Estimation of DSGE Models with the Effective Lower Bound

Gregor Boehl[a,*], Felix Strobel[b]

[a] *University of Bonn*
[b] *Deutsche Bundesbank*

October 30, 2023

Abstract

We propose a new approach for the efficient and robust Bayesian estimation of medium- and large-scale DSGE models with occasionally binding constraints. At its core lies the Ensemble Kalman filter, a novel nonlinear recursive filter, which allows for fast likelihood approximations even for models with large state spaces. We combine the filter with a computationally efficient solution method for piece-wise linear models a state-of-the-art MCMC sampler. Using artificial data, we demonstrate that our approach accurately captures the true parameters of models with a lower bound on nominal interest rates, even with very long lower bound episodes. We use the approach to analyze the US business cycle dynamics until the Covid-19 pandemic, with a focus on the long lower bound episode after the Global Financial Crisis.

Keywords: Effective Lower Bound, Bayesian Estimation, Great Recession, Nonlinear Likelihood Inference, Ensemble Kalman Filter
JEL: C11, C63, E31, E32, E44

1 Introduction

More than a decade ago, the Financial Crisis and the subsequent Great Recession did not only wreak havoc on the US economy, but it also shook the macroeconomic profession to the core. In

[⋆] Some of the content in this paper was previously circulating as part of a draft with the title "US Business Cycle Dynamics at the ZLB" and as "Solution, Filtering and Estimation of Models with the ZLB". We are grateful to Alex Clymo, Macro Del Negro, Simon Gilchrist, Gavin Goy, Alexander Meyer-Gohde, Daniel M. Rees, Alexander Richter, Mathias Trabandt, Carlos Zarazaga, two anonymous referees, and participants of the 2018 Stanford MMCI Conference, the 2018 EEA Annual Congress, the 2018 VfS Jahrestagung and a seminar at the Deutsche Bundesbank for discussions and helpful comments on the contents of this paper. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of Deutsche Bundesbank. Part of the research leading to the results in this paper has received financial support from the Alfred P. Sloan Foundation under the grant agreement G-2016-7176 for the MMCI at the IMFS, Frankfurt. Gregor Boehl gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG) under CRC-TR 224 (project C01 and C05) and under project number 441540692.

[*] Corresponding author. Address: Institute for Macroeconomics and Econometrics, University of Bonn, Adenauer-allee 24-42, 53113 Bonn, Germany

Email address: `gboehl@uni-bonn.de`

response, theoretical approaches to enhance our understanding of these dramatic events quickly flourished in large numbers. However, only few attempts were made to bring these models to the data. This is because the long-binding effective lower bound (ELB) on nominal interest rates presents a formidable challenge for the empirical evaluation of economic models: Conventional econometric methods for structural models are unable to handle the non-linearity implied by an occasionally-binding constraint such as the ELB, and most existing alternatives are highly demanding computationally.

In this paper, we offer a way forward by proposing a novel nonlinear Bayesian likelihood approach that allows us to estimate macroeconomic models while accounting for the nonlinear effects of an occasionally-binding constraint (OBC). At the heart of our approach we introduce the Ensemble Kalman filter (Evensen, 1994, 2009, *EnKF*) into the literature on estimating DSGE models. We demonstrate that the EnKF can be applied in the context of nonlinear DSGE models and delivers a good approximation of the likelihood even of high-dimensional models. We pair the filter with computationally efficient solution method and a novel MCMC sampler to be able to estimate even large models at very moderate computational costs: The piecewise-linear solution method developed in Boehl (2022a) solves for the ELB as an occasionally binding constraint and provides a significant increase in speed compared to alternative algorithms, whereas the DIME Markov chain Monte Carlo method of Boehl (2022b) allows us to quickly sample from – possibly bimodal – high-dimensional posterior distributions *in parallel*.

We provide an easy-to-use reference implementation of the set of methods presented here: the Pydsge package. Pydsge is freely available and actively developed on Github. Next to the solution method, the filter and estimation routines, we provide a model parser similar to the one in Dynare and ample documentation to make our methods easily accessible.[1]

Applying this new approach to the estimation of nonlinear DSGE models, we illustrate that analyzing the Great Recession through the lens of models that have been calibrated or estimated only on pre-2008 data can generate misleading conclusions.[2] We estimate the canonical medium-scale model of Smets and Wouters (2007) on a sample that extends to 2019, thereby also including the first exit from the ELB in our estimation. The results underline that including the observations of the ELB period in the estimation has highly relevant implications on the business cycle properties

[1] The package is developed on Github and documentation is located at `pydsge.readthedocs.io`. It is written in the powerful open-source multi-purpose language Python. We like to explicitly promote free and open software.

[2] Prominent models that have been calibrated or estimated on pre-crisis data only to study the post-2008 dynamics include, e.g., Gertler and Karadi (2011); Christiano et al. (2015); Carlstrom et al. (2017). Others use post-2008 data, but ignore the ELB constraint (e.g., Kollmann et al., 2016; Fratto and Uhlig, 2020). Recently, some researchers have accounted for both, post-2008 US data and the ELB (e.g., Gust et al., 2017; Kulish et al., 2017; Cai et al., 2019; Cozzi et al., 2021)). Boehl et al. (forthcoming) show that their model estimated to post-2008 data uncovers deflationary effects of quantitative easing that are absent when estimating the model only on pre-crisis data.

of the model. We compare the decomposition of macroeconomic dynamics derived from the model estimated on the full sample with a decomposition of these dynamics as implied by using pre-2008 data only. This exercise reveals that the sample choice substantially affects the quantitative contribution of the different driving forces in the model. In the full sample, elevated risk premiums in household financing are the dominant driver of the crisis. In contrast, the analysis based on pre-crisis data overstates the importance of shocks to firms' investment financing. In addition, we show that trying to circumvent the technical challenges posed by the ELB by estimating the model linearly on the full sample results in statistically different parmeter estimates. This cautions against ignoring the non-linearity associated with an occasionally binding constraint regardless of the application at hand.

The EnKF is a recursive filter that approximates the standard Kalman filter by representing the distribution of the state as a small ensemble of vectors, which traverses through time. Consequently, each likelihood evaluation requires only a small amount of state-space evaluations. For each newly available observation, the ensemble members are updated by linear shifting instead of re-weighting (as with the particle filter). It can hence be seen as a hybrid of the particle filter and the conventional Kalman filter. The ensemble representation and shifting-based updates make the EnKF computationally feasible for models even with extremely high-dimensional state spaces.

Although the EnKF – as pointed out by Katzfuss et al. (2016) – is remarkably unknown in the statistics and econometrics community, it has been used in in a wide range of applications in the fields of meteorology, oceanography and hydrology. Stroud and Bengtsson (2007) and Frei and Künsch (2012) successfully use it for the Bayesian parameter inference of nonlinear models.

Importantly, while we show that the EnKF works well for estimating models with the ELB, it is potentially applicable to the large class of models with nonlinearities, e.g. for models with aggregate uncertainty. In any of these applications the EnKF will be computationally more efficient than the particle filter, even compared to its more refined versions such as the adapted particle filter (Herbst and Schorfheide, 2016) or the tempered particle filter (Herbst and Schorfheide, 2019). While under the hood, the EnKF implicitly assumes a linear Gaussian state-space model, it has turned out to be surprisingly robust to deviations from linearity as well as from Gaussianity, even for applications with tens of millions of dimensions (Katzfuss et al., 2016). Yet, the performace of the EnKF will depend on the degree of nonlinearities of the model at hand.

To validate our approach, and to verify that it is able to recover a credible estimate of the model parameters, we test it on a large artificial dataset in the spirit of Atkinson et al. (2020). As the data generating process, we use the full medium-scale model of Smets and Wouters (2007) that we estimate on US data beforehand. We generate 100 artificial time series from simulating the model: 50 for which the ELB is not binding at all, and 50 in which the ELB is binding for exactly 30 quarters. This dataset allows for the direct comparison with the analysis of Atkinson et al. (2020), who compare the performance of several nonlinear filters in the estimation of a small-scale model.[3] Across datasets, the resulting parameter estimates suggest that our tools are indeed able to provide credible and precise parameter estimates at limited computational costs.

We further add to this literature a procedure of nonlinear path-adjustment, which extends the ensemble version of the Rauch-Tung-Striebel smoother (Rauch et al., 1965; Raanes, 2016). This is necessary for counterfactual analysis, which requires the series of shock innovations to fully respect the nonlinear transition function, while taking the smoothed distribution of states into account. Additionally, we propose a method to compute historic shock decompositions of models with occasionally binding constraints. Importantly, the weighing scheme that we suggest results in decompositions that are conditionally linear and independent of any ordering effects of the shocks.

We compare our approach to the *inversion filter* (IVF), a popular filter for the estimation of models with OBCs. (see, e.g. Cuba-Borda et al., 2019). In contrast to the EnKF, the IVF abstracts from measurement errors and any uncertainty surrounding the initial state, thereby allowing for a direct mapping between shocks and observables. The IVF then relies on root finding methods to solve for a shock vector that satisfies the transition function for a given vector of observables.[4] We document that this approach may have limitations that are related to the invertibility of the transition function: in models with occasionally binding constraints, for a given shock there often exist multiple spell durations which form an equilibrium (see, e.g., Holden, 2017). Hence, the mapping from observables to shocks may not be unique. In addition, in some cases, a mapping may exist but the root finding algorithm may simply not converge.[5]

Using our artificial datasets, we show in Section 5 that the above issues of the IVF can be very relevant: if the mapping is non-unique the IVF accepts any shock vector that satisfies the transition function independently of how likely it is.[6] Thus, the shock vector picked by the IVF may crucially depend on the initial guess of the spell duration or the root finding algorithm. This introduces noise into the likelihood function, which may make sampling from the posterior distribution rather

[3]The artificial data used by Atkinson et al. (2020) to evaluate their tools is generated by a calibrated model that includes capital and sticky wages. The model they estimate abstracts from these features. This allows them to additionally investigate the bias introduced by model-misspecification.

[4]Note, that the use of a root finding algorithm becomes necessary due to a lack of a closed-form solution for linearized models with an occasionally binding constraint (Cuba-Borda et al., 2019; Atkinson et al., 2020).

[5]A second concern is that ignoring the uncertainty regarding the initial state may introduce a bias into the filter. In 5, we compare the performance of the EnKF and the IVF in artificial datasets. Whereas our findings suggest that in datasets without binding ELB the bias introduced by the IVF is moderate and not systematic, ignoring uncertainty regarding initial states reduces the estimation accuracy. In comparison with the EnKF, normalized root mean squared errors are on average 30% larger in estimations with the IVF.

[6]Note that by construction, Bayesian filters such as particle filter and EnKF will select shock vectors that are likely given their covariance, uncertainty surrounding initial states and measurement errors.

difficult.[7]

Related literature

The likelihood inference of nonlinear DSGE models is an active branch of the literature. Recently, a small number of papers have estimated economic models with an *endogenously* binding ELB. Insightful early work on estimating small-scale structural models with an endogenously binding ELB includes Richter and Throckmorton (2016); Keen et al. (2017); Plante et al. (2018a). Gust et al. (2017) estimate a downsized and globally solved version of the medium-scale model using the particle filter. While this is a challenging task, it comes at a very high computational cost, which requires excellent computational skills. In comparison, researchers can easily adopt the approach presented here using the Pydsge package.

Conceptually, the beauty of the particle filter (PF) is that for infinitely many particles it yields unbiased parameter estimates, independently of the model's degree of non-linearity. In practice, however, it comes with some drawbacks that the EnKF does not have. A common problem with the PF is that with a finite number of particles, in the case of an sufficiently large model, the weights of all but one particle may essentially become zero, leading to a poor approximation of the state distribution. The shifting-based updating step of the EnKF avoids this degeneracy problem of such re-weighting-based algorithms. Naturally, the larger the model, the more it is prone to degeneracy, requiring a rapidly increasing number of particles due to the *curse of dimensionality*. This makes the use of the particle filter computationally very costly even for moderately large models.[8] Another advantage of the EnKF over the particle filter is that, if the initial state is sampled from a low-discrepancy sequence[9], the likelihood function is in fact continuous over the parameter space and contains no sampling noise, which eases posterior sampling.

Since the EnKF relies on stochastic linearization, the curse of dimensionsionality does not apply to it. This makes the EnKF preferable over the PF in terms of computational cost, since the number of particles required is several orders of magnitude below the number of particles needed by the PF. It is noteworthy that recent progress in this field (e.g. Herbst and Schorfheide (2016, 2019)) has managed to alleviate both the degeneracy problem as well as the curse of dimensionality and thus requires less particles. Yet, the computational costs of applying the EnKF remains significantly below that of these enhanced versions of the PF. When it comes to the IVF, the num-

[7]In our exercise, we employ samples, in which the ELB is binding for 30 periods. The acceptance ratio soon drops below 1%, therby preventing us to obtain a reliable posterior sample.

[8]Notable applications of the particle filter, include Gust et al. (2017) with 1,500,000 particles for a downsized version of the model of Smets and Wouters (2007), Atkinson et al. (2020) with 40,000 particles for a small-scale model, and Herbst and Schorfheide (2019).

[9]These are methods to construct a sample in such a way that, roughly speaking, it most perfectly represents a target distribution even for small samples.

ber of function evaluations needed by the EnKF is comparable but heavily depends on the number of iterations required by the IVF to converge.

Aruoba et al. (2021) provide a set of methods that alleviate some of the computational costs, in particular by suggesting a conditional optimal particle filter that is optimized to deal with models with OBCs. This approach has a number of advantages, especially as it allows to capture precautionary behavior. However, it is still subject to the curse of dimensionality and can, for the medium-scale model considered here, be expected to be considerably slower than our approach.

Kulish et al. (2017) suggest to circumvent the nonlinear filtering problem by treating the expected durations of the ELB as parameters in their estimation. Conditional on that, the estimated model is again linear. However, this approach may have limitations because the MCMC procedure may not always be capable to deal with such a high dimensionality of the parameter space. Additionally, introducing discrete parameters with potentially non-smooth effects may add further difficulties to the sampling procedure. In practice, these two points can potentially limit parameter identification. A similar approach is to directly feed survey data on interest rate expectations into a model augmented by news shocks and forecasting errors (see, e.g., Cai et al., 2019). As with the Kulish et al. (2017) approach, this results in a (conditionally) linear model. However, both procedures cannot capture the endogenous nonlinearity of the model and the shocks implied by the filter may actually imply different ELB durations than the ones initially imposed. This as well can potentially distort the parameter estimates.[10] In the context of the ELB, our methodology presents an alternative approach that allows for an endogenous generation of ELB spell durations. A powerful advantage of our approach also is that it can be used in the context of *any* occasionally binding constraint also when data on agents expectation on the duration of the binding constraint is not available or not reliable. This can, for example, be relevant in the context of downward nominal wage rigidities or financial constraints.

With the Extended Kalman filter (EKF, Smith et al., 1962; McElhoe, 1966) and the Unscented Kalman Filter (UKF, Julier et al., 2000) further alternatives exist for cases in which the nonlinearities are known to be rather mild. However, the EKF is known to easily diverge if nonlinearities become more severe, or if the time series of observables is very volatile. The performance of the UKF hinges strongly on the quality of the parametrization of its Sigma points and can be prone to divergence as well. Compared to the UKF, the EnKF does not rely on parameterized deterministic sampling techniques and is hence, apart from the choice of the number of particles, parameter-free. We experimented with both, the EKF and the UKF, and can confirm that both do not work well in the context of nonlinear DSGE models, and can return noisy likelihood estimates.

[10]The discrepancies between simulated spell durations and durations as imposed during estimation are exploited by Jones et al. (2018), who similar to Kulish et al. (2017), include the spell durations in the sampling procedure. They label such discrepancies as forward guidance shocks.

Our analysis of the post-2008 US macroeconomic dynamics confirms previous findings of Gust et al. (2017) and Kulish et al. (2017), who consider a binding ELB within comparable models but based on different methodology. This lends credence to our analysis and suggests, in comparison with the fully nonlinear method of Gust et al. (2017), that the loss of precision that might incur due to the use of a piecewise-linear solution method is small.[11]

Our finding that risk premium shocks have been the major drivers of the Great Recession is shared, e.g., by Kulish et al. (2017) and Cai et al. (2019). Several authors associate this shock with the importance of household financing for the Great Recession (see, e.g. Mian and Sufi, 2014, 2015; Kehoe et al., 2020).[12] Nevertheless, a large share of the previous literature attempts to explain the Great Recession via disturbances and frictions associated to firms' investment finance. This includes papers, which directly discuss risks on the firms' balance sheet such as, e.g., Christiano et al. (2014), and extends as well to contributions that focus on vulnerabilities in the banking sector, which in turn affect firm's investment financing such as, e.g., Gertler and Karadi (2011); Carlstrom et al. (2017). Many of these papers conduct their analysis by means of calibrated models, or models estimated on pre-2008 data. Our results suggests that, rather than focussing on firms' investment, a closer investigation on the role of household financing might be warranted.

We proceed as follows: Section 2 lays out the set of novel methods. Section 3 contains the application of the approach on US data and the resulting interpretations of the Great Recession through the lens of the estimated model with and without the use of post-crisis data. In Section 4 we test our set of methods on artificial data, and discuss its accuracy. Section 5 contains a comparison of the EnKF and the inversion filter. Section 6 concludes.

2 Conceptual Framework

Data samples in which the ELB binds pose a host of technical challenges for the estimation of DSGE models. These are related to the solution, likelihood inference, and posterior sampling of models in the presence of an occasionally binding constraint (OBC). While methods to solve models with OBCs exists, and – likewise – nonlinear filters are available, the combination of both is computationally very expensive for medium-scale models. Hence, very few examples in the literature were able to follow this approach (e.g., Gust et al., 2017; Kulish et al., 2017).[13]

[11]This is in line with Atkinson et al. (2020), who compare piecewise-linear OBC solutions with fully global methods. They acknowledge that the fully nonlinear solution entails some nice properties (e.g. capturing the effects of aggregate uncertainty), but prefer the piecewise-linear solution as it allows for larger, less misspecified models.

[12]Fisher (2015) provides another interpretation of the shock as an economy-wide increase in the demand for liquid or safe assets.

[13]The estimation of DSGE models with a binding ELB was pioneered by work on small-scale NK models. See, e.g., Keen et al. (2017); Aruoba et al. (2018, 2021); Plante et al. (2018b).

In this section, we summarize the set of novel methods that allow us to conduct the estimation of medium-scale models in the presence of an occasionally binding ELB. Next to the EnKF, this includes an ensemble version of the Kalman smoother and an extension to extract the precise series of shocks driving the dynamics. We also summarize the piece-wise linear solution method of Boehl (2022a) and, for posterior sampling, the differential-independence mixture ensemble Monte Carlo Markov chain method (DIME MCMC) developed for DSGE models in Boehl (2022b).

The advantage of the solution method over the widely used Occbin by Guerrieri and Iacoviello (2015) is its speed, as it is based on closed form solutions and circumvents simulations on anticipated trajectories and matrix inversions at runtime.[14] The main advantage of the DIME sampling method is that it uses a large number of chains, which are self-tuning, easy to parallelize, and – as an ensemble of chains – are robust against local maxima. The method also performs well on odd-shaped or multimodal distributions.

