(b)). \quad (\text{A.17})$$

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

$$\tilde{\omega}_t^j = p(y_t | s_{t-1}^j) = D(n) + D(b). \quad (\text{A.18})$$

This completes the proof of the proposition. ■

In the remainder of this section we consider two special cases: (i) y_t identifies the regime without error. This is the case, for instance, for a DSGE model with ELB constraint if the interest rate is observed without error, at least when it hits the ELB. (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.7) as

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

The formula for $D(n)$ in (A.12) remains unchanged. The formula for $D(b)$ in (A.13) 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.16) and (A.17), 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.13) 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.12) 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.12). 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.16) and (A.17) is replaced by setting $\eta_{1,t}$ and $\eta_{2,t}$ equal to their means.

E 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

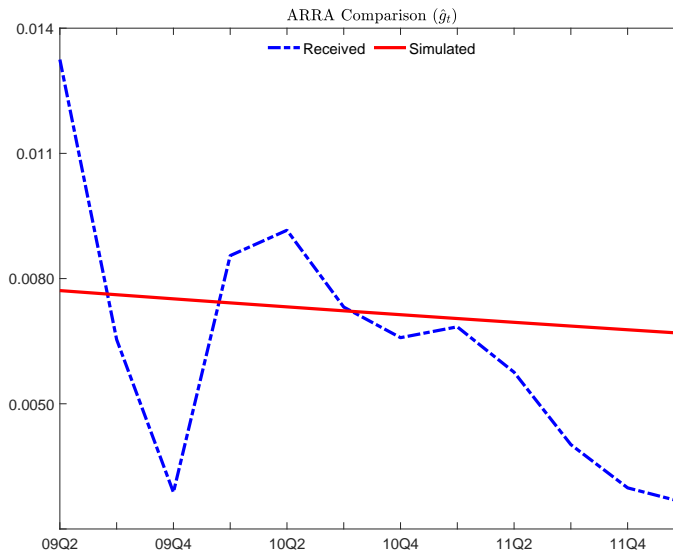
$$\hat{\zeta}_t^{ARRA} = \frac{1}{g_* - 1} \hat{g}_t^{ARRA},$$

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

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

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).²⁰

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.

²⁰Recall that $\sigma_g = 0.0029$, hence the ARRA impulse in our experiment is equal to $2.7 \times \sigma_g$.

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

$$IRF(x_t | \epsilon_t^I, \epsilon_t) = (\ln x_t^I - \ln x_t). \quad (\text{A.19})$$

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.

F Consumption-Savings Model with Borrowing Constraint

In this section we consider a simple consumption-savings model with occasionally-binding borrowing constraint. We compare the proposed PLC solution algorithm against OccBin and a time-iteration algorithm. The time-iteration algorithm delivers an exact solution, whereas the other two algorithms generate approximate solutions.

F.1 Model Specification

Consider the following model in which a representative agent chooses an path for consumption, C_t and borrowing B_{t+1} to maximizing the following expected utility function:

$$\max_{\{C_t, B_{t+1}\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\gamma} - 1}{1-\gamma},$$

where γ is the coefficient of relative risk aversion. The maximization problem is subject to the budget constraint and to an exogenous borrowing limit that states that the amount of debt chosen in the current period, B_{t+1} , cannot exceed a fraction m of the current period income Y_t :

$$C_t + RB_t = Y_t + B_{t+1} \tag{A.20}$$

$$B_{t+1} \leq mY_t. \tag{A.21}$$

In this economy, borrowing takes the form of one-period non-state contingent bonds that pay a fixed interest rate R . The borrowing constraint can become occasionally binding depending on the level of debt at the beginning of the period, B_t , and the realization of the exogenous stream of income that follows a stochastic process: $\ln Y_t = \rho \ln Y_{t-1} + \sigma \epsilon_t$, with $0 \leq \rho < 1$, $\sigma > 0$, and $\epsilon_t \sim N(0, 1)$. Denoting by λ_t the Lagrange multiplier associated to the borrowing constraint, we can define a competitive equilibrium allocation.