2.1 A method to deal with occasionally binding constraints efficiently

Throughout this paper, we apply the solution method for DSGE models with OBCs that is presented in Boehl (2022a). We refer to the original paper for details. The model is linearized around its steady state balanced growth path and thereby implicitly detrended. Respecting the ELB, the original model with variable vector $y_t \in \mathbb{R}^{n_y}$ and shock vector $\varepsilon_t \in \mathbb{R}^{n_z}$ can be represented as a piecewise linear model with

$$A \begin{bmatrix} c_t \\ s_{t-1} \end{bmatrix} + b \max\left\{ p \begin{bmatrix} E_t c_{t+1} \\ s_t \end{bmatrix} + m \begin{bmatrix} c_t \\ s_{t-1} \end{bmatrix}, \ \bar{r} \right\} = E_t \begin{bmatrix} c_{t+1} \\ s_t \end{bmatrix}, \tag{1}$$

where $\begin{bmatrix} c_t \\ s_{t-1} \end{bmatrix}$ is a re-ordering of $\begin{bmatrix} y_t \\ \varepsilon_t \end{bmatrix}$: s_{t-1} contains all the (latent) state variables and the current shocks, and c_t contains all forward looking variables. A is the system matrix and $\bar{r}$ is the minimum value of the constrained variable r_t (which is the nominal interest rate for our purpose). The constraint is included with $r_t = \max\left\{ pE_t \begin{bmatrix} c_{t+1} \\ s_t \end{bmatrix} + m \begin{bmatrix} c_t \\ s_{t-1} \end{bmatrix}, \ \bar{r} \right\}$. p and m measure how r_t is affected by other variables, and the vector b contains the effects of r_t onto all other variables. Then, denote by $(k, l) \in \mathbb{N}_0^+$ the expected duration of the ELB spell and the expected number of periods before the ELB binds.

It can be shown that the rational expectations solution to Equation (1) for the state s periods

[14]While in our application, we focus on the ELB constraint, in principle, the solution method can handle multiple constraints at the same time.

ahead, (c_{t+s}, s_{t+s-1}), can be expressed in terms of s_{t-1} and the expectations on k and l as

$$F_s(l,k,s_{t-1}) = A^{\max\{s-l,0\}}\hat{A}^{\min\{l,s\}}\begin{bmatrix} f\,(l,k,s_{t-1}) \\ s_{t-1} \end{bmatrix} + (I-A)^{-1}\left(I - A^{\max\{s-l,0\}}\right)b\bar{r}, \tag{2}$$

$$= E_t\begin{bmatrix} c_{t+s} \\ s_{t+s-1} \end{bmatrix}, \tag{3}$$

where $\hat{A} = (I-bp)^{-1}(A+bm)$ and

$$f(l,k,s_{t-1}) = \left\{ c_t : \Psi A^k \hat{A}\begin{bmatrix} c_t \\ s_{t-1} \end{bmatrix} = -\Psi(I-A)^{-1}(I-A^k)b\bar{r} \right\}. \tag{4}$$

Here, $\Psi = \begin{bmatrix} I & -\Omega \end{bmatrix}$ where $\Omega : c_t = \Omega s_{t-1}$ represents the linear rational expectations solution of the unconstrained system as given, e.g., in Blanchard and Kahn (1980).

Finding the equilibrium values of (l,k) must be done numerically. The crucial advantage of the above representation over alternative methods such as Guerrieri and Iacoviello (2015) is that the simulation of anticipated trajectories (and matrix inversions at runtime) can be avoided when iterating over (l,k). This lends a reduction in computation time by a factor of roughly 1,500, which is necessary for our application. Ultimately, the resulting transition function is a nonlinear state-space representation.

2.2 Filter

The nonlinear filtering methodology we apply is an adaptation of the Ensemble Kalman Filter (Evensen, 1994, EnKF) for the general type of nonlinear problems faced in macroeconomics. Denote a nonlinear hidden Markov-Model (HMM) by

$$x_t = g(x_{t-1}, \varepsilon_t) \tag{5}$$

$$z_t = h(x_t) + \nu_t \tag{6}$$

with exogenous economic innovations $\varepsilon_t \sim \mathcal{N}(0,Q)$ and measurement noise $\nu_t \sim \mathcal{N}(0,R)$. g is the state-transition function, i.e. in our case the function that implicitly assigns a set of (l,k) to a state-shock combination (x_{t-1}, ε_t). h is an observation function, mapping from states to observables. $x_t \in \mathbb{R}^n$ can, depending on the definition of g and h, either be the full variable vector y_t or just the state vector s_t

Let $\mathbf{X}_t = [\mathbf{x}_t^1, \cdots, \mathbf{x}_t^N] \in \mathbb{R}^{n\times N}$ be the *ensemble* at time t consisting of N vectors of the state. Denote by $(\bar{x}_t, P_t)$ the mean and the covariance matrix of the unconditional distribution of states for period t. Initialize the ensemble by drawing N samples from the prior state distribution (not to be confused with the parameter priors in the context of the Bayesian inference of parameter values, that we discuss below)

$$\mathbf{X}_0 \overset{N}{\sim} \mathcal{N}\left(\bar{x}_0, P_0\right). \tag{7}$$

Importantly, this distribution reflects any uncertainty about the initial state of the economy prior the first observation. We use latin hypercube sampling (McKay et al., 2000) to obtain $\mathbf{X}_0$. Such quasi-random low discrepancy series are a powerful tool to create prototypical samples of a target distribution, that are (almost) independent of the random seed (e.g., Niederreiter, 1988). During the estimations, we re-use the same underlying low-discrepancy sequence for the initial states for all likelihood draws to guarantee that the likelihood function is a smooth function over the parameter space.

Step 1: Predict

Predict the next (time-t) prior-ensemble $\mathbf{X}_{t|t-1}$ by applying the transition function to the posterior ensemble from last period. Use the observation function to obtain a prior ensemble of predicted observables $\mathbf{Z}_{t|t-1}$:

$$\mathbf{X}_{t|t-1} = g(\mathbf{X}_{t-1|t-1}, \boldsymbol{\varepsilon}_t), \tag{8}$$

$$\mathbf{Z}_{t|t-1} = h(\mathbf{X}_{t|t-1}) + \boldsymbol{\nu}_t, \tag{9}$$

where $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\nu}_t$ are each N realizations drawn from the respective distributions.

Step 2: Update

Denote by $\bar{\mathbf{X}}_t = \mathbf{X}_t(\mathbf{I}_N - \mathbf{1}\mathbf{1}^\intercal/N)$ the *anomalies* of the ensemble, i.e. the deviations from the ensemble mean. Define $\bar{\mathbf{Z}}_t$ likewise for the ensemble of predictions. Recall that the covariance matrix of the prior distribution at t is $\frac{\bar{\mathbf{X}}_t\bar{\mathbf{X}}_t^\intercal}{N-1}$. The Kalman mechanism then yields an *update*-step of

$$\mathbf{X}_{t|t} = \mathbf{X}_{t|t-1} + \bar{\mathbf{X}}_{t|t-1}\bar{\mathbf{Z}}_{t|t-1}^\intercal\left(\bar{\mathbf{Z}}_{t|t-1}\bar{\mathbf{Z}}_{t|t-1}^\intercal\right)^{-1}\left(z_t\mathbf{1}^\intercal - \mathbf{Z}_{t|t-1}\right). \tag{10}$$

The mechanism is similar to the unscented Kalman filter (UKF), developed by Julier and Uhlmann (1997), but with a particle representation of the state distribution instead of deterministic Sigma points, and statistical linearization instead of the unscented transform. The advantage of the EnKF over the UKF is that its output does not depend on the parametrization of the sigma points, which can be quite sensitive. Conceptionally the procedure suggested here can be seen as a *transposition* of the EnKF.[15]

[15]Notationally both are equivalent. The regular EnKF assumes the size of the state spaces to be larger than N, and accordingly the term $\left(\bar{\mathbf{Z}}_{t|t-1}\bar{\mathbf{Z}}_{t|t-1}^\intercal\right)$ to be rank deficient. The mechanism then builds on the properties of the pseudoin-

The likelihood at each iteration can then be determined by

$$\mathcal{L}_t = \varphi\left(z_t \middle| \bar{z}_t, \frac{\bar{\mathbf{Z}}_t\bar{\mathbf{Z}}_t^\intercal}{N-1} + R\right), \tag{11}$$

where φ denotes the PDF of the multivariate Gaussian distribution and $\bar{z}_t$ is the mean over $\mathbf{Z}_{t|t-1}$. Note that the calculation of the likelihood requires one prediction-updating loop for each observation. Each prediction step in turn requires N state-space evaluations. For all estimations and for the numerical analysis we use ensembles of $N = 400$ particles. For 120 observations, this would amount to 48,000 state-space evaluations – that is, calculations of (l, k) – per likelihood evaluation. This underlines why we require the very fast OBC solution method of Boehl (2022a).[16]

Strictly speaking, the EnKF only delivers the exact likelihood in linear systems (Katzfuss et al., 2016), as each state distribution – and thereby the inference of the likelihood – is based on a linear approximation around the ensemble mean.[17] This stands in contrast to the particle filter (PF), which can be shown to be an unbiased estimator also for non-linear transition functions. Nonetheless, as we show in Section 4, the bias of the EnKF is negligible in samples with a binding ELB. As an advantage over the PF, the EnKF avoids degeneracy issues (see e.g. Binning and Maih, 2015), a problem which is commonly mitigated by assuming counterfactually high measurement errors (MEs). This bears the risk of likelihood misspecification, where the misspecification error involved in PFs grows with the size of the assumed MEs if the true DGP has no or only small MEs (see, Cuba-Borda et al., 2019; Canova et al., 2020). In contrast, the EnKF can generally be used with very small MEs, and variants exist that allow filtering and likelihood inference without MEs.

More importantly, however, the EnKF enables us to estimate large-scale nonlinear systems, for which an estimation with particle filter is too costly. This facilitates the estimation of models with a rich set of features and helps to avoid the model-misspecification that may be the price for using smaller models, which the PF can estimate in an acceptable time frame. As Atkinson et al. (2020) highlight, in practice this type of model misspecification turns out to be far more severe.

2.3 *Smoothing and iterative path-adjusting*

For economic analysis we are also interested in the series of shocks, $\{\varepsilon_t\}_{t=0}^{T-1}$, that fully recovers the mode of the smoothened states. The econometric process of using *all* available information on *all* estimates is called smoothing. For this purpose, we extend the Rauch-Tung-Striebel smoother

verse (the latter provides a least squares solution to a system of linear equations), which is used instead of the regular matrix inverse.

[16]The number of particles is chosen to minimize the standard deviation of the likelihood approximation across random seeds. For the estimations in this paper, an average likelihood evaluation then takes a bit less than two seconds.

[17]For linear systems the EnKF gives results identical to the standard Kalman Filter.

(Rauch et al., 1965) in its ensemble formulation similar to Raanes (2016).

Denote by T the period of the last available observation and update each ensemble according to the backwards recursion[18]

$$\mathbf{X}_{t|T} = \mathbf{X}_{t|t} + \bar{\mathbf{X}}_{t|t}\bar{\mathbf{X}}^{+}_{t+1|t}\left[\mathbf{X}_{t+1|T} - \mathbf{X}_{t+1|t}\right]. \tag{13}$$

This creates a series $\{\mathbf{X}_{t|T}\}_{t=0}^{T}$ of representatives of the distributions of states at each point in time, reflecting all the available information. We now want to ensure that the mode of the distribution fully reflects the nonlinearity of the transition function while retaining a reasonably good approximation of the full distribution. We call this process *nonlinear path-adjustment*. It is important that the smoothened distributions are targeted instead of, e.g., just the distributions of observables and shocks. Only when the full smoothened distributions are targeted it can be maintained that *all* available information from the observables is taken into account. This procedure implicitly assumes that the smoothened distributions approximate the actual transition function sufficiently well and only minor adjustments remain necessary. Since in general there are (many) more states than exogenous shocks, the fitting problem is underdefined and matching precision will depend on the size of the relative (co)variance of each variable. Small observation errors lead to small variances around observable states and tight fitting during path-adjustment while loosely identified states grant more leeway.

Initialize the algorithm with $\hat{x}_0 = \mathrm{E}\{\mathbf{X}_{0|T}\}$ (the mean vector over the ensemble members), define $P_{t|T} = \mathrm{Cov}\{\mathbf{X}_{t|T}\}$ and for each period t recursively find

$$\hat{\varepsilon}_t = \arg\max_{\varepsilon} \left\{ \log \varphi\left(g(\hat{x}_{t-1}, \varepsilon) \middle| \bar{x}_{t|T}, P_{t|T} \right) \right\}, \tag{14}$$

$$\hat{x}_t = g(\hat{x}_{t-1}, \hat{\varepsilon}_t), \tag{15}$$

where φ again denotes the PDF of the multivariate Gaussian distribution. This can be done using standard iterative optimization methods.

The resulting series of $\hat{x}_t$ corresponds to the estimated mode given the initial mean and approximated covariances and is completely recoverable by $\hat{\varepsilon}_t$. Naturally, it represents the nonlinearity of the transition function while taking all available information into account. Since the deviation between mode $\hat{x}_t$ and mean $\bar{x}_t$ is in general marginal, we refer to $\{\hat{x}_t, P_t\}_{t=0}^{T}$ as the *path-adjusted*

[18] Although it is formally correct that

$$\bar{\mathbf{X}}_{t|t}\bar{\mathbf{X}}^{\intercal}_{t+1|t}\left(\bar{\mathbf{X}}_{t+1|t}\bar{\mathbf{X}}^{\intercal}_{t+1|t}\right)^{+} = \bar{\mathbf{X}}_{t|t}\bar{\mathbf{X}}^{+}_{t+1|t}, \tag{12}$$

the implementation using the LHS of this equation is numerically more stable when using standard implementations of the *pseudo-inverse* based on the SVD.

smoothed distributions.[19]

2.4 Posterior sampling

For posterior sampling we apply the differential-independence mixture ensemble Monte Carlo Markov chain method (DIME MCMC) developed in Boehl (2022b), which builds on ter Braak (2006). The DIME method is a member of the class of ensemble MCMC methods which, instead of relying on a single or small number of state-dependent chains (as e.g. in the random walk Metropolis-Hastings algorithm, RWMH), uses a large number of chains (also called the "ensemble", but in the context of posterior sampling). While, e.g., the conventional RWMH generates new proposals using a multivariate normal jump distribution centered at the iterate, the differential evolution algorithm generates new proposals for each chain by adding to the current point a fraction of the difference of two randomly chosen chains from the ensemble. Thus, the proposal density is endogenous and adapts to each updating step, thereby yielding a high speed of convergence. At the same time, the use of many chains ensures a broad search over the parameter space.[20]

As shown in Boehl (2022b), the main advantage of DIME is that it is self-tuning, easy to parallelize, and relatively robust against local maxima, which allows to use it to sample from potentially bimodal distributions. DIME explicitly renders any posterior mode search prior to sampling (as with RWMH) unnecessary because the algorithm itself is based on a heuristic global optimizer (i.e., the differential evolution method). The sampler even works well if large regions of the parameter space do not have a likelihood due to indeterminacy or explosive dynamics. The fact that DIME is easy to parallelize is very useful because, even when using the methodology above, the evaluation of the likelihood function is computationally expensive. Hence, DIME allows us to further reduce the computational burden considerably. For each estimation, we initialize an ensemble of 200 particles with the prior distribution and run 2500 iterations. Of these, we keep 500 as a representation of the posterior distribution. Thus, the posterior is represented by a sample of 200x500 = 100,000 parameter vectors. All estimations are conducted on a machine with 40 Intel Xeon E5 CPUs (3.10GHz) and 32 GB RAM and take about 3 hours in average.[21]

---

[19]Unfortunately the adjustment step can not be done during the filtering stage already. Iterative adjustment before the prediction step, would bias the transition of the covariance. Likewise, adjusting after the prediction step will require the repeating the prediction and updating step leading to a potentially infinite loop. See e.g. Ungarala (2012) for details.

[20]ter Braak (2006) provides a well-written introduction into the DE-MCMC and a comprehensive comparison to the conventional RWMH. Similar ensemble methods have been extensively applied in particular in astrophysics (see, e.g., Foreman-Mackey et al., 2013).

[21]Parallelization scales almost linearly, implying that one of the estimations presented here would take about 30 hours on a common quad-core PC.

3 Business Cycle Dynamics and the Effective Lower Bound

In this section, we apply our methods to analyze the business cycle dynamics at the ELB through the lens of a standard medium-scale DSGE model. We start with a brief look at the employed structural framework, followed by a discussion the data and the empirical treatment of the effective lower bound. We then discuss the parameter estimates and present the main implications of the estimated model for the dynamics of the great recession. Finally, we show that the additional post-2008 data points are crucial for the interpretation of the data, and lead to significantly different model dynamics compared to the model estimated on pre-crisis data only.

3.1 Model

In our analysis, we employ the canonical medium-scale framework by Smets and Wouters (2007) as a data generating process and use it to interpret the Great Recession. Following Del Negro and Schorfheide (2013), we detrend all non-stationary variables by

$$Z_t = e^{\gamma t + \frac{1}{1-\alpha}\widetilde{z}_t}, \tag{16}$$

where, γ is the steady-state growth rate of the economy and α is the output share of capital. $\widetilde{z}_t$ is the linearly detrended log productivity process that follows the autoregressive law of motion $\widetilde{z}_t = \rho_z \widetilde{z}_{t-1} + \sigma_z \epsilon_z$. For z_t, the growth rate of technology in deviations from γ, it holds that $z_t = \frac{1}{1-\alpha}(\rho_z - 1)\widetilde{z}_t + \frac{1}{1-\alpha}\sigma_z \epsilon_z$. We take into account the fact that the central bank is constrained in its interest rate policy by a zero lower bound (ELB) on the nominal interest rate. Therefore, in the linear model, it is that

$$r_t = \max\{\bar{r}, r_t^n\}, \tag{17}$$

with $\bar{r}$ being the lower bound value. Whenever the policy rate is away from the constraint, it corresponds to the notional rate, r_t^n, which, as in Smets and Wouters (2007), follows the feedback rule

$$r_t^n = \rho r_{t-1}^n + (1-\rho)\left(\phi_\pi \pi_t + \phi_y \widetilde{y}_t\right) + \phi_{dy}\Delta\widetilde{y}_t + v_{r,t}. \tag{18}$$

Here, π_t is the inflation rate, $\widetilde{y}_t$ is the output gap and $\Delta\widetilde{y}_t = \widetilde{y}_t - \widetilde{y}_{t-1}$ its growth rate. Parameter ρ expresses an interest rate smoothing motive by the central bank. ϕ_π, ϕ_y and ϕ_{dy} are feedback coefficients. The stochastic process $v_{r,t}$ follows an AR(1) process. Whenever the economy is at the ELB, the design of the central bank's policy rule allows agents to form expectations on when the notional rate will re-enter positive territory. That is, the design of the central bank's policy rule combined with the state of the economy governs agents expectations of the duration of the ELB spell. In addition to this systematic component of forward guidance, the innovation $v_{r,t}$ alters the path of the notional rate and, at the ELB in effect alters the expected duration of the lower bound spell. It can hence be viewed as a forward guidance shock whenever the economy is at the ELB. As the model is well known, we delegate a short description and the full set of linearized equilibrium conditions to Appendix B.

3.2 Data and Priors

For the quantitative analysis of the Great Recession and its aftermath, we use data from 1964:Q1 to 2019:Q4. Thereby we also capture the exit from the ELB at the end of 2015. The inclusion of the ELB period in the sample employed in the estimation matters for the model-implied interpretation of the Great Recession. To show this, we additionally consider a pre-crisis sample in our analysis, which extends from 1964:Q1 to 2008:Q4.

We estimate the model on seven observables. Those are real GDP growth, real consumption growth, real investment growth, labor hours, the log change of the GDP deflator, real wage growth, and the Federal Funds Rate.

The measurement equations that relate the model variables to our data series are

$$\text{Real GDP growth} = \overline{\gamma} + (y_t - y_{t-1} + z_t), \tag{19}$$

$$\text{Real consumption growth} = \overline{\gamma} + (c_t - c_{t-1} + z_t), \tag{20}$$

$$\text{Real investment growth} = \overline{\gamma} + (i_t - i_{t-1} + z_t), \tag{21}$$

$$\text{Real wage growth} = \overline{\gamma} + (w_t - w_{t-1} + z_t), \tag{22}$$

$$\text{Labor hours} = \bar{l} + l_t, \tag{23}$$

$$\text{Inflation} = \overline{\pi} + \pi_t, \tag{24}$$

$$\text{Federal funds rate} = 100\left(\frac{\overline{\pi}}{\beta\gamma^{-\sigma_c}} - 1\right) + r_t. \tag{25}$$

The construction of the observables is mostly standard and delegated to Appendix A. Consistent with the detrending of nonstationary variables, the growth rate of technology, z_t in deviations from its steady state enters the measurement equations.

Notably, we set the empirical lower bound of the nominal interest rate within the model to 0.05% quarterly. Setting it exactly to zero would imply that the ELB never binds in our estimations, as the observed series for the FFR stays strictly above zero. The value is chosen such that the ELB is considered binding throughout the period from 2009:Q1 to 2015:Q4. For the observable Federal Funds Rate we cut off any value below 0.05. This maintains that any observable value is also in the domain of model.[22]

[22] The lower bound for the quarterly nominal rate is $\bar{r} = -100(\frac{\overline{\pi}}{\beta\gamma^{-\sigma_c}} - 1) + 0.05$, where $\overline{\pi}$ is gross inflation and the parameters γ and σ_c denote the steady state growth rate and the coefficient of relative risk aversion, respectively.

We assume small measurement errors for all variables with a variance that is 0.01 times the variance of the respective series. Since the Federal Funds rate is directly observable we divide the measurement error variance here again by 100. Hence, the observables are de facto matched perfectly.

In the calibration of some parameters and the choice of the priors for the estimation of the others we mostly adopt the choices of Smets and Wouters (2007). An exception is our prior for $\overline{\gamma}$. Here, we follow Kulish et al. (2017). Importantly, they opt for a tighter prior for this parameter than Smets and Wouters (2007). Arguably the economy deviated strongly and persistently from its steady state during the Great Recession. In order to dampen the data's pull of the parameter down to the sample mean, we prefer the tight prior as well.[23]

3.3 Parameter estimates

The summary statistics of the posteriors for the structural parameters are presented in Table 1. We present estimates for the full sample, a pre-crisis sample without the post-crisis data, and the full sample when appying a linear filter only. Overall, both for the sample with the ELB period and for the pre-crisis sample, our results are broadly in line with the findings of previous work that estimates versions of this model. This lends credence to the results generated with the novel set of methods. However, there are some crucial differences between the estimates *across* samples.

We find that the coefficient of relative risk aversion, σ_c, is slightly above unity in the full sample whereas its mean is higher in the pre-crisis sample (1.5). Similarly, Kulish et al. (2017), who also include the last decade in their estimation, find σ_c to be close to unity, similar as in our pre-crisis sample. A value of σ_c close to one mutes the effect of variations in labor hours on consumption via the Euler equation, which is introduced through the nonseperabilities in preferences. The reduction of this channel prevents the strong drop in labor hours during the crisis to exert an excessive downwards pull on consumption.

Another difference lies in the estimate of the slope of the Phillips Curve. The pre-crisis estimate of $\zeta_p = 0.714$ is close to the value in the estimation by Smets and Wouters (2007). In contrast, the Calvo parameter of $\zeta_p = 0.904$ in the full sample supports the general notion that the New Keynesian Phillips Curve has flattened in the last decades. This finding is corroborated by estimates of Kulish et al. (2017).

Importantly, the persistence of structural shocks appears to have changed over the last decades. While the estimates of these parameters for the pre-crisis sample are well aligned with the results by Smets and Wouters (2007), the risk premium shock in the sample which includes the ELB episode displays a substantially higher persistence. This points to the increased importance of risk

[23]For wider priors we confirm unrealistically low estimates of the trend growth rate.

	Prior			Posterior								
				1964–2019			1964–2008			1964–2019: no OBC		
	distribution	mean	sd/df	mean	sd	mode	mean	sd	mode	mean	sd	mode
σ_c	normal	1.500	0.375	1.156	0.121	1.023	1.500	0.150	1.539	1.333	0.157	1.272
σ_l	normal	2.000	0.750	3.333	0.416	3.490	2.411	0.471	2.468	1.948	0.511	2.733
β_{tpr}	gamma	0.250	0.100	0.147	0.044	0.146	0.148	0.045	0.175	0.141	0.060	0.168
h	beta	0.700	0.100	0.635	0.042	0.667	0.590	0.054	0.560	0.484	0.050	0.500
S''	normal	4.000	1.500	5.140	0.637	5.574	4.435	0.890	4.444	3.340	0.793	3.383
ι_p	beta	0.500	0.150	0.657	0.058	0.651	0.425	0.109	0.395	0.343	0.103	0.234
ι_w	beta	0.500	0.150	0.528	0.092	0.586	0.493	0.106	0.582	0.600	0.131	0.531
α	normal	0.300	0.050	0.173	0.015	0.157	0.213	0.017	0.222	0.175	0.018	0.169
ζ_p	beta	0.500	0.100	0.904	0.016	0.900	0.714	0.042	0.670	0.845	0.032	0.860
ζ_w	beta	0.500	0.100	0.817	0.018	0.823	0.773	0.051	0.743	0.809	0.036	0.852
Φ_p	normal	1.250	0.125	1.440	0.058	1.412	1.591	0.067	1.629	1.421	0.070	1.380
ψ	beta	0.500	0.150	0.502	0.077	0.460	0.617	0.083	0.685	0.759	0.091	0.826
ϕ_π	normal	1.500	0.250	2.190	0.128	2.198	1.958	0.164	1.987	1.866	0.164	1.709
ϕ_y	normal	0.125	0.050	0.173	0.018	0.194	0.072	0.029	0.054	0.120	0.025	0.114
ϕ_{dy}	normal	0.125	0.050	0.254	0.018	0.258	0.250	0.023	0.263	0.268	0.027	0.262
ρ	beta	0.750	0.100	0.870	0.012	0.876	0.820	0.027	0.804	0.865	0.018	0.861
ρ_r	beta	0.500	0.200	0.098	0.039	0.111	0.192	0.068	0.231	0.142	0.061	0.139
ρ_g	beta	0.500	0.200	0.949	0.017	0.939	0.972	0.010	0.968	0.961	0.014	0.967
ρ_z	beta	0.500	0.200	0.985	0.002	0.985	0.968	0.009	0.965	0.985	0.004	0.984
ρ_u	beta	0.500	0.200	0.836	0.022	0.845	0.499	0.141	0.486	0.895	0.033	0.886
ρ_p	beta	0.500	0.200	0.167	0.059	0.160	0.808	0.127	0.882	0.896	0.070	0.906
ρ_w	beta	0.500	0.200	0.990	0.003	0.986	0.936	0.030	0.942	0.986	0.010	0.982
ρ_i	beta	0.500	0.200	0.651	0.038	0.637	0.822	0.053	0.844	0.783	0.069	0.819
σ_g	inv.gamma	0.100	2.000	0.467	0.023	0.469	0.496	0.025	0.495	0.457	0.025	0.441
σ_u	inv.gamma	0.100	2.000	0.574	0.070	0.586	1.088	0.339	0.972	0.334	0.056	0.336
σ_z	inv.gamma	0.100	2.000	0.437	0.027	0.467	0.395	0.025	0.381	0.406	0.031	0.409
σ_r	inv.gamma	0.100	2.000	0.197	0.010	0.200	0.223	0.012	0.223	0.203	0.012	0.201
σ_p	inv.gamma	0.100	2.000	0.143	0.010	0.135	0.119	0.012	0.110	0.179	0.017	0.188
σ_w	inv.gamma	0.100	2.000	0.340	0.016	0.338	0.258	0.021	0.274	0.355	0.019	0.334
σ_i	inv.gamma	0.100	2.000	0.387	0.030	0.386	0.365	0.033	0.350	0.362	0.039	0.322
μ_p	beta	0.500	0.200	0.140	0.077	0.077	0.646	0.129	0.706	0.976	0.034	0.980
μ_w	beta	0.500	0.200	0.968	0.005	0.966	0.851	0.064	0.850	0.971	0.011	0.972
ρ_{gz}	normal	0.500	0.250	1.316	0.089	1.299	1.394	0.100	1.386	1.360	0.109	1.516
$\bar{\gamma}$	normal	0.440	0.050	0.351	0.013	0.346	0.402	0.017	0.399	0.347	0.022	0.374
$\bar{l}$	normal	0.000	2.000	3.257	0.760	2.711	1.653	0.849	1.266	1.372	0.981	1.922
$\bar{\pi}$	gamma	0.625	0.100	0.936	0.097	0.986	0.973	0.084	0.979	0.661	0.094	0.603

Table 1: Estimation results for the samples: 1964–2019, 1964–2008 and again for 1964–2019 while ignoring the nonlinearities implied by the ELB.

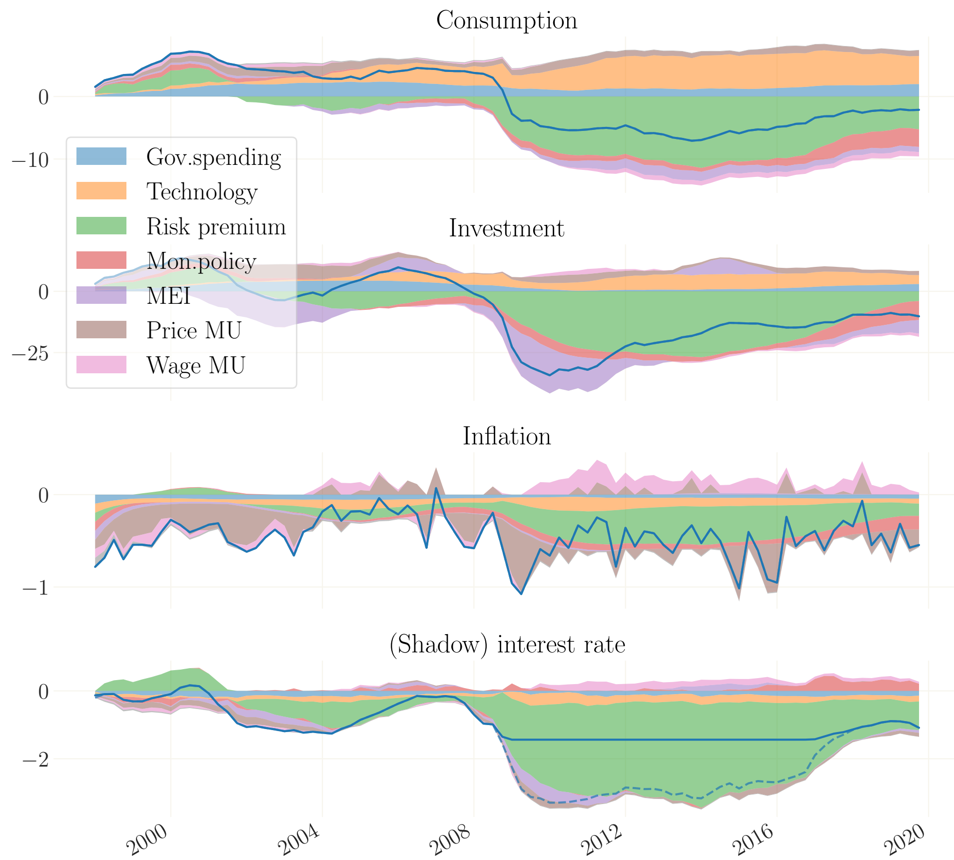

Figure 1: Historical Shock Decomposition of the Great Recession using the model estimated on the full sample from 1964–2019. Consumption and Investment: percentage deviations from their steady state growth path. Inflation and (shadow) interest rate: percentage points deviation from steady state. The decomposition in the bottom panel is made with respect to the shadow interest rate (dashed line), which corresponds to the notional interest rate r_t^n. *Note:* Means over 250 simulations drawn from the posterior. The contribution of each shock is normalized as in Appendix C.

premium shocks in the Great Recession. In turn, the persistence of shock to the marginal efficiency of investment, ρ_i and that of the price markup shock, ρ_p are estimated to be lower in the full sample than in the pre-crisis sample. Additionally, the inclusion of the Great Recession lowers the trend growth rate of the economy, γ.

3.4 The Great Recession Through the Lens of the model

We continue by briefly presenting the economic implications from the estimated model. These are largely in line with the results presented in Gust et al. (2017) and Kulish et al. (2017).[24]

Our estimation suggests that risk premiums shocks ϵ_t^u are the most prominent driver of the joint

[24]For a more in-depth discussion of the dynamics of the great financial crisis in the context of an estimated model with financial frictions, see Boehl and Strobel (2022).

dynamics of key variables following the financial crisis.[25] Figure 1 illustrates the dominant role of this shock for macroeconomic dynamics following the Great Recession.[26] It presents the historical shock decompositions of key variables during the Great Recession based on estimates using the full sample. From 2009 on, persistently elevated risk premiums account for almost the entire drop of aggregate consumption, weigh on aggregate investment and inflation, and consequently are responsible for the long duration of the ELB spell for the nominal interest rate.

However, high risk premiums cannot fully account for the sharp drop in investment during the Great Recession. While recessionary risk premium shocks do trigger a simultaneous downturn of consumption and investment, they fail to match the drop differential of these components, creating the need for an extra driver to make up for the missing decline in investment. In the case at hand, the initial decline of investment is triggered by recessionary MEI shocks, ϵ_t^i, which at the trough account for roughly half of the collapse in investment. Similarly, the decline of inflation during the Great Recession can only partly be attributed to the increase in risk premiums. The estimated flat Phillips Curve prevents the decline in real activity from generating substantial deflation.[27] It requires price markup shocks, ϵ_t^p, to account for the high-frequency movements of inflation in the sample and account for the dip in inflation during the Great Recession.

3.5 Implied expected ELB durations

A vivid debate has spanned around the question of whether the long ELB duration was an active policy choice or the result from a prolonged and persistent slump after the GFC. In our estimates the long duration of the ELB is largely interpreted by our estimation as an endogenous response of the central bank to the deterioration of fundamentals via the Taylor rule, rather than to an active lower-for-longer policy.[28] Figure 2 shows the dynamics and the distribution of the expected duration of the ELB spell over the sample. Although we do not target, nor use any prior information on the actual expectations of market participants on the duration of the ELB, they

[25] This shock is most prominently featured in Smets and Wouters (2007), who compare the effects of the shock to those of disturbances to net worth of entrepreneurs in a model with financial frictions as in Bernanke et al. (1999). Christiano et al. (2015) label this shock *consumption wedge* contrasting it with the *financial wedge* that is captured by the MEI shocks in our analysis. Fisher (2015) offers a structural interpretation of the risk premium shock as a shock to the demand for safe and liquid assets. Each of these interpretations share the notion that the risk premium shock is a short cut for capturing some financial disturbances, which makes its prominent role in the Great Recession plausible.

[26] The dominant role of risk premium shocks is corroborated by the generalized forecast error variance decomposition. It accounts for roughly half of the variation of output and 60 percent of the variation of the notional rate.

[27] This modest inflation response triggered a debate on the *missing disinflation puzzle*. See, e.g. Christiano et al. (2015); Gilchrist et al. (2017) for a discussion in the context of structural models.

[28] In principle, our specification of the shadow rate allows us to interpret monetary policy shocks at the ELB as forward guidance shocks. However, in the absence of additional data input such as, e.g., term premiums, we find substantial uncertainty surrounding our estimate of the shadow rate. For this reason we abstain from any statement regarding the effects of such policy. For a discussion of the effects of unconventional monetary policy, see Boehl et al. (forthcoming).

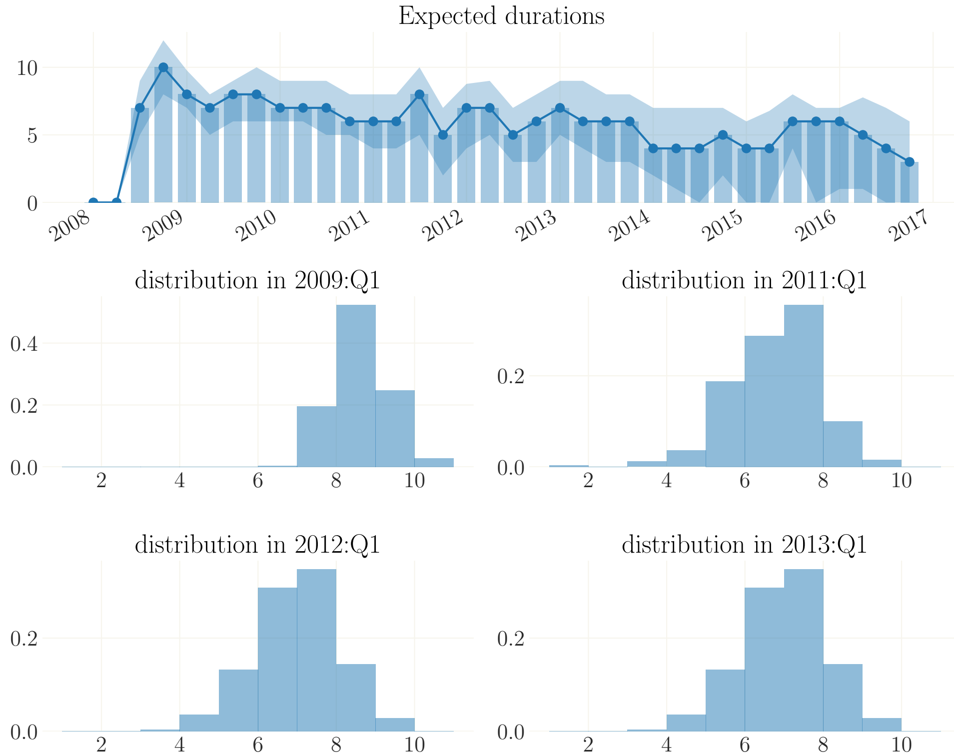

Figure 2: Estimated expected ELB durations based on the benchmark estimation. Bars in the top panel mark the mean estimate. The shaded area represents 90% credible sets reflecting parameter and filtering uncertainty. The lower panels show histograms of the distribution of ELB durations. The last bar to the right marks the probability of a duration of 10 or more quarters.

are broadly comparable to the average expected durations reported by the Blue Chip Financial Forecast and the Federal Reserve Bank of New York's Survey of Primary Dealers. The lower panels of Figure 2 show the distributions of expected ELB durations at different points in time. In 2009:Q1, most of the probability mass lies on a duration of 8 quarters, which is between the 75th and 90th percentile of the distribution implied by survey data. For 2011:Q1, where our mean expected duration of six quarters slightly exceeds the mean implied by the Primary Dealer Survey, our estimation allots a considerable probability mass to lower expected durations and the survey mean is within the credible set of the estimation. In the first quarters of 2012 and 2013, for which survey data shows expected durations of ten to eleven quarters, our estimates allots most of the probability mass to seven or six quarters, which still implies a substantial role of the ELB.