Definition 1 *A competitive equilibrium in this economy is a sequence of consumption, borrowing decisions, and Lagrange multipliers $\{C_t, B_{t+1}, \lambda_t\}_{t=0}^{\infty}$, that given initial conditions B_0*

and the exogenous sequence of income $\{Y_t\}_{t=0}^{\infty}$, satisfy the following equations:

$$(C_t)^{-\gamma} = \beta RE_t (C_{t+1})^{-\gamma} + \lambda_t \quad (\text{A.22})$$

$$C_t + RB_t = Y_t + B_{t+1} \quad (\text{A.23})$$

$$\lambda_t (mY_t - B_{t+1}) = 0. \quad (\text{A.24})$$

F.2 Solution Algorithms

F.2.1 Time Iteration

The time iteration procedure delivers a global approximation to the recursive equilibrium associated with the optimality conditions of dynamic optimization problem. We follow the implementation in Mendoza and Villalvazo (2020) and implement the Fixed-Point Iteration algorithm (FiPIt).

The time iteration solves for recursive policy functions of the form $B_{t+1} = b'(B, Y)$ and $C_t = c(B, Y)$ that define the optimal amount of borrowing and consumption given the current level of debt B and income Y . Denote a candidate policy function as $\hat{b}'_j(B, Y)$, where j denotes the sub-index that helps track the proposed policy function on the j th step of the algorithm outlined below. The associated consumption decision rule $c_j(B, Y) = Y + \hat{b}'_j(B, Y) - RB$. Note that $c_j(\hat{b}'_j(B, Y), Y') = Y' - R\hat{b}'_j(B, Y) + \hat{b}'_j(\hat{b}'_j(B, Y), Y')$. Using the Euler equation (A.22), we can write the recursion under the assumption that the constraint is slack in period $t + 1$:

$$c_{j+1}(B, Y) = \left\{ \beta R \mathbb{E}_{Y'|Y} \left[Y' - R\hat{b}'_j(B, Y) + \hat{b}'_j(\hat{b}'_j(B, Y), Y') \right]^{-\gamma} \right\}^{-\frac{1}{\gamma}}. \quad (\text{A.25})$$

The updated decision rule for bond holdings can be recovered directly from the budget constraint combined with borrowing constraint: $\hat{b}'_{j+1}(B, Y) = \max \{c_{j+1}(B, Y) - Y + RB, mY\}$.

Algorithm 3 (FiPIt Algorithm)

1. Define a grid of values for the state variables $\mathcal{S} \subset [B_{min}, B_{max}] \times [Y_{min}, Y_{max}]$. For current debt levels we define an equally-spaced grid $\mathbf{B} = \{b_1 < b_2 < \dots < b_{N_b}\}$ with a total of $N_b = 200$ grid points. For the income grid, we discretize the AR(1) process for $Z_t = \ln(Y_t)$ using the method of Kopecky and Suen (2010) and obtain a grid $\mathbf{Y} =$

$\{y_1 < y_2 < \dots < y_{N_z}\}$ with a total of $N_z = 11$ grid points. This method also generates a transition matrix, $P_{N_z \times N_z}$ for the Markov-chain approximation for the evolution of Z_t that we can use to compute expectations. Finally we construct the solution grid using a Cartesian product $\mathbf{B} \otimes \mathbf{Y}$.

2. Start with guess $\hat{b}'_j(B, Y)$ for every grid points. $(B, Y) \in \mathcal{S}$. To evaluate $\hat{b}'_j(\hat{b}'_j(B, Y), Y')$ off the grid points in \mathcal{S} , we use one-dimensional linear interpolation.
3. Compute $c_{j+1}(B, Y)$ using equation (A.25).
4. Obtain the implied borrowing decision rule: $\hat{b}'_{j+1}(B, Y) = \max \{c_{j+1}(B, Y) - Y + RB, mY\}$.
5. Compute the distance $\vartheta = \|\hat{b}'_{j+1}(B, Y) - \hat{b}'_j(B, Y)\|$. If $\vartheta \leq 10^{-7}$ stop, otherwise set $j = j + 1$ and go back to step 2.
6. Once the recursion converges, for every (B, Y) , we can obtain $c(B, Y)$ using (A.23) and the $\hat{b}'(B, Y)$ that we obtain from the recursion.