Whereas the Fed exited the ELB in 2015:Q4, our mean estimates of the expected durations remain positive until 2017:Q1. At the same time, the uncertainty surrounding our estimates increases strongly with the 90% credible set including values of k slightly above zero. The reason is that in the linear model, the output gap and the inflation rate are still far below the detrended balanced growth path, giving rise to very low interest rates via the monetary policy rule (see Figure 1).

Hence, expectations of the ELB duration are driven by the model-implied large and persistently negative output gap after the Great Recession and the low inflation rate.[29] Although agents observe the FFR to climb above the ELB, they interpret this as a contractionary monetary policy shock and expect the FFR to return to the ELB in the very near future.

The resulting estimated average expected durations are higher than those by Gust et al. (2017), who obtain an average ELB spell of merely 3.5 quarters. A potential reason for the difference in the resulting expected durations might be the treatment of the ELB in the estimation. As mentioned in Section 3.2, we set the empirical ELB to 0.05% quarterly, whereas Gust et al. (2017) choose exactly zero percent. This may be problematic as the Federal Funds Rate never actually went all the way down to zero. In theory, their model is hence capable of matching the observables without forcing the model to the zero lower bound.[30] Another difference in our estimation to that from Gust et al. (2017) is that we allow to estimate the persistence of the risk premium shock ρ_u wheras this parameter is fixed at a value of 0.85 in their estimation. Since our results suggest a prominent role of this shock – not least because its persistence is a major driver of the duration of the ELB spells – fixing the parameter ex-ante may bias the estimation results.[31]

3.6 The Merits of Using Post-Crisis Data in the Estimation

Accounting for the ELB in the estimation of a DSGE model is non-trivial (c.f. subsection 2). Thus, many model-driven analyses of the macroeconomic dynamics during the crisis are based on models that are either purely calibrated or estimated on pre-ELB data only (see, e.g., Christiano et al., 2015; Carlstrom et al., 2017). Others cover just a few ELB-quarters and also abstain from accounting for the ELB in their estimations (e.g., Chen et al., 2012; Christiano et al., 2014). Based on this, prominent results were generated that have shaped our understanding of the Great Recession, the role of financial frictions or the effects of unconventional monetary policy. In this subsection, we illustrate that omitting the ELB period can yield misleading implications: first, we estimate our model on a sample that ends in 2008Q4, thus ignoring the ELB-period. We then use these parameter estimates to confront the model with *post*-2008 data.

Figure 3 shows the historical shock decomposition of key variables in the Great Recession, but

[29] The finding of such a large, enduring output gap is is neither exclusive for the US data nor for estimated DSGE models. E.g., the OECD reports consistently negative output gaps for all years between the Global Financial Crisis and the Corona pandemic for most of its member states (OECD, 2021).

[30] From this angle it is surprising that in their smoothed state estimates, they hit the ELB at all. We suspect that this is due to the assumption of relatively large observation errors, which is often necessary when employing the particle filter (see e.g. Atkinson et al., 2020). Their measurement errors variances are assumed to at least 10% of the variance of data sample, which is a full magnitude higher than our assumed measurement errors, and even three magnitudes for the Federal Funds Rate.

[31] A potential reason for fixing this parameter is that for more persistent risk premium shocks, the global solution method employed by Gust et al. (2017) may not yield a unique solution. The piece-wise linear solution approach employed here does not confront this issue.

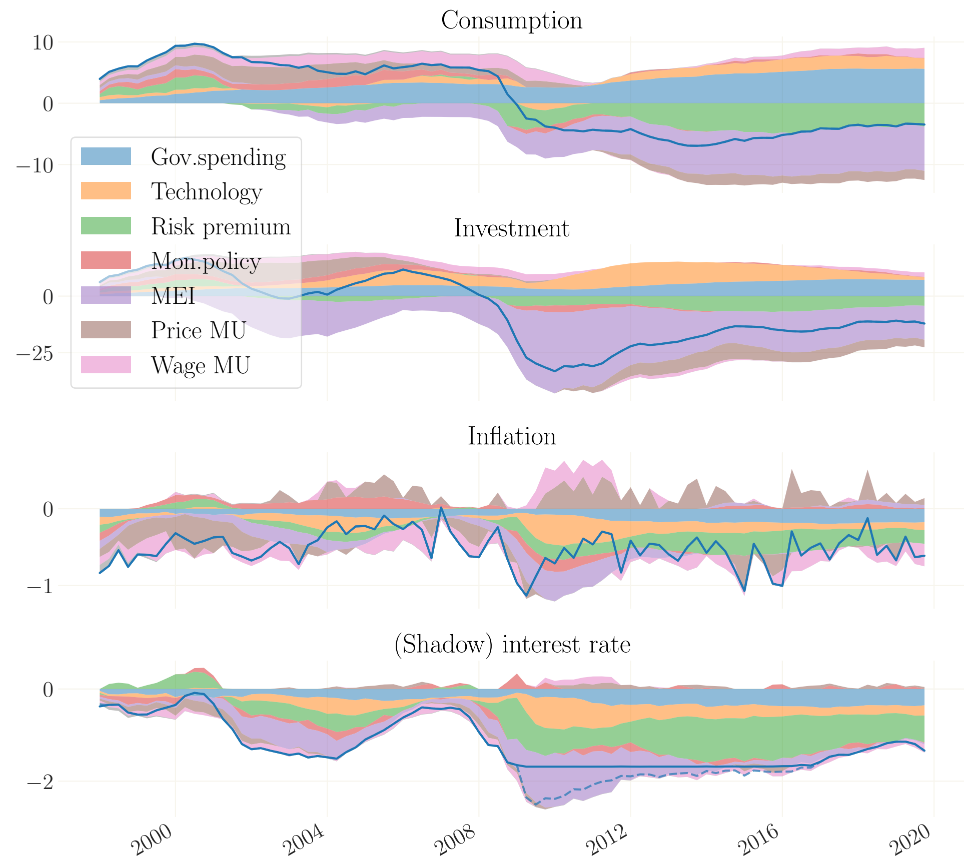

Figure 3: Historical Shock Decomposition of the Great Recession using the model estimated on the pre-crisis sample w/o ELB period from 1964–2008. The estimation results are then applied to the data including the post-crisis data. Consumption and Investment: percentage deviations from their steady state growth path. Inflation and (shadow) interest rate: percentage points deviation from steady state. The decomposition in the bottom panel is made with respect to the shadow interest rate (dashed line), which corresponds to the notional interest rate r_t^n. *Note:* Means over 250 simulations drawn from the posterior. The contribution of each shock is normalized as in Appendix C.

based on the model estimated on the pre-crisis sample , i.e. without the ELB period. Compared to the full sample, the importance of disturbances to the firms investment decision is highly overtaxed, thereby pointing to such disturbances as a major explanation for the Great Recession. Indeed, and likely consequentially, many studies focus in their explanation of the Great Recession on frictions that affect firms' investment financing.[32]

The different interpretation can be traced back to the difference in the estimates of the persistence parameters of risk premium shocks and MEI shocks. Figure 4 illustrates that in the full sample, the effects of risk premium shocks are far more persistent. Additionally, the figure shows that the fall of investment relative to the decline in consumption in the face of this shock is far less pronounced when the model is estimated on the pre-crisis sample. This is largely due to the dif-

[32]See, e.g. Gertler and Karadi (2011); Carlstrom et al. (2017) or Christiano et al. (2014).

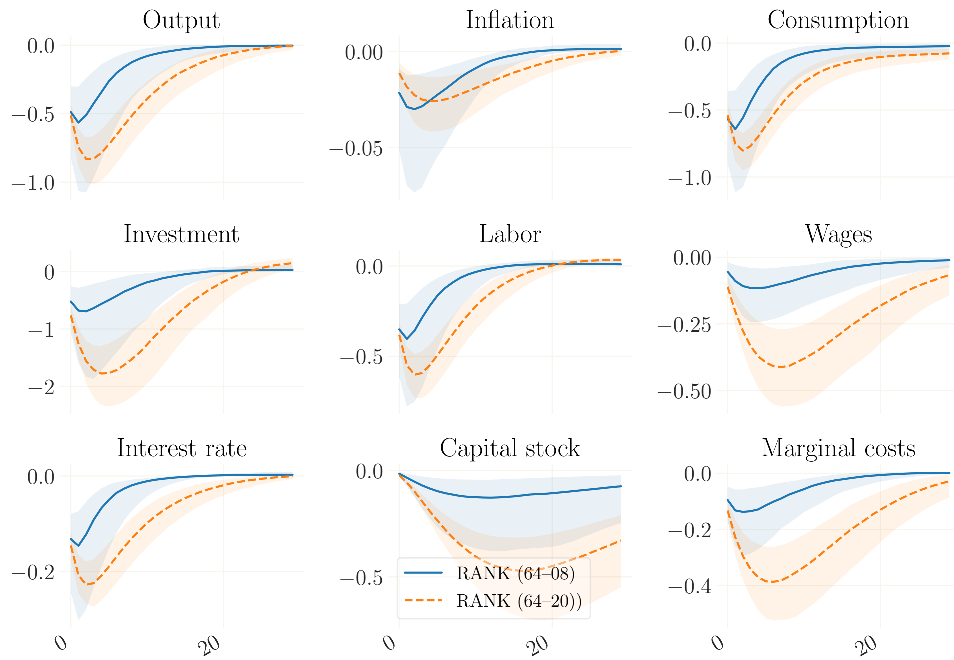

Figure 4: Impulse responses to a risk premium shock of one standard deviation.
Note: Medians over 250 simulations drawn from the posterior. The shaded area depicts the 90% credible set. The shock size equals the posterior mean standard deviation of the shock.

ference in the estimates of the coefficient of relative risk aversion, σ_c. In the full sample estimate, its posterior mean is close to unity while the pre-crisis mean estimate lies at 1.5. The larger value of σ_c means that the decline in labor hours in the Great Recession pulls consumption down further through the non-separabilities in the utility function. In turn, the additional drag on consumption implies that for a given decline of output that is caused by a risk premium shock, the decline of investment is reduced. This makes it less likely that risk premium shocks can account for the Great Recession, which was characterized by a collapse in investment.

In contrast, Figure 5 shows that MEI shocks become more attractive when post-2008 data is omitted from the estimation. In the model estimated on the full sample, a negative MEI shock initially increases consumption: by lowering aggregate demand, MEI shocks weigh on the policy interest rate, which in turn stimulates consumption on impact. This negative co-movement of consumption and investment is at odds with the observed dynamics in the Great Recession. In the pre-crisis sample, however, both consumption and investment decline with a negative MEI shock. Again, this can be traced back to the difference in the estimate of σ_c. In the pre-crisis sample, the higher value of σ_c strengthens the non-separabilities between labor and consumption. This implies that the decline in labor induces a drop in consumption as well. Notably, the pre-crisis estimate of σ_c is very close to the prior mean and it is hard to reject that this estimate is a matter of poor identification. On the contrary, the full sample estimate of this parameter is almost two

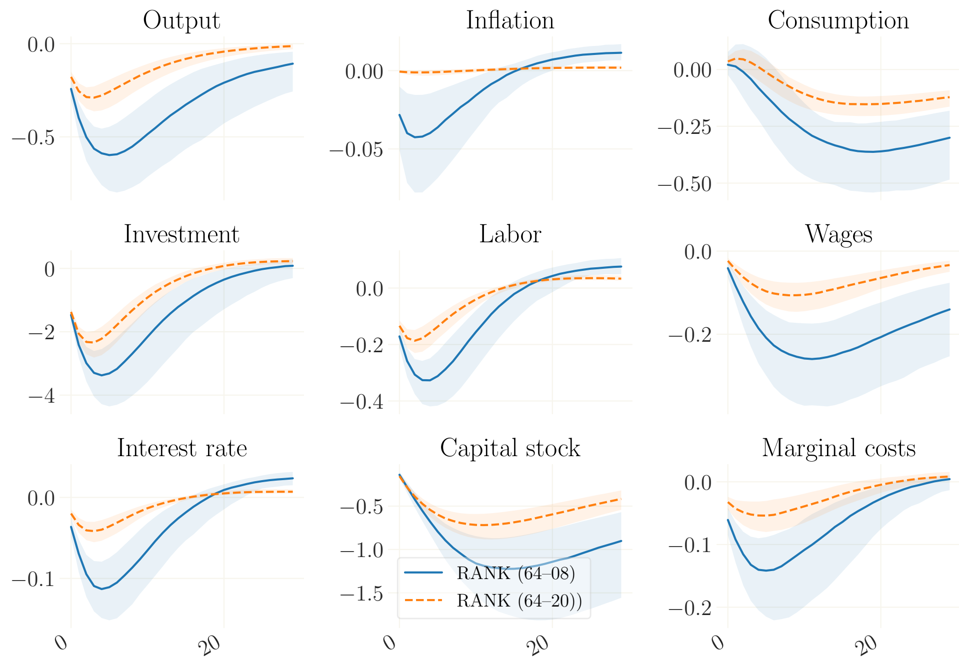

Figure 5: Impulse responses to a MEI shock of one standard deviation.
Note: Medians over 250 simulations drawn from the posterior. The shaded area depicts the 90% credible set. The shock size equals the posterior mean standard deviation of the shock.

standard deviations distant from the prior mean, which suggests that the value is driven by the data. Hence, through the lens of our pre-crisis estimates, MEI shocks – and other financial wedge type of shocks which share similar properties – appear more attractive than they are when including post-2008 data in the estimation.

In summary, the account of the Great Recession offered by our exercise based on the pre-crisis sample differs sharply from the interpretation based on the full sample. Here, elevated risk premiums play a dominant role for business cycles. Apart from the question, which modeling choices prove to be the best fit to capture the events of the recent decade, the exercise in this section highlights the importance of making use of post-2008 data, when analyzing macroeconomic dynamics during this time.

3.7 Ignoring the ELB

To circumvent the technical challenges posed by the ELB in the estimation, one can also simply ignore the ELB and estimate the model linearly on the full sample.[33] The right-hand columns of Table 1 show the parameter estimates resulting from this approach. A comparison with our baseline parameter estimates shows that ignoring the occasionally binding constraint introduces additional

[33] Fratto and Uhlig (2020) take this approach.

imprecision. For a number of parameters, the mean estimates are statistically different from the estimates that result from the estimates accounting for the ELB. The linear estimation implies among others, a lower risk-aversion (σ_c) lower habit formation (h), lower investment adjustment cost (S''), a higher frequency of price adjustment (ζ_p), and a lower responsiveness of monetary policy to movements in inflation. Among the exogenous processes, the most notable difference is a change in the nature of the price markup shocks, which are far more persistent (ρ_p) in the linear setting and feature a lower standard deviation (σ_p).

This result strongly advises against ignoring the ELB. Ignoring the non-linearity implied by an occasionally binding constraint will introduce a mistake into the analysis, regardless of the application. Since the severity of the mistake for the application of interest is ex-ante unknown the researcher needs to estimate the model including the ELB. Our approach offers a way forward.[34]

4 Estimation Accuracy

In this section, we test the estimation performance and accuracy of our approach on artificial data. To ease comparison, we closely follow Atkinson et al. (2020, henceforth ART). The authors compare the estimation performance of a fully nonlinear solution combined with the particle filter, and the piece-wise solution method of Guerrieri and Iacoviello (2015) in conjunction with the inversion filter (IVF) of Cuba-Borda et al. (2019). Their results are obtained for a small-scale DSGE model with a relatively small number of parameters and only two endogenous states (interest rate inertia and consumption). They conclude that the advantages of the fully nonlinear solution (including agents that take aggregate uncertainty into account) are small and outweighed by the benefits of using much faster methods such as OccBin with the IVF, which enables the researcher to estimating richer and hence less mis-specified models.

As in ART, we simulate a large set of artificial datasets to test our set of tools. Other than ART, we use the medium-scale model introduced in Section 3.1 as the data generating process (DGP) and set the parameters of the DGP to the posterior mean from the estimation in the previous section (cf. Table 1). Also different than ART, we abstract from the effects of model misspecification: the estimated model and the DGP are the same model. Each dataset spans over 240 quarters, of which we omit the first 120 quarters. We then take the first 50 datasets in which the ELB is not binding at all, and the first 50 sets in which the ELB is successively binding for exactly 30

[34] Another way to avoid the ELB could be to use "shadow rates" (e.g. Krippner (2013) or Wu and Xia (2016)) as a replacement for the FFR and to proceed with a linear estimation. However, shadow rates are sensitive to the strong assumptions required in their construction. Accordingly, prominent shadow rates have substantially different paths. An important drawback of using shadow rates in the context of Bayesian analysis is that it implies hard-wiring the assumed effects of unconventional monetary policy into the analysis. Our results in Boehl et al. (forthcoming) suggest that the use of shadow rates may overestimate the effects of unconventional monetary policy on consumption.

quarters.[35] As documented by the first columns in Table 2 and in line with ART we also set the prior mean to the true parameter values to eliminate potential biases that are orthogonal to the filtering methodology.[36] The standard deviations of the prior distribution are exactly as before, which reflect the original priors of SW. Note that for the ensemble-MCMC sampling procedure this also implies that we initialize the estimation around the true parameter values, which implies that any deviation in the parameter estimates must come from filtering bias.

To measure the accuracy of parameter estimates we use normalized root-mean squared errors (NRMSE) as in ART. For parameter j, the error is the difference between the posterior mean estimate for dataset k, $\hat{\theta}_{j,k}$ and the true parameter θ_j. For the number of datasets N the measure is given by

$$\mathrm{NRMSE}_j = \frac{1}{\theta_j}\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\hat{\theta}_{j,k}-\theta_j\right)^2}, \tag{26}$$

which normalizes the standard root-mean square error by the true parameter θ_j to remove scale differences.

Table 2 presents the results of our accuracy check. Overall, we find the means of the simulations to be closely aligned with the true parameter values. This suggests that the EnKF indeed approximates the true likelihood very well. The results do not indicate any severe bias in either direction. As discussed in Section 2, the EnKF is an exact Bayesian filter for linear models and replicated the exact results of the linear Kalman filter. We can use this fact to determine parameters that are likely to be generally badly identified from the estimations where the model is actually linear. Examples are ρ_r, ρ_p or μ_p, which display slightly elevated NRMSE. It turns out that exactly these parameters are not very well identified in the nonlinear estimation, while all others display NRMSEs in a similar range as for the linear estimation.

We also go one step ahead and benchmark the EnKF against the filter of Cuba-Borda et al. (2019) in 5. We repeat the exact same setup as above and use the same datasets. As argued in the introduction, the IVF has two potential shortcomings: it ignores uncertainty about the initial states, and the inverse of the transition function may either not exist, or not be unique. Regarding the first problem we find that the IVF still delivers acceptable parameter estimates for the datasets in which the ELB is *not* binding, however with a considerably larger dispersion in mean estimates (NRMSEs are about 30% larger than with the EnKF). This suggests that ignoring the uncertainty about the initial states does indeed cause a loss on estimation accuracy.

However, The more severe problem seems to be the non-uniqueness of the transition function

[35] While datasets in which the ELB is not binding occur quite frequently, sets in which the ELB is binding for exactly 30 periods are quite rare events. To obtain 50 of these datasets, we need a total number of almost one million draws.

[36] Sole exceptions are the priors of ρ_z, ρ_g, ρ_w and μ_w, which we set to 0.9 since sampling from a Beta distribution with a mean close to one poses difficulties.

	Prior			No ELB				ELB binding for 30 periods			
	type	mean	std	mean	NRMSE	HDP: 5%	95%	mean	NRMSE	HDP: 5%	95%
σ_c	normal	1.156	0.375	1.219	0.618	1.086	1.321	1.207	0.602	1.046	1.322
σ_l	normal	3.333	0.750	3.366	0.357	3.067	3.585	3.252	0.589	2.712	3.591
β_{tpr}	gamma	0.147	0.100	0.154	1.295	0.107	0.184	0.140	1.390	0.088	0.179
h	beta	0.635	0.100	0.608	0.405	0.573	0.651	0.628	0.332	0.573	0.660
S''	normal	5.140	1.500	5.147	0.555	4.533	5.815	5.574	0.918	4.829	6.505
ι_p	beta	0.657	0.150	0.674	0.516	0.596	0.746	0.708	0.750	0.635	0.777
ι_w	beta	0.528	0.150	0.532	0.481	0.477	0.578	0.521	0.825	0.413	0.612
α	normal	0.173	0.050	0.162	0.712	0.140	0.181	0.156	0.864	0.141	0.179
ζ_p	beta	0.904	0.100	0.896	0.178	0.861	0.928	0.905	0.120	0.878	0.929
ζ_w	beta	0.817	0.100	0.817	0.241	0.776	0.863	0.811	0.239	0.768	0.862
Φ_p	normal	1.440	0.125	1.468	0.275	1.397	1.548	1.472	0.260	1.376	1.520
ψ	beta	0.502	0.150	0.498	0.689	0.406	0.566	0.475	0.875	0.404	0.595
ϕ_π	normal	2.190	0.250	2.225	0.235	2.115	2.323	2.352	0.696	2.121	2.564
ϕ_y	normal	0.173	0.050	0.177	0.785	0.151	0.205	0.202	1.649	0.157	0.244
ϕ_{dy}	normal	0.254	0.050	0.261	0.540	0.235	0.294	0.235	0.898	0.187	0.269
ρ	beta	0.870	0.100	0.866	0.173	0.835	0.902	0.863	0.208	0.824	0.898
ρ_r	beta	0.098	0.200	0.102	4.534	0.028	0.180	0.090	4.252	0.029	0.172
ρ_g	beta	0.900	0.200	0.936	0.340	0.896	0.969	0.930	0.321	0.891	0.963
ρ_z	beta	0.900	0.200	0.983	0.657	0.972	0.995	0.980	0.642	0.967	0.995
ρ_u	beta	0.836	0.200	0.836	0.294	0.775	0.890	0.874	0.380	0.840	0.922
ρ_p	beta	0.167	0.200	0.143	2.014	0.080	0.211	0.146	1.881	0.084	0.197
ρ_w	beta	0.900	0.200	0.952	0.465	0.894	0.980	0.956	0.505	0.930	0.989
ρ_i	beta	0.651	0.200	0.654	0.655	0.583	0.771	0.665	0.561	0.579	0.731
μ_p	beta	0.140	0.200	0.104	2.434	0.054	0.154	0.120	2.128	0.068	0.174
μ_w	beta	0.900	0.200	0.946	0.403	0.912	0.975	0.934	0.367	0.907	0.968
ρ_{gz}	normal	1.316	0.250	1.313	0.443	1.214	1.473	1.303	0.531	1.197	1.459
σ_g	IG	0.467	0.250	0.454	0.428	0.415	0.489	0.458	0.457	0.414	0.506
σ_u	IG	0.574	0.250	0.564	0.876	0.448	0.691	0.536	0.876	0.453	0.642
σ_z	IG	0.437	0.250	0.385	1.010	0.343	0.450	0.360	1.391	0.284	0.398
σ_r	IG	0.197	0.250	0.195	0.552	0.173	0.220	0.186	0.697	0.166	0.208
σ_p	IG	0.143	0.250	0.139	0.713	0.110	0.156	0.139	0.787	0.116	0.162
σ_w	IG	0.340	0.250	0.347	0.533	0.303	0.383	0.337	0.553	0.282	0.374
σ_i	IG	0.387	0.250	0.388	0.721	0.327	0.437	0.375	0.744	0.317	0.434
$\bar{\gamma}$	normal	0.351	0.050	0.350	0.357	0.327	0.384	0.352	0.418	0.311	0.378
$\bar{l}$	normal	3.257	2.000	3.157	1.283	2.150	4.019	3.645	1.638	2.410	4.542
$\bar{\pi}$	gamma	0.936	0.100	0.953	0.243	0.909	1.001	0.904	0.359	0.844	0.956

Table 2: Estimation results for our set of methods across 50 artificial datasets in which the ELB is not binding at all (center columns) and binding for 30 subsequent periods (right columns).

once we allow for a binding ELB. We document that for the datasets in which the ELB is binding for 30 subsequent periods, the estimate of the likelihood is very noisy. We conclude that this renders sampling from the posterior distribution hardly possible for our medium scale model. Note that, given the size of the state space of the model, it is cumbersome to also benchmark against the particle filter with a fully nonlinear solution. The potential disadvantages of the particle filter are, however, already documented in ART and Cuba-Borda et al. (2019).

5 The Ensemble Kalman Filter and the Inversion Filter

A natural benchmark for the EnKF is the inversion filter (IVF, henceforth), which was first suggested in Guerrieri and Iacoviello (2017) for the estimation of a model with occasionally binding constraints (OBCs). The filter was initially proposed by Fair and Taylor (1980) as a simple device for likelihood inference of nonlinear models. Two recent papers (Cuba-Borda et al., 2019; Atkinson et al., 2020) discuss its performance for models with the ELB. The filter is implemented in the most recent version of Dynare (Dynare 5.0).

For convenience we here repeat equations 5 and 6 from the main body of the text, where we denote a nonlinear hidden Markov-Model (HMM) by

$$x_t = g(x_{t-1}, \varepsilon_t), \tag{27}$$

$$z_t = h(x_t) + \nu_t, \tag{28}$$

with exogenous economic innovations $\varepsilon_t \sim \mathcal{N}(0, Q)$ and measurement errors $\nu_t \sim N(0, R)$. Given x_{t-1} and in the absence of measurement errors, (27) and (28) imply a direct mapping $f_{\text{IVF}} : \varepsilon_t \rightarrow z_t$ with $f_{\text{IVF}} = h \circ g$. Invertibility of f_{IVF} implies a mapping $f_{\text{IVF}}^{-1} : z_t \rightarrow \varepsilon_t$ from observables to shocks. In other words, if the initial state x_0 is known, the "hidden"-property of the HMM becomes irrelevant. Proposition 1 gives a formal statement of the filter.

Proposition 1. *Iff*

a) x_0 *is known with certainty,*

b) f_{IVF} *is invertible (and thereby,* f_{IVF}^{-1} *is unique), and*

c) *there is no measurement error (*$R = 0_{z\times z}$*),*

then for any time series data $\{z_t\}_{t=1}^{T}$ *the mapping* f_{IVF}^{-1} *can be used to find a series of shocks* $\{\varepsilon_t\}_{t=1}^{T}$ *that perfectly explains the data. The unbiased likelihood of the model is then given by*

$$\log\left(p(y_{1:T})\right) = -\frac{Tn_z}{2}\log(2\pi) - \frac{T}{2}\log(\det(Q) - \frac{1}{2}\sum_{t=1}^{T}\varepsilon_t' Q^{-1}\varepsilon_t + \sum_{t=1}^{T}\log\left(\left|\det\frac{\partial\varepsilon_t}{\partial z_t}\right|\right). \tag{29}$$

Proof. See Appendix A.2.1 in Guerrieri and Iacoviello (2017). While no formal proof is provided, this claim is easy to verify. ■

For the linearized model with the occasionally binding ELB, there exists no known closed form expression for f_{IVF} (and, hence, not for its inverse). As in Guerrieri and Iacoviello (2017); Cuba-Borda et al. (2019); Atkinson et al. (2020) we instead use a standard root finding algorithm to find a shock ε_t that satisfies f_{IVF} for a given z_t.[37] Additionally, as suggested by Guerrieri and Iacoviello (2017) we set $\epsilon_r = 0$ whenever the observed FFR approaches zero to avoid underdeterminancy.[38] Lastly, we can find the determinant of the Jacobian of ε_t wrt. z_t by acknowledging that $\log\left(\left|\det\frac{\partial\varepsilon_t}{\partial z_t}\right|\right) = -\log\left(\left|\det\frac{\partial z_t}{\partial\varepsilon_t}\right|\right)$ where $\frac{\partial z_t}{\partial\varepsilon_t}$ is a direct byproduct of evaluating f.

In practical applications, x_0 (that is, the initial state) is unobservable. This clearly violates the necessary conditions in Proposition 1 and hence biases the estimate of the likelihood. That may or may not be a serious problem in the context of a Bayesian estimation. The applications considered in Cuba-Borda et al. (2019); Atkinson et al. (2020) all feature small scale models with only few endogenous steady states. In these models, the bias is likely to be rather limited.

We test the extent of this bias in the standard medium-scale model. For this purpose, we use the set of artificial data in which the ELB is *not* binding. Since trivially, the inverse (almost always) exists for a linear transition function, using data in which the ELB is not binding circumvents the second problem of the IVF, which is that f_{IVF} may not be invertible. This helps us to single out the effects that only stem from ignoring uncertainty about the initial state at $t = 0$.

In the first exercise, we use the same prior as in Section 4 and evaluate the likelihood around

[37] We use the "hybr" method implemented in Pythons Scipy library, which uses MINPAKS hybrd and hybrj routines. These are established and well tested routines used as a backend for many high level languages. As a practical matter, we let the log-likelihood be $-\infty$ if the root finding algorithm does not converge.

[38] Note that this is a limitation of the filter – a Bayesian filter can determine ε_t even if $n_\varepsilon > n_z$.

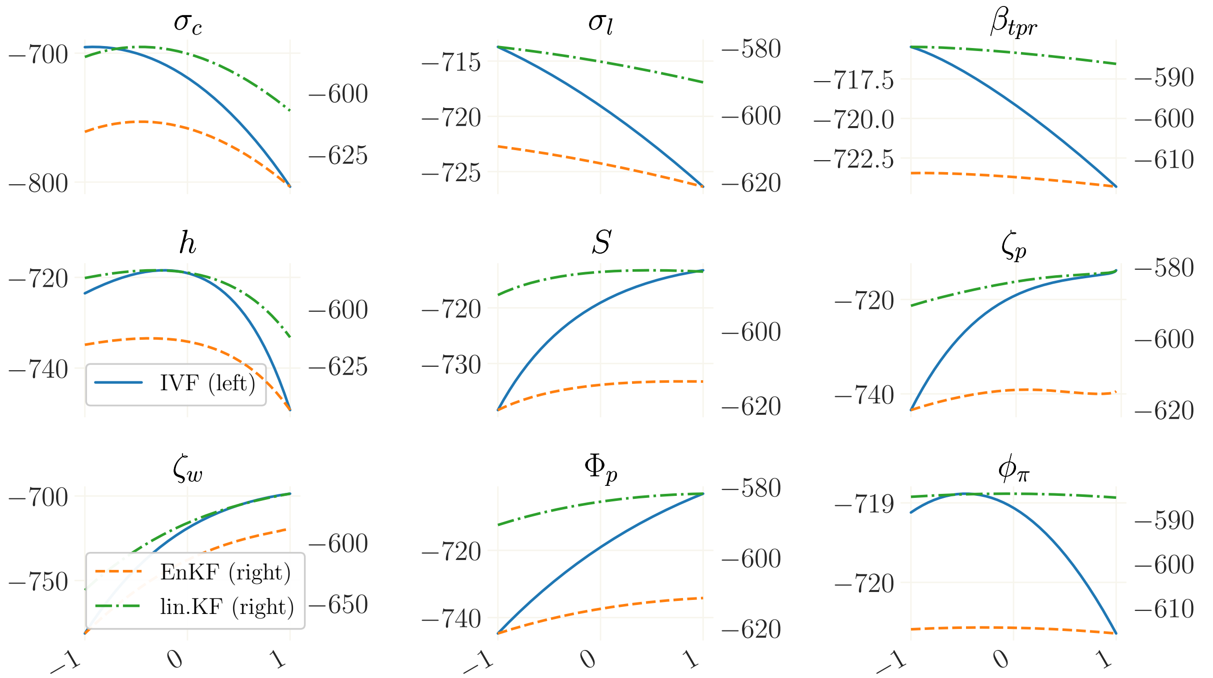

Figure 6: Likelihood evaluations for an artificial dataset without binding ELB. For each panel, all parameters are set to the prior mean while one parameter is varied within one standard deviation of its prior distribution (x-axis). Left y-axis: likelihood evaluations with the IVF; right y-axis: likelihood evaluations with the EnKF and the KF.

the prior mean for the IVF, the KF and the EnKF in one of the artificial datasets in which the ELB is *not* binding (note that the prior means are the true parameters of the DGP). The result is shown in Figure 6. In each of the panels, we vary exactly one parameter within the range of one standard deviation of its prior distribution (from -1 to 1), while leaving all others at the prior mean (zero, at the x-axis). Overall, there is a considerable difference in scale between IVF (left-axis) and EnKF (right axis). At the same time, apart from one exception (ζ_p) the EnKF matches the KF (also right y-axis) up to a constant. Still, IVF and (En)KF suggest similar positions of the maximum of the (marginal) likelihood function. A notable exception is σ_c, where the IVF suggests a lower mode than KF and EnKF.

Secondly, we repeat the exercise from Section 4 using artificial data in which the ELB is *not* binding. Table E.3 in Appendix E shows the resulting parameter estimates. Similar as with the EnKF, the means over all 50 simulations are relatively close to the true parameters of the DGP (which are also the prior mean). This suggests that the IVF is not systematically biased in any direction. However, a comparison of normalized root-mean squared errors (NRMSEs) indicates that ignoring uncertainty about the initial states does indeed have impact on estimation accuracy. NRMSEs for the IVF are on average 30.2% larger, with extreme cases such as β_{tpr} and $\bar{l}$ in which they are more than three times larger.

Our second concern regards the invertibility of f_{IVF}. As argued above, the mapping is clearly (almost always) invertible if it is linear. However, it is hard to argue that for any z_t, there is a *unique* ε_t that satisfies f_{IVF} for a given x_{t-1}. The reason is that given ε_t, there are potentially multiple sets of spell durations that form a valid equilibrium (see especially Holden (2017) but also, e.g., Carlstrom et al. (2015)). Hence, if g is possibly not unique this implies that f_{IVF} is not unique, in turn suggesting that there is no unique mapping $z_t \rightarrow \varepsilon_t$. An additional point is that f_{IVF} may even not exist or the root finding algorithm may simply not converge. We find that these issues are very relevant in practice. Note that a Bayesian filter works in the opposite direction than the IVF: shocks are drawn according to their distribution and then passed through the transition function. A Bayesian filter selects shocks that are more likely given their covariance, uncertainty about previous states, and measurement noise. In contrast, the IVF will accept any ε_t that satisfies f_{IVF}, independent of how likely it is. If there are several spell durations that form an equilibrium, ε_t may crucially depend on the initial guess for the spell duration, the initial guess for the root finding procedure, or both.

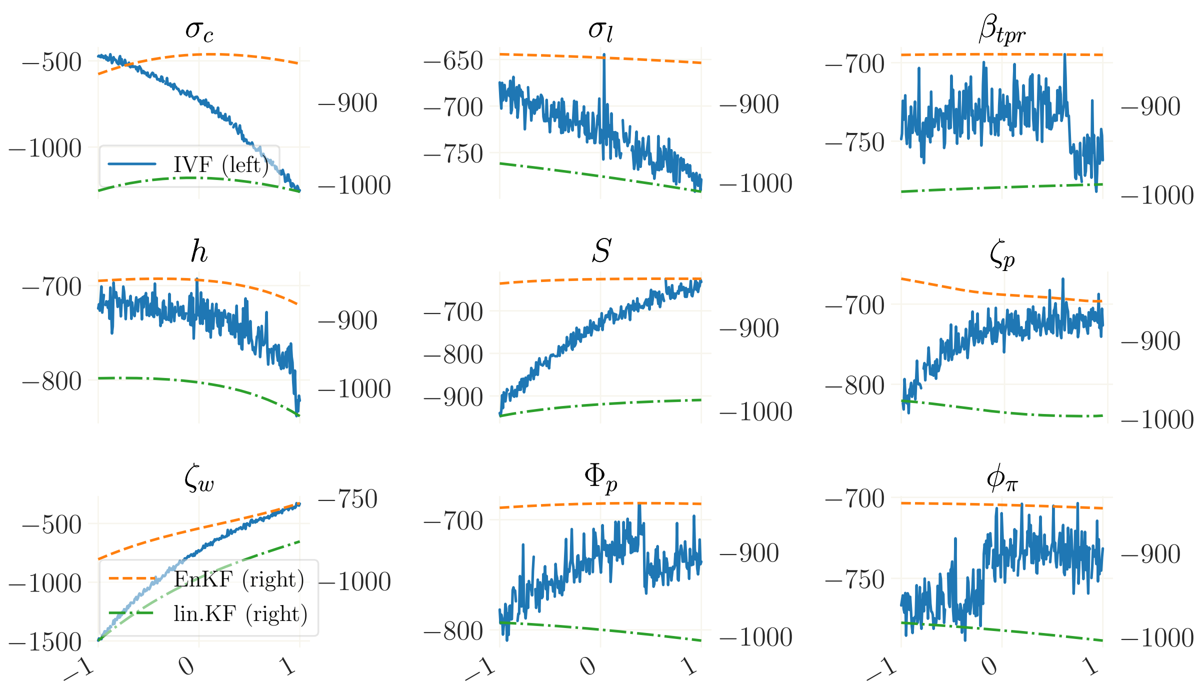

Figure 7: Likelihood evaluations given for an artificial dataset where the ELB binds for 30 subsequent quarters. For each panel, all parameters are set to the prior mean while one parameter is varied within one standard deviation of its prior distribution (x-axis). Left y-axis: likelihood evaluations with the IVF; right y-axis: likelihood evaluations with the EnKF and the KF.