F.2.2 PLC Approximation

Let $\mathbb{X} = \{x_1, X_2'\}$ be the $n \times 1$ vector of state variables. Specifically: $x_1 = \{B_t\}$ and $X_2 = \{1, Y_t\}$. We proceed approximating the decision rule for the level of debt that will be carried into the next period $B_{t+1} = b'(x_1, X_2)$ and in the standard generic notation we have $Y = \{B'\}$. The PLC decision rule takes the following form:

$$b'(x_1, X_2) = \begin{cases} \alpha_{1,1}x_1 + \alpha'_{1,2}X_2, & \text{if } x_1 \leq \delta X_2 \\ \alpha_{2,1}x_1 + \alpha'_{2,2}X_2 & \text{if } x_1 > \delta X_2 \end{cases}, \quad (\text{A.26})$$

where the linear function $x_1 = \delta X_2$ characterizes the locus of points in the state space that satisfy borrowing constraint just exactly.

The borrowing constraint in (A.21) is linear and we can write it as:

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

Given the general form of the kink function $h(x_1, X_2, y) \equiv \gamma_1 x_1 + \gamma_2' X_2 + \gamma_Y' y$, we deduce that the constants entering the PLC constraint are: $\gamma_1 = 0$, $\gamma_2' = [0 \quad m]$ and $\gamma_Y' = -1$.

Plugging the decision rule coefficients into the kink functions yields the δ coefficients:

$$\delta' = \frac{-(\gamma'_2 + \gamma_Y \alpha'_{1,2})}{\gamma_1 + \gamma_Y \alpha_{1,1}} = \frac{(\gamma'_2 - \alpha'_{1,2})}{\alpha_{1,1}}.$$

Continuity at the kink requires:

$$\alpha'_{2,2} = (\alpha_{1,1} - \alpha_{2,1}) \delta' + \alpha'_{1,2}.$$

Replacing δ :

$$\alpha'_{2,2} = (\alpha_{1,1} - \alpha_{2,1}) \frac{(\gamma'_2 - \alpha'_{1,2})}{\alpha_{1,1}} + \alpha'_{1,2}.$$

From the complementary slackness condition we know that when the constraint binds $b' = mY$, implies that borrowing does not depend on the previous level of debt $\{B_{-1}\}$, so we can set $\alpha_{2,1} = 0$ and obtain:

$$\begin{aligned} \alpha'_{2,2} &= \alpha_{1,1} \frac{(\gamma'_2 - \alpha'_{1,2})}{\alpha_{1,1}} + \alpha'_{1,2} \\ \alpha'_{2,2} &= \gamma'_2. \end{aligned}$$

Which correctly imposes the constraint. The last step is to verify which side of the constraint corresponds to the non-binding regime. Following the derivation in the main text, it is easy to show that, $c(\vartheta) = -\alpha_{1,1} < 0$. Hence for the constraint to be non-binding, $h(\cdot) > 0$, we need $x_1 < \delta' X_2$.

The free coefficients $\alpha_{1,1}, \alpha_{2,1}$ defining the decision rules, can be obtained through the following numerical algorithm:

Algorithm 4 (PLC Algorithm)

1. Define a grid $\mathcal{S} \subset [B_{min}, B_{max}] \times [Y_{min}, Y_{max}]$. For the PLC algorithm we construct the solution grid using the Smolyak algorithm for two dimensions and with an approximation order of $\mu = 2$, delivering a total of 13 grid points—see Maliar and Maliar (2014) for details. The bounds of the hypercube are constructed using simulated series under the initial guess of the model, such that we cover 99% of the distribution of the endogenous and exogenous state variables.

2. Start with guess $\theta = \{\alpha_{1,1}, \alpha_{1,2}\}$ and use (A.26) to construct $b'(B, Y)$ for every grid point $(B, Y) \in \mathcal{S}$. Use the resource constraint to compute $C(B, Y) = Y + b'(B, Y) - RB$ which will also be piece-wise linear by construction.

3. For every grid point (B, Y) , compute the following residuals:

(a) When the constraint is slack ($x_1 > \delta'X_2$):

$$\begin{aligned}\lambda(B, Y) &= 0 \\ \mathcal{R}(B, Y) &= C(B, Y)^{-\gamma} - \beta R \mathbb{E}_{Y'|Y} (C(b'(B, Y), Y')^{-\gamma})\end{aligned}$$

(b) When the constraint is binding ($x_1 \leq \delta'X_2$):

$$\begin{aligned}\lambda(B, Y) &= C(B, Y)^{-\gamma} - \beta R \mathbb{E}_{Y'|Y} (C(b'(B, Y), Y')^{-\gamma}) \\ \mathcal{R}(B, Y) &= \lambda(B, Y) (mY - b'(B, Y))\end{aligned}$$

The expectations are approximated using a Gaussian-Quadrature integration rule of order 5.