We find that when replicating the exercise from Section 4 for the IVF with the data in which the ELB binds for 30 subsequent periods, the acceptance rate soon drops down to 1% and below.

Consequently, we were unable to obtain a reliable posterior sample since the sampler does not move away from the initial ensemble. When examining the problem, we noted a very high dispersion of the likelihood, even if we initialized all chains very close to the true parameter values. This indeed suggests that the estimate of the likelihood is quite noisy. Figure 7 repeats the exercise from Figure 6 but with an artificial dataset in which the ELB binds for 30 subsequent periods. Since the transition function is now (at times) nonlinear, the estimates from EnKF and KF are not equal. The noisiness of the likelihood estimate for the IVF varies across parameters, but is clearly large enough to make proper sampling from the posterior distribution impossible. Note, that the selection of ε_t is a crucial difference to the EnKF the KF (and, for that matter, also to the particle filter): the Bayesian filters will propose those shocks that are likely given the state at period t. Respectively, the filter will ex-post reject those shock vectors ε_t which are very unlikely.

6 Conclusion

This paper proposes a novel approach for the efficient and robust Bayesian estimation of medium- and large-scale DSGE models with occasionally binding constraints. It combines a novel nonlinear recursive filter with a piece-wise linear solution method for models with OBCs and a state-of-the art MCMC sampler that allows for an easy in-parallel sampling from high dimensional posterior distributions. Our discussion of the novel methods is accompanied by an accessible reference implementation: the Pydsge package. We validate our methods on artificial data in which the ELB is binding for a prolonged time. Our toolkit can easily be extended to the estimation of larger models with OBCs, as e.g. in Boehl et al. (forthcoming).

A further advantage of the methods presented here is that they enable researchers to estimate models with occasionally binding constraint even in the absence of reliable data on the expected duration of the binding constraint. We illustrate this along the example of the Great Recession in the US and the long-binding ELB on nominal interest rates. Our approach to endogenize the ELB durations generates similar parameter estimates and historical shock decompositions as previous papers that use external survey data on expectations of the ELB durations. This lends additional credence to our methods.

We find that through the lens of the canonical medium-scale model, post-2008 dynamics are dominated by elevated risk premiums on household borrowing rates, in line with the importance of increased mortgage rates in the financial crisis. In contrast, we find that using pre-crisis-only estimates to analyze the post-2008 period yields the conclusion that shocks to the cost of investment were a main driver for the Great Recession and the US economy's post-2008 trajectory. This difference in results is a cautionary tale that should discourage from empirically investigating on the Great Recession with models tuned to match the pre-2008 experience.

References

Aruoba, S. Borağan, Pablo Cuba-Borda, and Frank Schorfheide, "Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries," *The Review of Economic Studies*, 2018, *85* (1), 87–118.

_ , _ , Kenji Higa-Flores, Frank Schorfheide, and Sergio Villalvazo, "Piecewise-linear approximations and filtering for DSGE models with occasionally-binding constraints," *Review of Economic Dynamics*, 2021, *41*, 96–120. Special Issue in Memory of Alejandro Justiniano.

Atkinson, Tyler, Alexander W Richter, and Nathaniel A Throckmorton, "The zero lower bound and estimation accuracy," *Journal of Monetary Economics*, 2020, *115*, 249–264.

Bernanke, Ben S, Mark Gertler, and Simon Gilchrist, "The financial accelerator in a quantitative business cycle framework," *Handbook of Macroeconomics*, 1999, *1*, 1341–1393.

Binning, Andrew and Junior Maih, "Sigma point filters for dynamic nonlinear regime switching models," Technical Report 2015.

Blanchard, Olivier Jean and Charles M Kahn, "The solution of linear difference models under rational expectations," *Econometrica: Journal of the Econometric Society*, 1980, pp. 1305–1311.

Boehl, Gregor, "Efficient solution and computation of models with occasionally binding constraints," *Journal of Economic Dynamics and Control*, 2022, *143*, 104523.

_ , "An Ensemble MCMC Sampler for Robust Bayesian Inference," *Available at SSRN 4250395*, 2022.

_ and Felix Strobel, "The Empirical Performance of Financial Frictions Since 2008," Discussion Paper Series CRC TR 224, University of Bonn and University of Mannheim, Germany 2022.

_ , Gavin Goy, and Felix Strobel, "A structural investigation of quantitative easing," *Review of Economics and Statistics*, forthcoming.

Cai, Michael, Marco Del Negro, Marc P. Giannoni, Abhi Gupta, Pearl Li, and Erica Moszkowski, "DSGE forecasts of the lost recovery," *International Journal of Forecasting*, 2019, *35* (4), 1770–1789.

Calvo, Guillermo A, "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 1983, *12* (3), 383–398.

Canova, Fabio, Filippo Ferroni, and Christian Matthes, "Detecting and analyzing the effects of time-varying parameters in DSGE models," *International Economic Review*, 2020, *61* (1), 105–125.

Carlstrom, Charles T, Timothy S Fuerst, and Matthias Paustian, "Inflation and output in New Keynesian models with a transient interest rate peg," *Journal of Monetary Economics*, 2015, *76*, 230–243.

Carlstrom, Charles T., Timothy S. Fuerst, and Matthias Paustian, "Targeting Long Rates in a Model with Segmented Markets," *American Economic Journal: Macroeconomics*, January 2017, *9* (1), 205–42.

Chen, Han, Vasco Cúrdia, and Andrea Ferrero, "The Macroeconomic Effects of Large-scale Asset Purchase Programmes," *The Economic Journal*, 2012, *122* (564), F289–F315.

Christiano, Lawrence J., Martin S. Eichenbaum, and Mathias Trabandt, "Understanding the Great Recession," *American Economic Journal: Macroeconomics*, January 2015, *7* (1), 110–67.

_ , Roberto Motto, and Massimo Rostagno, "Risk Shocks," *American Economic Review*, January 2014, *104* (1), 27–65.

Cozzi, Guido, Beatrice Pataracchia, Philipp Pfeiffer, and Marco Ratto, "How much Keynes and how much Schumpeter?," *European Economic Review*, 2021, *133*, 103660.

Cuba-Borda, Pablo, Luca Guerrieri, Matteo Iacoviello, and Molin Zhong, "Likelihood evaluation of models with occasionally binding constraints," *Journal of Applied Econometrics*, 2019, *34* (7), 1073–1085.

Del Negro, Marco and Frank Schorfheide, "DSGE Model-Based Forecasting," in G. Elliott, C. Granger, and A. Timmermann, eds., *Handbook of Economic Forecasting*, Vol. 2 of Handbook of Economic Forecasting, Elsevier, 2013, chapter 0, pp. 57–140.

Evensen, Geir, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research: Oceans*, 1994, *99* (C5), 10143–10162.

_ , *Data assimilation: the ensemble Kalman filter*, Vol. 2, Springer, 2009.

Fair, Ray C and John B Taylor, "Solution and Maximum Likelihood Estimation of Dynamic Nonlinear RationalExpectations Models," Technical Report, National Bureau of Economic Research 1980.

Fisher, Jonas D.M., "On the Structural Interpretation of the Smets–Wouters "Risk Premium" Shock," *Journal of Money, Credit and Banking*, 2015, *47* (2-3), 511–516.

Foreman-Mackey, Daniel, David W Hogg, Dustin Lang, and Jonathan Goodman, "EMCEE: the MCMC hammer," *Publications of the Astronomical Society of the Pacific*, 2013, *125* (925), 306.

Fratto, Chiara and Harald Uhlig, "Accounting for Post-Crisis Inflation: A Retro Analysis," *Review of Economic Dynamics*, January 2020, *35*, 133–153.

Frei, Marco and Hans R Künsch, "Sequential state and observation noise covariance estimation using combined ensemble Kalman and particle filters," *Monthly Weather Review*, 2012, *140* (5), 1476–1495.

Gertler, Mark and Peter Karadi, "A model of unconventional monetary policy," *Journal of Monetary Economics*, 2011, *58*, 17–34.

Gilchrist, Simon, Raphael Schoenle, Jae Sim, and Egon Zakrajšek, "Inflation Dynamics during the Financial Crisis," *American Economic Review*, March 2017, *107* (3), 785–823.

Goodman, Jonathan and Jonathan Weare, "Ensemble samplers with affine invariance," *Communications in applied mathematics and computational science*, 2010, *5* (1), 65–80.

Guerrieri, Luca and Matteo Iacoviello, "OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily," *Journal of Monetary Economics*, 2015, *70*, 22–38.

_ and _ , "Collateral constraints and macroeconomic asymmetries," *Journal of Monetary Economics*, 2017, *90* (C), 28–49.

Gust, Christopher, Edward Herbst, David López-Salido, and Matthew E Smith, "The empirical implications of the interest-rate lower bound," *American Economic Review*, 2017, *107* (7), 1971–2006.

Herbst, Edward and Frank Schorfheide, "Tempered particle filtering," *Journal of Econometrics*, 2019, *210* (1), 26–44.

Herbst, Edward P and Frank Schorfheide, *Bayesian estimation of DSGE models*, Princeton University Press, 2016.

Holden, Tom D, "Existence and uniqueness of solutions to dynamic models with occasionally binding constraints," Technical Report 2017.

Jones, Callum, Mariano Kulish, and Daniel M Rees, *International spillovers of forward guidance shocks*, International Monetary Fund, 2018.

Julier, Simon J and Jeffrey K Uhlmann, "New extension of the Kalman filter to nonlinear systems," in "Signal processing, sensor fusion, and target recognition VI," Vol. 3068 International Society for Optics and Photonics 1997, pp. 182–193.

Julier, Simon, Jeffrey Uhlmann, and Hugh F Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on automatic control*, 2000, *45* (3), 477–482.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti, "Investment Shocks and the Relative Price of Investment," *Review of Economic Dynamics*, January 2011, *14* (1), 101–121.

Katzfuss, Matthias, Jonathan R Stroud, and Christopher K Wikle, "Understanding the ensemble Kalman filter," *The American Statistician*, 2016, *70* (4), 350–357.

Keen, Benjamin D, Alexander W Richter, and Nathaniel A Throckmorton, "Forward Guidance and the State of the Economy," *Economic Inquiry*, 2017, *55* (4), 1593–1624.

Kehoe, Patrick J, Pierlauro Lopez, Virgiliu Midrigan, and Elena Pastorino, "Credit Frictions in the Great Recession," Working Paper 28201, National Bureau of Economic Research December 2020.

Kimball, Miles S., "The Quantitative Analytics of the Basic Neomonetarist Model," NBER Working Papers 5046, National Bureau of Economic Research, Inc February 1995.

Kollmann, Robert, Beatrice Pataracchia, Rafal Raciborski, Marco Ratto, Werner Roeger, and Lukas Vogel, "The post-crisis slump in the Euro Area and the US: Evidence from an estimated three-region DSGE model," *European Economic Review*, 2016, *88* (C), 21–41.

Krippner, Leo, "Measuring the stance of monetary policy in zero lower bound environments," *Economics Letters*, 2013, *118* (1), 135–138.

Kulish, Mariano, James Morley, and Tim Robinson, "Estimating DSGE models with zero interest rate policy," *Journal of Monetary Economics*, 2017, *88*, 35 – 49.

McElhoe, B.A., "An assessment of the navigation and course corrections for a manned flyby of mars or venus," *IEEE Transactions on Aerospace and Electronic Systems*, 1966, *AES-2*.

McKay, Michael D, Richard J Beckman, and William J Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, 2000, *42* (1), 55–61.

Mian, Atif and Amir Sufi, "What Explains the 2007–2009 Drop in Employment?," *Econometrica*, 2014, *82* (6), 2197–2223.

__ **and** __ , *House of Debt* number 9780226271651. In 'University of Chicago Press Economics Books.', University of Chicago Press, 2015.

Niederreiter, Harald, "Low-discrepancy and low-dispersion sequences," *Journal of number theory*, 1988, *30* (1), 51–70.

OECD, "Statistical Appendix to the OECD Economic Outlook, December 2021," 2021.

Plante, Michael, Alexander W Richter, and Nathaniel A Throckmorton, "The zero lower bound and endogenous uncertainty," *The Economic Journal*, 2018, *128* (611), 1730–1757.

__ **, Alexander W. Richter, and Nathaniel A. Throckmorton**, "The Zero Lower Bound and Endogenous Uncertainty," *The Economic Journal*, 2018, *128* (611), 1730–1757.

Raanes, Patrick Nima, "On the ensemble Rauch-Tung-Striebel smoother and its equivalence to the ensemble Kalman smoother," *Quarterly Journal of the Royal Meteorological Society*, 2016, *142* (696), 1259–1264.

Rauch, Herbert E, CT Striebel, and F Tung, "Maximum likelihood estimates of linear dynamic systems," *AIAA journal*, 1965, *3* (8), 1445–1450.

Richter, Alexander W and Nathaniel A Throckmorton, "Is Rotemberg pricing justified by macro data?," *Economics Letters*, 2016, *149*, 44–48.

Smets, Frank and Raf Wouters, "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American Economic Review*, 2007, *97* (3), 586–606.

Smith, G.L., S.F. Schmidt, and L.A. McGee, "Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle," Technical Report, National Aeronautics and Space Administration 1962.

Stroud, Jonathan R and Thomas Bengtsson, "Sequential state and variance estimation within the ensemble Kalman filter," *Monthly weather review*, 2007, *135* (9), 3194–3208.

ter Braak, Cajo JF, "A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces," *Statistics and Computing*, 2006, *16* (3), 239–249.

Ungarala, Sridhar, "On the iterated forms of Kalman filters using statistical linearization," *Journal of Process Control*, 2012, *22* (5), 935–943.

Wu, Jing Cynthia and Fan Dora Xia, "Measuring the macroeconomic impact of monetary policy at the zero lower bound," *Journal of Money, Credit and Banking*, 2016, *48* (2-3), 253–291.

Appendix (For Online-Publication)

Appendix A Data

Our measurement equations contain the following variables:

- GDP: ln(GDP/GDPDEF/CNP16OV)*100
- CONS: ln((PCEC)/GDPDEF/CNP16OV)*100
- INV: ln((FPI)/GDPDEF/CNP16OV)*100
- LAB: ln((AWHNONAG*CE16OV)/CNP16OV)*100
- INFL: ln(GDPDEF)
- WAGE: ln(COMPNFB/GDPDEF)*100
- FFR: FEDFUNDS/4

For GDP, CONS, INV, INFL and WAGE we use the log changes in our measurement equations. We demean LAB in our measurement equation.

Data sources:

- GDP: Gross Domestic Product, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate, FRED
- GDPDEF: Gross Domestic Product: Implicit Price Deflator, Index 2012=100, Quarterly, Seasonally Adjusted, FRED
- PCEC: Personal Consumption Expenditures, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate, FRED
- FPI: Fixed Private Investment, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate, FRED
- AWHNONAG: Average Weekly Hours of Production and Nonsupervisory Employees: Total private, Hours, Quarterly, Seasonally Adjusted, FRED.
- CE16OV: Civilian Employment Level, Thousands of Persons, Seasonally Adjusted, FRED.
- CNP16OV: trailing MA(5) of the Civilian Noninstitutional Population, Thousands of Persons, Quarterly, Not Seasonally Adjusted, FRED.

- COMPNFB, Nonfarm Business Sector: Compensation Per Hour, Index 2012=100, Quarterly, Seasonally Adjusted, FRED
- FEDFUNDS: Effective Federal Funds Rate, Percent, FRED.

Appendix B Model Descriptions

We adopt the framework by Smets and Wouters (2007) as a baseline model to interpret the Great Recession. Following Del Negro and Schorfheide (2013), we detrend all nonstationary variables by $Z_t = e^{\gamma t + \frac{1}{1-\alpha}\widetilde{z}_t}$, where, γ is the steady-state growth rate of the economy and α is the output share of capital. $\widetilde{z}_t$ is the linearly detrended log productivity process that follows the autoregressive law of motion $\widetilde{z}_t = \rho_z \widetilde{z}_{t-1} + \sigma_z \epsilon_z$. For z_t, the growth rate of technology in deviations from γ, it holds that $z_t = \frac{1}{1-\alpha}(\rho_z - 1)\widetilde{z}_t + \frac{1}{1-\alpha}\sigma_z \epsilon_z$.

Labor is differentiated by unions with monopoly power that face nominal rigidities for their wage setting process. Intermediate good producers employ labor and capital services and sell their goods to final goods firms. Final good firms are monopolistically competitive and face nominal rigidities as in . The model further allows for exogenous government spending and features a monetary authority that sets the short-term nominal interest rate according to a monetary policy rule.

This subsection briefly presents the linearized equilibrium conditions. A detailed derivation of the linearized equations is discussed e.g. in the appendix to Smets and Wouters (2007). All variables in this section are expressed as a log-deviation from their respective steady state values. The consumption Euler equation of the households is given by

$$\begin{aligned} c_t = \frac{h/\gamma}{(1+h/\gamma)}(c_{t-1} - z_t) + \frac{1}{1+h/\gamma}E_t[c_{t+1} + z_{t+1}] + \frac{(\sigma_c - 1)(W^h L/C)}{\sigma_c(1+h/\gamma)}(l_t - E_t[l_{t+1}]) \\ - \frac{(1-h/\gamma)}{(1+h/\gamma)\sigma_c}(r_t - E_t[\pi_{t+1}] + u_t), \end{aligned} \tag{B.1}$$

where c_t is consumption, and l_t is their supply of labor. Parameters h, σ_c and σ_l are, respectively, the degree of external habit formation in consumption, the coefficient of relative risk aversion, and the inverse of the Frisch elasticity. γ denotes the steady-state growth rate of the economy. r_t is the nominal interest rate, π_t is the inflation rate, and u_t is an exogenous risk premium shock, which drives a wedge between the lending/savings rate and the riskless real rate.