4. The PLC solution solves the following minimization problem: $\min_{\theta} \frac{1}{M} \sum_{i=1}^M \mathcal{R}(B, Y)^2$

F.2.3 OccBin Algorithm

The toolkit OccBin (Guerrieri and Iacoviello, 2015) implements a non-linear solution procedure based on a modified shooting algorithm. To solve the model, it is necessary to define two sets of linearized equilibrium conditions that correspond to a reference regime, with a binding constraint, and an alternative regime, when the constraint does not bind but it is expected to bind again the future. OccBin produces a decision rule of the form:

$$X_t = \mathbf{P}(X_{t-1}, \epsilon_t) X_{t-1} + \mathbf{D}(X_{t-1}, \epsilon_t) + \mathbf{Q}(X_{t-1}, \epsilon_t) \epsilon_t \quad (\text{A.27})$$

in which the evolution of all endogenous variables in the model, X_t , can be expressed as function of their past X_{t-1} and the realization of an exogenous shock ϵ_t . The solution of the model with occasionally binding constraints from periods $t = 1, \dots, T$, is characterized by a sequence of time-varying matrices $\{\mathbf{P}(\cdot)_t, \mathbf{D}(\cdot)_t, \mathbf{Q}(\cdot)_t\}_{t=1}^T$. In terms of the simple consumption-savings problem, we have $X_t = [B_{t+1}, C_t, \lambda_t, Y_t]'$.

The decision rule in equation (A.27) shows that the solution matrices are state-dependent and that the solution need not be piece-wise linear, despite the fact that the time-varying matrices are obtained from linear approximations around the binding and non-binding regimes of the original problem. The construction of the time-varying matrices requires an iterative procedure that is briefly described in the following algorithm:

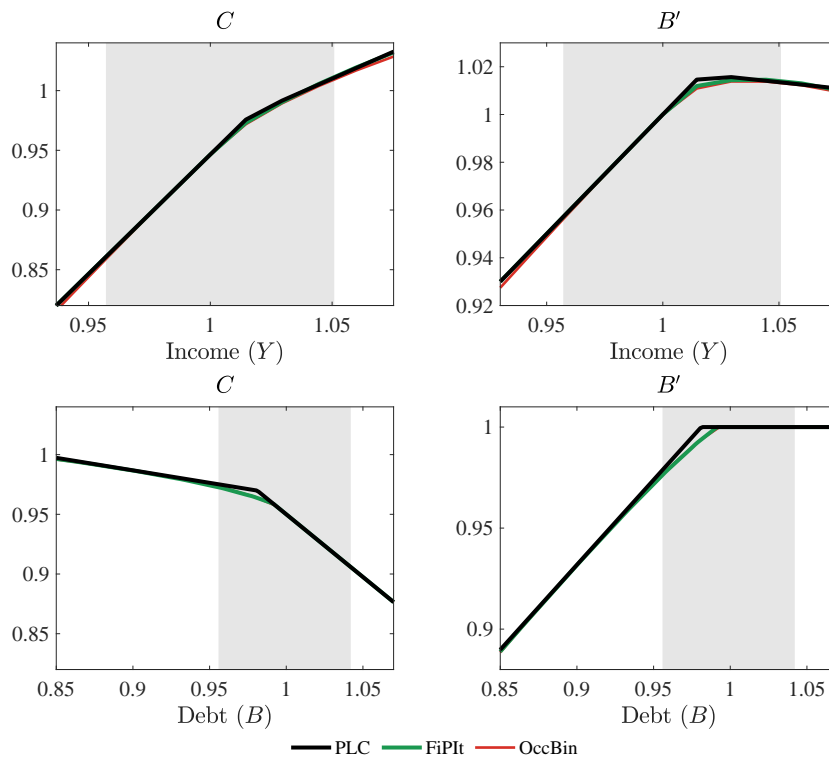
Algorithm 5 (OccBin Algorithm)