Equation (B.2) is the linearized relationship between investment and the relative price of capital,

$$i_t = \frac{1}{1+\overline{\beta}}[(i_{t-1} - z_t] + \frac{\overline{\beta}}{1+\overline{\beta}}E_t[i_{t+1} + z_{t+1}] + \frac{1}{(1+\overline{\beta})\gamma^2 S''}q_t + v_{i,t}. \tag{B.2}$$

Here, i_t denotes investment in physical capital and q_t is the price of capital. It holds that $\overline{\beta} = \beta\gamma^{(1-\sigma_c)}$ where β is the households' discount factor. Investment is subject to adjustment costs, which are governed by S'', the steady-state value of the second derivative of the investment adjustment cost function, and an exogenous process, $v_{i,t}$. While Smets and Wouters (2007) interpret $e_{i,t}$ as an investment specific technology disturbance, Justiniano et al. (2011) stress that this shock can as well be viewed as a reduced-form way of capturing financial frictions, as it drives a wedge between aggregate savings and aggregate investment. We henceforth refer to this disturbance as a shock on the marginal efficiency of investment (MEI).

The accumulation equation of physical capital is given by

$$\overline{k}_t = (1-\delta)/\gamma(\overline{k}_{t-1} - z_t) + (1-(1-\delta)/\gamma)i_t + (1-(1-\delta)/\gamma)(1+\overline{\beta})\gamma^2 S'' v_{i,t}, \tag{B.3}$$

where $\overline{k}$ denotes physical capital, and parameter δ is the depreciation rate. The following Equation (B.4) is the no-arbitrage condition between the rental rate of capital, r_t^k, and the riskless real rate:

$$r_t - E_t[\pi_{t+1}] + u_t = \frac{r^k}{r^k + (1-\delta)} E_t[r_{t+1}^k] + \frac{(1-\delta)}{r^k + (1-\delta)} E_t[q_{t+1}] - q_t. \tag{B.4}$$

As the use of physical capital in production is subject to utilization costs, which in turn can be expressed as a function of the rental rate on capital, the relation between the effectively used amount of capital k_t and the physical capital stock is

$$k_t = \frac{1-\psi}{\psi} r_t^k + \overline{k}_{t-1}, \tag{B.5}$$

where $\psi \in (0,1)$ is the parameter governing the costs of capital utilization. Equation (B.6) is the aggregate production function

$$y_t = \Phi(\alpha k_t + (1-\alpha)l_t + z_t) + (\Phi - 1)\frac{1}{1-\alpha}\widetilde{z}_t. \tag{B.6}$$

Intermediate good firms employ labor and capital services. Let z_t be the exogenous process of total factor productivity. Parameter α is the elasticity of output with respect to capital and Φ enters the production function due to the assumption of a fixed cost in production. Real marginal costs for producing firms, mc_t, can be written as

$$mc_t = w_t - z_t + \alpha(l_t - k_t). \tag{B.7}$$

w_t denotes the real wage, which are set by labor unions. Furthermore, cost minimization for intermediate good producers results in condition (B.8):

$$k_t = w_t - r_t^k + l_t. \tag{B.8}$$

The aggregate resource constraint (B.9) contains an exogenous demand shifter, g_t, which comprises exogenous variations in government spending and net exports, as well as the resource costs of capital utilization:

$$y_t = \frac{G}{Y}g_t + \frac{C}{Y}c_t + \frac{I}{Y}i_t + \frac{R^k K}{Y}\frac{1-\psi}{\psi}r_t^k + \frac{1}{1-\alpha}\widetilde{z}_t.. \tag{B.9}$$

Final good producers are assumed to have monopoly power and face nominal rigidities as in Calvo (1983) when setting their prices. This gives rise to a New Keynesian Phillips Curve (NKPC) of the form

$$\pi_t = \frac{\overline{\beta}}{1+\iota_p\overline{\beta}}E_t\pi_{t+1} + \frac{\iota_p}{1+\iota_p\overline{\beta}}\pi_{t-1} + \frac{(1-\zeta_p\overline{\beta})(1-\zeta_p)}{(1+\overline{\beta}\iota_p)\zeta_p((\Phi-1)\epsilon_p+1)}mc_t + v_{p,t}. \tag{B.10}$$

Here, ζ_p is the probability that a firm cannot update its price in any given period. In addition to Calvo pricing, we assume partial price indexation, governed by the parameter ι_p. The Phillips Curve is hence both, forward and backward looking. ϵ_p denotes the curvature of the Kimball (1995) aggregator for final goods. Due to the Kimball aggregator, the sensitivity of inflation to fluctuations in marginal cost is affected by the market power of firms, represented by the steady state price markup, $\Phi - 1$.[39] Furthermore, the curvature of the Kimball aggregator affects the adjustment of prices to marginal cost as the higher ϵ_p, the higher is the degree of strategic complementarity in price setting, dampening the price adjustment to shocks. The last term in the NKPC, $v_{p,t}$, represents exogenous fluctuations in the price markup.

While final good producers set prices on the good market, wages are set by labor unions. Unions bundle labor services from households and offer them to firms with a markup over the frictionless wage, w_t^h, which reads

$$w_t^h = \frac{1}{(1-h)}(c_t - h/\gamma c_{t-1} + h/\gamma z_t.) + \sigma_l l_t. \tag{B.11}$$

As with price setting, we assume that the nominal rigidities in the wage setting process are of the

[39]Note that in equilibrium, the steady state price markup is tied to the fixed cost parameter by a zero profit condition.

Calvo type, and include partial wage indexation. The wage Phillips curve thus is

$$w_t = \frac{1}{1+\bar{\beta}\gamma}(w_{t-1} - z_t + \imath_w \pi_{t-1}) + \frac{\bar{\beta}\gamma}{1+\bar{\beta}\gamma}E_t[w_{t+1} + z_{t+1} + \pi_{t+1}] - \frac{1+\imath_w\bar{\beta}\gamma}{1+\bar{\beta}\gamma}\pi_t \\ + \frac{(1-\zeta_w\bar{\beta}\gamma)(1-\zeta_w)}{(1+\bar{\beta}\gamma)\zeta_w((\lambda_w - 1)\epsilon_w + 1)}(w_t^h - w_t) + v_{w,t}. \tag{B.12}$$

The term $w_t^h - w_t$ is the inverse of the wage markup. Analogous to equation (B.10), the terms λ_w and ϵ_w are the steady state wage markup and the curvature of the Kimball aggregator for labor services, respectively. The term $v_{w,t}$ represents exogenous variations in the wage markup.

We take into account the fact that the central bank is constrained in its interest rate policy by a zero lower bound (ELB) on the nominal interest rate. Therefore, in the linear model, it is that

$$r_t = \max\{\bar{r}, r_t^n\} \tag{B.13}$$

with $\bar{r}$ being the lower bound value. Whenever the policy rate is away from the constraint, it corresponds to the notional rate, r_t^n, which follows the feedback rule

$$r_t^n = \rho r_{t-1}^n + (1-\rho)\left(\phi_\pi \pi_t + \phi_y \widetilde{y}_t\right) + \phi_{dy}\Delta\widetilde{y}_t + v_{r,t}. \tag{B.14}$$

Here, $\widetilde{y}_t$ is the output gap and $\Delta\widetilde{y}_t = \widetilde{y}_t - \widetilde{y}_{t-1}$ its growth rate. Parameter ρ expresses an interest rate smoothing motive by the central bank. ϕ_π, ϕ_y and ϕ_{dy} are feedback coefficients. When the economy is away from the ELB, the stochastic process $v_{r,t}$ represents a regular interest rate shock. When the nominal interest rate is zero, however, $v_{r,t}$ may not directly affect the level of the nominal interest rate. However, through the persistence of the stochastic process that drives $v_{r,t}$, it affects the expected path of the notional rate and can therefore alter the expected duration of the lower bound spell. It can hence be viewed as a forward guidance shock whenever the economy is at the ELB.

Finally, the stochastic drivers in our model are the following seven processes:

$$u_t = \rho_u u_{t-1} + \epsilon_t^u, \tag{B.15}$$

$$z_t = \rho_z z_{t-1} + \epsilon_t^z, \tag{B.16}$$

$$g_t = \rho_g g_{t-1} + \epsilon_t^g + \rho_{gz}\epsilon_t^z, \tag{B.17}$$

$$v_{r,t} = \rho_r v_{r,t-1} + \epsilon_t^r, \tag{B.18}$$

$$v_{i,t} = \rho_i v_{i,t-1} + \epsilon_t^i, \tag{B.19}$$

$$v_{p,t} = \rho_p v_{p,t-1} + \epsilon_t^p - \mu_p \epsilon_{t-1}^p, \tag{B.20}$$

$$v_{w,t} = \rho_w v_{w,t-1} + \epsilon_t^w - \mu_w \epsilon_{t-1}^w, \tag{B.21}$$

$$\tag{B.22}$$

where $\epsilon_t^k \overset{iid}{\sim} N(0, \sigma_k^2)$ for all $k = \{r, i, p, w\}$, and likewise for $\{u_t, z_t, g_t\}$.

Appendix C Normalization of historic shock decompositions for models with OBCs

We are interested in quantifying the contribution of a each type of shock to the sequence of model variables. Such quantification is called the historic shock decomposition (HSD). If the model feature one or several occasionally binding constraints (OBCs), the model is nonlinear and the HSD is generally not unique. To illustrate this, imagine a deflationary MEI shock ε_t^i and a risk premium shock u_t, which together cause the ELB to bind. Assume that each, the MEI shock and the risk premium shock alone are insufficiently strong to force the ELB to hold. Then, the effect of u_t conditional on the realization of ε_t^i will have a different dynamic effect than just u_t taken alone, and it is unclear which value to assign to u_t within the context of a HSD. This appendix offers a way to quantify the historic shock contributions in models with OBCs.

More precisely, we are interested in the sequence of vectors

$$\{\mathbf{h}_{t,z}\}_0^T \tag{C.1}$$

where $z \in \{1, 2, \cdots, n_z\}$ is in the set of all n_z types of shocks and where each $\mathbf{h}_{t,z}$ is the cumulative dynamic contribution of type-z shocks to time-t model variables y_t. $\mathbf{h}_{t,z}$ is hence recursive. By definition, $\varepsilon_t = (\varepsilon_t^1, \varepsilon_t^2, \cdots, \varepsilon_t^{n_z})$ is the vector of all n_z shocks in the model at time t. We require for each period t that

$$\sum_{z=1}^{n_z} \mathbf{h}_{t,z} = y_t \tag{C.2}$$

and at least that

$$\{\mathbf{h}_{t,z} = \mathbf{0} \wedge \mathbf{h}_{t-1,z} = \mathbf{0} \iff \varepsilon_t^z = 0\} \; \forall z = 1, 2, \cdots, n_z \tag{C.3}$$

i.e. that any zero shock has a zero net contribution to the HSD. Further, we require the HSD to be unique and the attributions to each shock to be proportional.

We propose a normalization method for historic shock decomposition that is specific to models with OBCs. Importantly, the normalization is such that the result is independent of any ordering effects. For convenience, let us repeat Equation (2) from the main body:

$$F_s(l,k,s_{t-1}) = N^{\max\{s-l,0\}}\hat{N}^{\min\{l,s\}}\begin{bmatrix} f\,(l,k,s_{t-1}) \\ s_{t-1}\end{bmatrix} + (I-N)^{-1}\left(I - N^{\max\{s-l,0\}}\right)b\bar{r}, \tag{C.4}$$

$$= E_t\begin{bmatrix} c_{t+s} \\ s_{t+s-1}\end{bmatrix}, \tag{C.5}$$

where $\hat{N} = (I-bp)^{-1}(N+bm)$ and

$$f(l,k,s_{t-1}) = \left\{c_t : \Psi N^k \hat{N}\begin{bmatrix} c_t \\ s_{t-1}\end{bmatrix} = -\Psi(I-N)^{-1}(I-N^k)b\bar{r}\right\}. \tag{C.6}$$

Define latent states net of shocks as $\tilde{s}_{t-1}$ and remember that the state vector consist of latent states and current shocks, $w_{t-1} = (\tilde{s}_{t-1}, \varepsilon_t)^\intercal$. Take as given the time sequence of smoothed shocks $\{\varepsilon_t\}_0^T$ that fully reproduces $\{y_t\}_0^T$. This implies that we also have obtained the sequence of all $\{l,k\}$. The law-of-motion from period t to $t+1$ is then given by $F_1(l,k,s_{t-1})$, i.e. $F_s(\cdot)$ for $s=1$. From Equation (C.6), $f(l,k,s_{t-1})$ can be decomposed in a coefficient matrix $\bar{f}_w(l,k)$ that is to be pre-multiplied to s_{t-1}, and a constant vector $\bar{f}_c(k)$ that only depends on k. To ease notation, define both such that s_{t-1} is returned:

$$\begin{bmatrix} f\,(l,k,s_{t-1}) \\ s_{t-1}\end{bmatrix} = \bar{f}_w(l,k)s_{t-1} + \bar{f}_c(l,k), \tag{C.7}$$

That means the bottom part of $\bar{f}_w(l,k)$ is a (n_y+n_z) dimensional identity matrix and the bottom part of $\bar{f}_c(k)$ is a $(n_y+n_z)\times 1$ zero vector.

From this we can rewrite $F_1(\cdot)$ as

$$E_t(c_{t+1}, s_t)^\intercal = \tag{C.8}$$

$$\begin{aligned} F_1(l,k,s_{t-1}) = {} & N^{\max\{1-l,0\}}\hat{N}^{\min\{l,1\}}\left(\bar{f}_w(l,k)\begin{bmatrix}\tilde{s}_{t-1} \\ \varepsilon_t\end{bmatrix} + \bar{f}_c(k)\right) \\ & +(I-N)^{-1}\left(I-N^{\max\{s-l,0\}}\right)b\bar{r}, \end{aligned} \tag{C.9}$$

where we are more explicit about the shocks.

Denote by I_z the n_z-dimensional identity matrix and by $I_z^{n_z}$ its z-th column. For each z we define

$\mathbf{h}_{t,z}$ by the recursion

$$
\begin{aligned}
(\mathfrak{x}_{t+1,z}, \mathbf{h}_{t,z})^{\intercal} = & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(C.10)} \\
F_1(l, k, \mathbf{h}_{t-1,z}, \varepsilon_t^z) = & N^{\max\{1-l,0\}} \hat{N}^{\min\{l,1\}} \bar{f}_w(l,k) \begin{bmatrix} \mathbf{h}_{t-1,z} \\ I_z^{n_z} \varepsilon_t^z \end{bmatrix} \\
& + \omega_{t,z} N^{\max\{1-l,0\}} \hat{N}^{\min\{l,1\}} \bar{f}_c(k) \qquad\qquad\qquad\qquad\qquad \text{(C.11)} \\
& + \omega_{t,z} (I - N)^{-1} \left(I - N^{\max\{s-l,0\}} \right) b\bar{r},
\end{aligned}
$$

where, from the linearity of the first two terms at the RHS, it is easy to show that Condition (C.2) is satisfied as long as $\sum_z^n \omega_{t,z} = 1 \ \forall t$. [40]

The first terms of the RHS of (C.11) is the recursion of $\mathbf{h}_{t,z}$, and also attributes the effects of the current shock to $\mathbf{h}_{t,z}$. For the two other terms the remaining task is to assign weights $\omega_{t,z}$ such that Condition (C.3) is satisfied.

Define

$$
\omega_{t,z} = \frac{bN^{\max\{1-l,0\}} \hat{N}^{\min\{l,1\}} \bar{f}_w(l,k) \begin{bmatrix} \mathbf{h}_{t-1,z} \\ I_z^{n_z} \varepsilon_t^z \end{bmatrix}}{bN^{\max\{1-l,0\}} \hat{N}^{\min\{l,1\}} \bar{f}_w(l,k) \begin{bmatrix} \tilde{s}_{t-1} \\ \varepsilon_t \end{bmatrix}}, \tag{C.12}
$$

i.e. set $\omega_{t,z}$ proportional to the relative contribution of ε_t^z to the constraint value r_t.

Intuitively, this acknowledges that the values of $\{l, k\}$ depend on the magnitude of the scalar r_t relative to $\bar{r}$. The deeper below r_t is of $\bar{r}$, the longer the constraint will bind, and the higher is k (note that the constant term will be zero for any $l > 0$). If the contribution of ε_t^z to a negative r_t is large, then the respective weight $\omega_{t,z}$ of the constant terms in (C.11) attributed to ε_t^z will be high, and vice versa.

For our application with the ELB this means that the weight of constant terms for each shock is proportional to the contribution of the shock to the total level of the shadow rate. Further note that by (C.2)

$$
\sum_z^{n_z} \left\{ N^{\max\{1-l,0\}} \hat{N}^{\min\{l,1\}} \bar{f}_w(l,k) \begin{bmatrix} \mathbf{h}_{t-1,z} \\ I_z^{n_z} \varepsilon_t^z \end{bmatrix} \right\} = N^{\max\{1-l,0\}} \hat{N}^{\min\{l,1\}} \bar{f}_w(l,k) \begin{bmatrix} \tilde{s}_{t-1} \\ \varepsilon_t \end{bmatrix}, \tag{C.13}
$$

and hence $\sum_e \omega_{t,e} = 1$, i.e. the weights sum up to unity.

[40] Additionally, note that $\mathfrak{x}_{t+1,z}$ is the time-t decomposition of controls.

Finally, acknowledge that for $\mathbf{h}_{t-1,z} = 0$ and $\varepsilon_t^z = 0$ we have

$$
\begin{aligned}
F_1(l,k,0,0) =& \omega_{t,z} N^{\max\{1-l,0\}} \hat{N}^{\min\{l,1\}} \bar{f}_c(k) \\
&+ \omega_{t,z}(I-N)^{-1}\left(I - N^{\max\{s-l,0\}}\right) b\bar{r}, \\
=&(0,0),
\end{aligned}
\tag{C.14}
$$

which follows from the fact that $\omega_{t,z} = 0$ whenever $\mathbf{h}_{t-1,z}$ and ε_t^z both are zero. This shows that Condition (C.3) is also satisfied.

Appendix D The shape of the posterior distribution

The figures in this section show the 200 chains used for the estimation of the benchmark model. We have a total of 2500 samples, of which we keep the last 500. That means that the posterior contains $500 \times 200 = 100,000$ parameter draws. See Boehl (2022b) for details on the differential-independence mixture ensemble Monte Carlo Markov chain method (DIME MCMC) we use for posterior sampling. For each model, we run a total of 2500 iterations, of which we keep the last 500. That means that the posterior contains $500 \times 200 = 100,000$ parameter draws. We check for convergence using the method of integrated autocorrelation time with a window size of $c = 50$, as suggested by Goodman and Weare (2010). Note that it is not trivial to find a sufficient statistics for convergence since the samples in the chain are not independent. The figures strongly suggest that the estimation is converged from iteration 2000 onwards.