1. Take a generic period t , guess a sequence of regimes, such that for $t \geq T$, the model returns and forever remains in the reference regime. In our case, this corresponds to the case in which the borrowing constraint is binding.
2. For period $t \geq T$, solve the system of expectational equations that correspond to the reference regime: $\mathcal{A}_1 E_t X_{t+1} + \mathcal{A}_0 X_t + \mathcal{A}_{-1} X_{t-1} + \mathcal{B} \epsilon_t = 0$. Where the matrices $\mathcal{A}_1, \mathcal{A}_0, \mathcal{A}_{-1}, \mathcal{B}$ are functions of the model parameters and E_t is the expectations operator. Standard solution procedures, yield a decision rule of the form: $X_t = \mathcal{D} + \mathcal{P} X_{t-1} + \mathcal{Q} \epsilon_t$. Given we assumed there are no additional shocks beyond period t and that we approximate the solution around the steady state of the reference regime, we have $\mathcal{P}_T = \mathcal{P}$, $\mathcal{D}_T = \mathbf{0}$, and $X_T = \mathcal{P} X_{T-1}$.
3. For periods $t < T$ proceed as follow to compute the solution for X_t :
 - (a) If the guess implies the reference regime in period $T-1$, solve the matrix equation: $\mathcal{A}_1 \mathcal{P}_T X_{T-1} + \mathcal{A}_0 X_{T-1} + \mathcal{A}_{-1} X_{T-2} + \mathcal{B} \epsilon_{T-1} = 0$.
 - (b) If the guess implies a change to the alternative regime, solve the matrix equation: $\mathcal{A}_1^* \mathcal{P}_T X_{T-1} + \mathcal{A}_0^* X_{T-1} + \mathcal{A}_{-1}^* X_{T-2} + \mathcal{C}^* + \mathcal{B}^* \epsilon_{T-1} = 0$. The matrices $\mathcal{A}_1^*, \mathcal{A}_0^*, \mathcal{A}_{-1}^*, \mathcal{B}^*, \mathcal{C}^*$ correspond to the linearized system of equations in the alternative regime when the borrowing constraint is slack.
 - (c) The solution yields matrices $\mathcal{P}_{T-1}, \mathcal{D}_{T-1}, \mathcal{Q}_{T-1}$. Given X_{T-1} , solve for X_{T-2} and repeat this step for all $t < T$. Recall the assumption of no future shocks, hence \mathcal{Q}_t is irrelevant beyond period t and can be set to the null matrix.
4. Given X_{t-1}, ϵ_t , and the sequence of matrices $\mathcal{P}_t, \mathcal{D}_t$ and \mathcal{Q}_t , simulate the model forward using $X_t = \mathcal{D}_t + \mathcal{P}_t X_{t-1} + \mathcal{Q}_t \epsilon_t$. If the implied regimes in the forward simulation coincide with the regimes guessed in Step 1, a solution is given by $\mathbf{P}(X_{t-1}, \epsilon_t) = \mathcal{P}_t$, $\mathbf{D}(X_{t-1}, \epsilon_t) = \mathcal{D}_t$ and $\mathbf{Q}(X_{t-1}, \epsilon_t) = \mathcal{Q}_t$. Otherwise, update the guess of regimes and return to Step 1.

Table A-3: Solution Times

	FiPIt	OccBin	PLC
Time in seconds	10.6	3.9	0.5

Figure A-2: Decision Rules



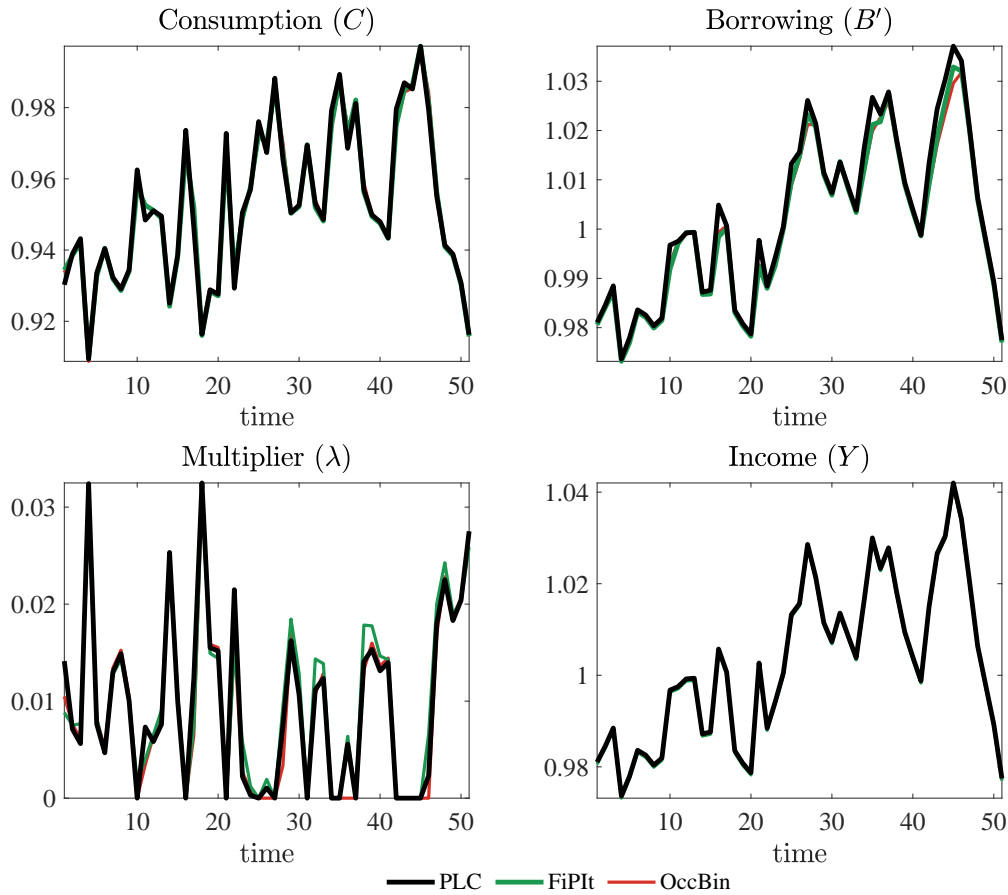
F.3 Quantitative Comparison

We set the following parameters to perform a numerical evaluation of the three solution algorithms for the consumption-savings problem. $R = 1.05$, $\beta = 0.945$, $\rho = 0.9$, $\sigma = 0.010$, $m = 1$, $\gamma = 1$. The parameter configuration satisfy the condition $\beta R < 1$ which ensures the existence of a stationary equilibrium in the incomplete market setting of this model.

We solve the model and compute solution times for the three methods. To compare with OccBin, we record the time it takes to construct decision rules over the same solution grid used to for the global solution implemented with the FiPIt method. The solution times for the three algorithm are reported in Table A-3.

Figure A-2 shows two different slices of the decision rules implied by each of the three

Figure A-3: Simulated Paths



solution algorithms. The top row shows the decision rules for consumption and borrowing decisions as a function of income, while holding the current period level of debt at its steady state of $B = 1$. The second row shows the decision rules for consumption and borrowing decisions as a function of the current level of debt, while holding the income realization at its steady state level of $Y = 1$. In each panel, the grade shaded areas show 95% of the ergodic distribution of the associated state variable.

Figure A-3 shows a simulated path of 50 observations for the three solution algorithms. We show the time path for consumption, borrowing and the Lagrange multiplier as well as the exogenous sequence of income realizations.

One may also want to see how the three methods compare in terms of where the constraint binds in the state space. To investigate this, we compare the three algorithms in terms of their predictions about the constraint binding. Taking the global solution as the “truth”, we

find that when the constraint is slack in the global solution, the other two algorithms also predict that the constraint will be slack with over 99.7% probability. When the constraint in the global solution is binding, that is when $\lambda > 0$, then for a vast majority of the cases, both solutions find that the constraint will be binding as well. In particular, if we look at quartiles of λ values obtained in the global solution conditional on $\lambda > 0$, in the third and fourth quartiles, the match is virtually perfect, and in the second quartile match is over 92%. Only in the first quartile – when λ is small – both methods deliver a positive λ in about 20% of the cases.

Looking at these results, we conclude that the PLC algorithm produces results that are very close to OccBin. Relative to the global solution, the decision rules show small differences around the kink, but the simulations show that this is not consequential: consumption simulation is virtually indistinguishable while the borrowing simulation shows slightly bigger but still small deviations. Similarly the multiplier is slightly different especially when it is positive and small, but in terms of identification of when the constraint binds, the three methods broadly agree. At the end of the day, for estimation purposes the simulation is really what matters and the three methods very much agree, with PLC taking about 1/20 of the time of the global solution and about 1/8 of the time of OccBin.