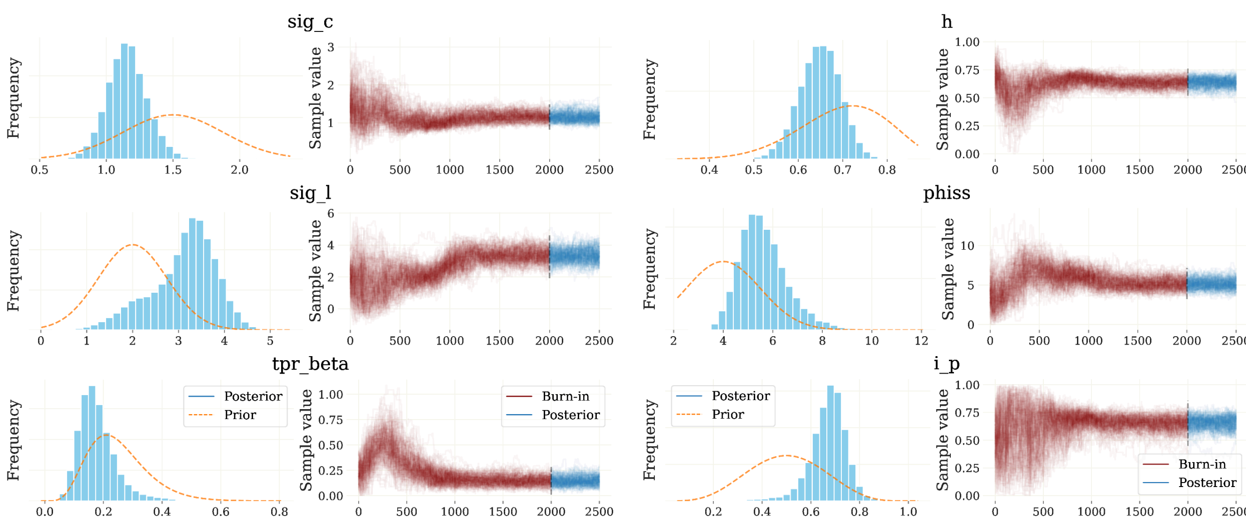

Figure D.8: Traceplots of the 200 DIME chains for selected parameters. Estimation of the benchmark model. The left panel shows a KDE of the parameter distribution. The right displays the trace of each of the chains over time.

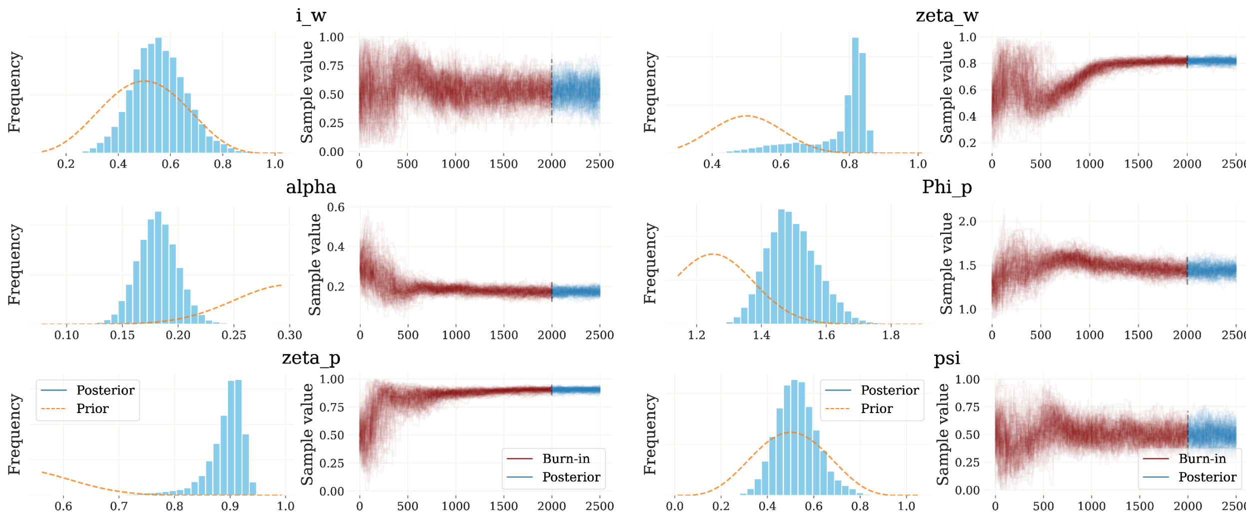

Figure D.9: Traceplots of the 200 DIME chains for selected parameters. Estimation of the benchmark model. The left panel shows a KDE of the parameter distribution. The right displays the trace of each of the chains over time.

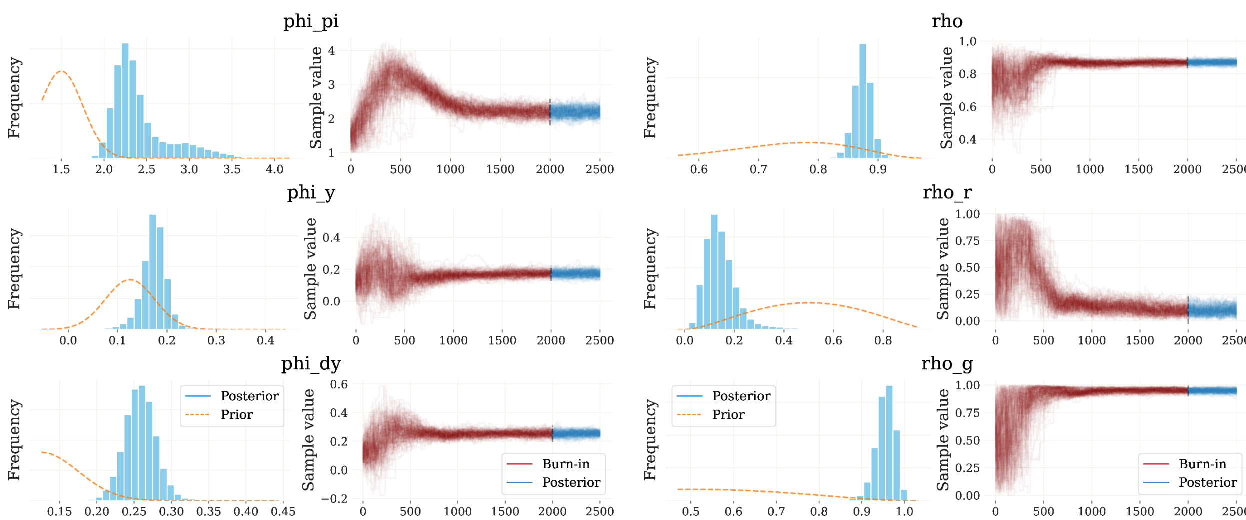

Figure D.10: Traceplots of the 200 DIME chains for selected parameters. Estimation of the benchmark model. The left panel shows a KDE of the parameter distribution. The right displays the trace of each of the chains over time.

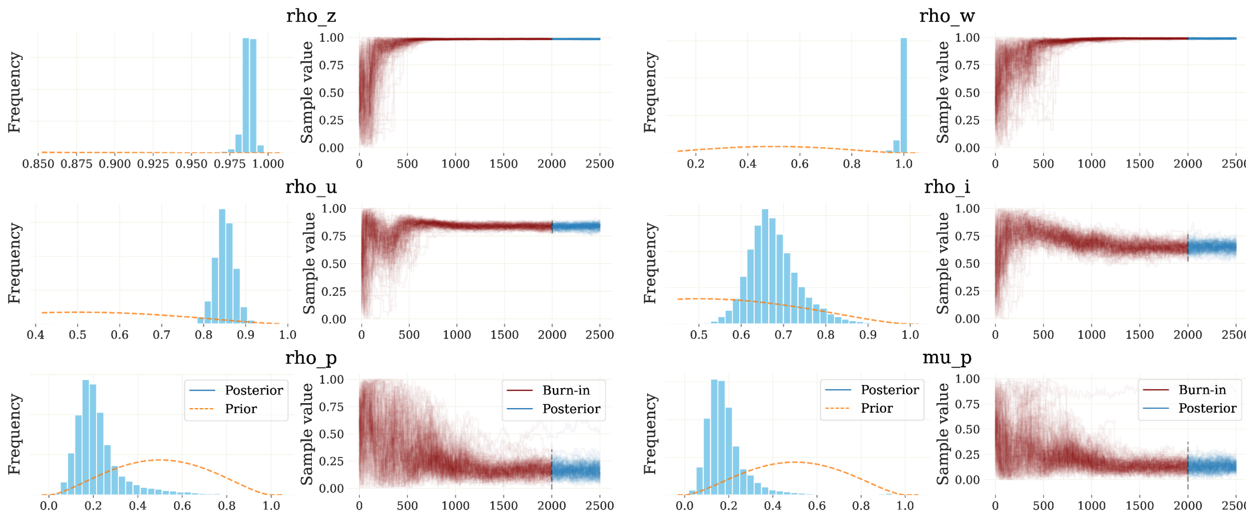

Figure D.11: Traceplots of the 200 DIME chains for selected parameters. Estimation of the benchmark model. The left panel shows a KDE of the parameter distribution. The right displays the trace of each of the chains over time.

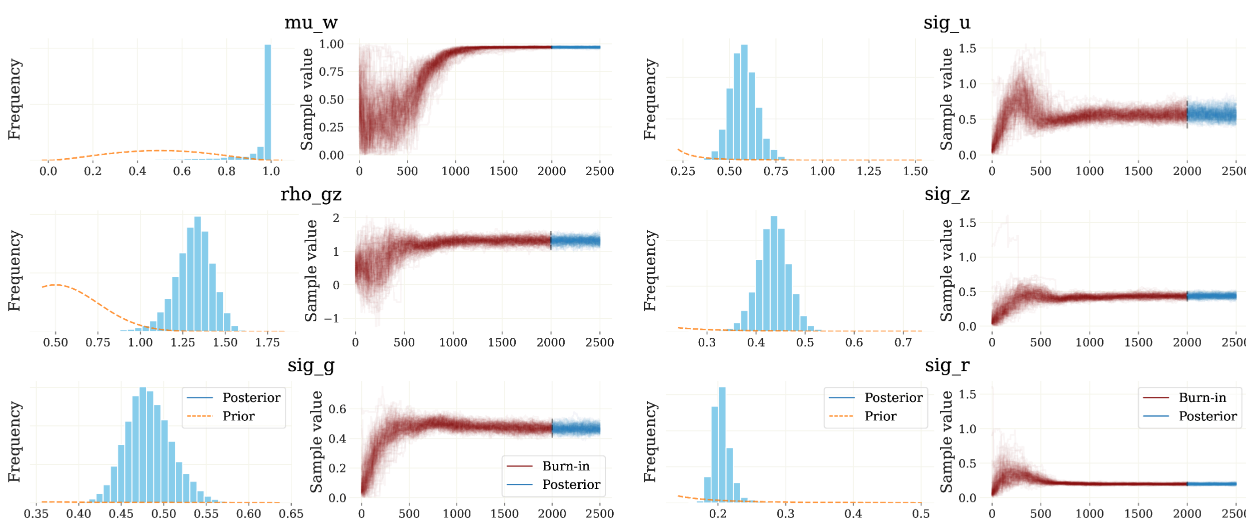

Figure D.12: Traceplots of the 200 DIME chains for selected parameters. Estimation of the benchmark model. The left panel shows a KDE of the parameter distribution. The right displays the trace of each of the chains over time.

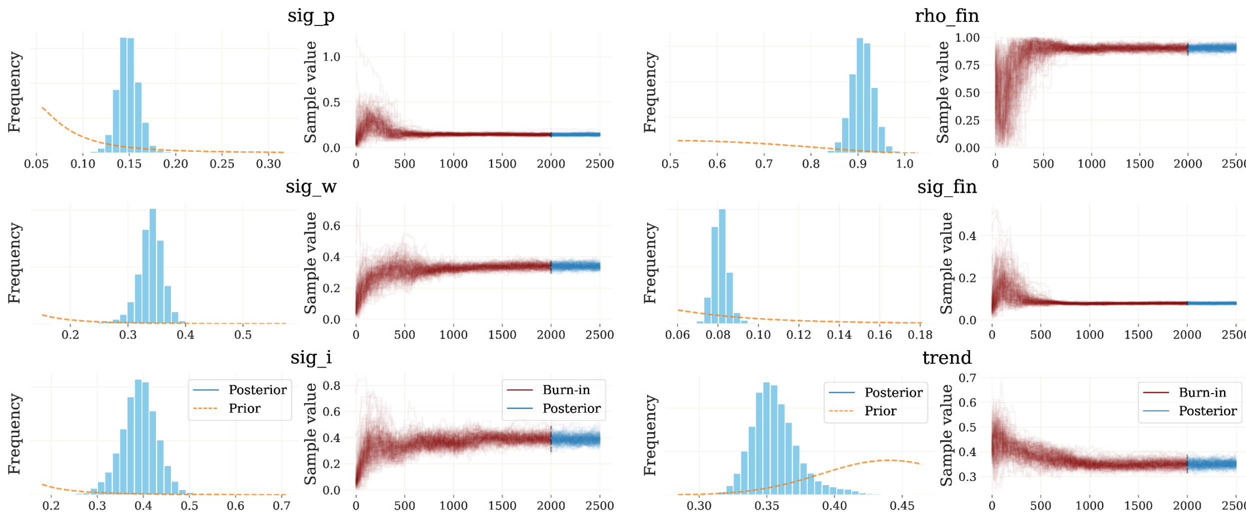

Figure D.13: Traceplots of the 200 DIME chains for selected parameters. Estimation of the benchmark model. The left panel shows a KDE of the parameter distribution. The right displays the trace of each of the chains over time.

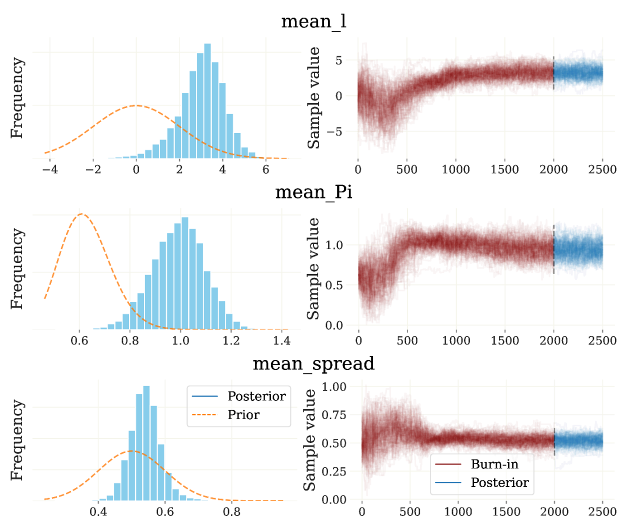

Figure D.14: Traceplots of the 200 DIME chains for selected parameters. Estimation of the benchmark model. The left panel shows a KDE of the parameter distribution. The right displays the trace of each of the chains over time.

Appendix E Comparison of the Ensemble Kalman filter and the Inversion filter in artificial data sets in which the ELB is not binding

	Prior			**EnKF no ELB**				**IVF no ELB**			
	type	mean	std	mean	NRMSE	HDP: 5%	95%	mean	NRMSE	HDP: 5%	95%
σ_c	normal	1.156	0.375	1.219	0.618	1.086	1.321	1.231	0.962	1.051	1.450
σ_l	normal	3.333	0.750	3.366	0.357	3.067	3.585	3.277	0.566	2.888	3.731
β_{tpr}	gamma	0.147	0.100	0.154	1.295	0.107	0.184	0.179	3.910	0.078	0.306
h	beta	0.635	0.100	0.608	0.405	0.573	0.651	0.626	0.358	0.583	0.682
S''	normal	5.140	1.500	5.147	0.555	4.533	5.815	5.359	0.704	4.671	6.074
ι_p	beta	0.657	0.150	0.674	0.516	0.596	0.746	0.680	0.845	0.554	0.760
ι_w	beta	0.528	0.150	0.532	0.481	0.477	0.578	0.526	0.621	0.458	0.596
α	normal	0.173	0.050	0.162	0.712	0.140	0.181	0.158	0.882	0.131	0.181
ζ_p	beta	0.904	0.100	0.896	0.178	0.861	0.928	0.916	0.263	0.862	0.967
ζ_w	beta	0.817	0.100	0.817	0.241	0.776	0.863	0.823	0.285	0.789	0.895
Φ_p	normal	1.440	0.125	1.468	0.275	1.397	1.548	1.446	0.261	1.331	1.516
ψ	beta	0.502	0.150	0.498	0.689	0.406	0.566	0.483	1.103	0.358	0.593
ϕ_π	normal	2.190	0.250	2.225	0.235	2.115	2.323	2.232	0.325	2.134	2.417
ϕ_y	normal	0.173	0.050	0.177	0.785	0.151	0.205	0.164	1.122	0.127	0.209
ϕ_{dy}	normal	0.254	0.050	0.261	0.540	0.235	0.294	0.259	0.622	0.218	0.286
ρ	beta	0.870	0.100	0.866	0.173	0.835	0.902	0.875	0.210	0.836	0.915
ρ_r	beta	0.098	0.200	0.102	4.534	0.028	0.180	0.115	4.285	0.038	0.192
ρ_g	beta	0.900	0.200	0.936	0.340	0.896	0.969	0.960	0.505	0.923	0.998
ρ_z	beta	0.900	0.200	0.983	0.657	0.972	0.995	0.988	0.691	0.974	0.998
ρ_u	beta	0.836	0.200	0.836	0.294	0.775	0.890	0.832	0.284	0.772	0.879
ρ_p	beta	0.167	0.200	0.143	2.014	0.080	0.211	0.168	2.082	0.097	0.246
ρ_w	beta	0.900	0.200	0.952	0.465	0.894	0.980	0.960	0.609	0.913	0.996
ρ_i	beta	0.651	0.200	0.654	0.655	0.583	0.771	0.717	1.044	0.636	0.858
μ_p	beta	0.140	0.200	0.104	2.434	0.054	0.154	0.126	2.112	0.065	0.192
μ_w	beta	0.900	0.200	0.946	0.403	0.912	0.975	0.924	0.508	0.871	0.976
ρ_{gz}	normal	1.316	0.250	1.313	0.443	1.214	1.473	1.290	0.443	1.171	1.400
σ_g	IG	0.467	0.250	0.454	0.428	0.415	0.489	0.471	0.376	0.434	0.503
σ_u	IG	0.574	0.250	0.564	0.876	0.448	0.691	0.601	0.944	0.502	0.739
σ_z	IG	0.437	0.250	0.385	1.010	0.343	0.450	0.443	0.481	0.405	0.492
σ_r	IG	0.197	0.250	0.195	0.552	0.173	0.220	0.207	0.683	0.180	0.238
σ_p	IG	0.143	0.250	0.139	0.713	0.110	0.156	0.144	0.748	0.115	0.167
σ_w	IG	0.340	0.250	0.347	0.533	0.303	0.383	0.330	0.586	0.291	0.372
σ_i	IG	0.387	0.250	0.388	0.721	0.327	0.437	0.362	0.829	0.300	0.416
$\bar{\gamma}$	normal	0.351	0.050	0.350	0.357	0.327	0.384	0.351	0.467	0.305	0.381
$\bar{l}$	normal	3.257	2.000	3.157	1.283	2.150	4.019	2.624	5.666	-1.189	6.143
$\bar{\pi}$	gamma	0.936	0.100	0.953	0.243	0.909	1.001	1.029	1.505	0.804	1.342

Table E.3: Comparison of the EnKF with results obtained using the IVF, using 50 artificial datasets in which the ELB is *not* binding.