Efficient and robust inference of models with occasionally binding constraints

Massimo Giovannini[a], Philipp Pfeiffer[b], Marco Ratto[a]

[a] *European Commmission, JRC*
[b] *European Commmission, DG ECFIN*

Abstract

This paper proposes a piecewise-linear Kalman filter (PKF) to estimate DSGE models with occasionally binding constraints. This method expands the set of models suitable for nonlinear estimation. It straightforwardly handles missing data, non-singularity (more shocks than observed time series), and large-scale models. We provide several applications to highlight its efficiency and robustness compared to existing methods. Our toolkit integrates the PKF into Dynare, the most popular software in DSGE modeling.

Keywords: DSGE, occasionally binding constraints, nonlinear estimation, Piecewise Kalman Filter
JEL classification: C11, C32, C51

*The opinions and views expressed in this paper are those of the authors only and should not be considered as official positions of the European Commission.

**Corresponding author: Marco Ratto, European Commission, Joint Research Centre, TP058, Via E. Fermi 2749, 21027 Ispra VA, Italy.

Acknowledgments: We thank Alexander Richter for sharing his code with us.

1. Introduction

Dynamic Stochastic General Equilibrium (DSGE) models with occasionally binding constraints (OBCs) have become central to economic analysis and policy advice. For example, the nonlinearities imposed by the zero lower bound (ZLB) on nominal interest rates, borrowing constraints, or downward nominal wage rigidity can alter standard policy prescriptions.[1] Given current projections on nominal interest rates, the ZLB is likely to stay a recurring feature of advanced economies.[2] At the same time, sound policy advice requires models with satisfactory empirical performance. While it is now common practice to estimate linear models using historical data, ignoring OBCs can lead to a biased inference by missing the stronger internal propagation mechanism.[3] Yet, when taking DSGE models with OBCs to the data, applied researchers often find the existing estimation methods computationally too expensive or not applicable to their model of interest.

To deal with this challenge, we propose a general, efficient, and easily implementable estimation procedure for DSGE models with OBCs. Our approach builds on OccBin (Guerrieri and Iacoviello, 2015), a version of the perfect-foresight extended path (Fair and Taylor, 1983).[4] This solution method handles OBCs as different regimes in which the constraints are either slack or binding and represents the model as a time-varying state-space system. As emphasized in Guerrieri and Iacoviello (2015), the interaction of the expected regime length and state variables can result in highly nonlinear dynamics. The piecewise-linear solution can thus capture important economic aspects of OBCs. At the same time, it can easily handle models with many state variables. The basic idea of our proposal is to embed this fast iterative algorithm into a diffuse Kalman filter, resulting in a piecewise-linear Kalman filter (PKF). Since likelihood-based model estimations require solving the model for many parameter draws from a proposal distribution, the PKF exploits the efficiency of the piecewise-linear solution. It is worth noting that while we build on the OccBin toolkit, the PKF is a general filter suitable for any piecewise-linear

[1] See Christiano et al. (2011) for an application to the ZLB. Guerrieri and Iacoviello (2017) discuss the impact of household borrowing constraints. Burgert et al. (2021) study fiscal policy in a monetary union under downward nominal wage rigidity.

[2] See e.g. the OECD (2021) short-term interest rates forecast.

[3] For example, Gust et al. (2017) find that a nonlinear model with a ZLB produces to significantly different parameter estimates compared to its linearized counterpart. In particular, the linear model overstates the importance of exogenous sources of business cycles. See also Hirose and Inoue (2016); Atkinson et al. (2019).

[4] Kulish et al. (2017) also apply the extended path but treat the duration of ZLB as an exogenous process. By contrast, the regime length is endogenous in the OccBin toolkit and its implementation in the PKF.

model.[5]

Compared to the two main existing approaches for estimating models with OBCs, namely the inversion filter (IF) and the particle filter, the PKF is a general and flexible alternative.[6] The IF is closest to our approach (e.g. Guerrieri and Iacoviello 2017). This method also builds on the piecewise-linear solution but proceeds differently for the estimation. It solves for the innovations that minimize the distance between the data and the model predictions. While the IF is relatively fast, one drawback is its tight restriction on the innovation structure: the number of shocks needs to be equal to the number of observables. The direct mapping from shocks to observed variables also precludes accounting for the uncertainty about data (measurement errors) and initial state variables.

The second method solves the model fully nonlinearly and evaluates the likelihood function with a particle filter. Unlike the deterministic OccBin algorithm, the full nonlinear solution captures the role of risk for economic behavior and other model nonlinearities. However, in this case, filtering becomes highly expensive as it requires a costly model solution for each parameter draw. Moreover, particle filtering requires sensitive assumptions about measurement error.[7] In practice, these factors currently limit this approach to small models, which, depending on the application, miss crucial economic transmission channels.[8]

Against this backdrop, we highlight two main contributions of the PKF. First, the PKF is robust and generalizable to a large class of models that feature more shocks than observables. Such "non-singularity" is a property of many standard models. We mention here three common examples. One, in models with news shocks or heterogeneous beliefs, agents typically do not observe innovations to exogenous shock processes. Instead, they form expectations based on partial information by solving a signal extraction problem. As a consequence, this class of models features additional shocks and more shocks than observables.[9] Two, richer (policy) models such as Smets and Wouters (2003) often fea-

[5]For example, Boehl (2021) proposes a promising and very fast piecewise-linear solution algorithm.

[6]Farmer (2021) provides a fast nonlinear approximation of the likelihood in non-Gaussian state-space models. Aruoba et al. (2020) exploit the piecewise-linear nature of OBCs for high-order solution and particle filtering. These methods are particularly suitable for low-dimensional state-spaces but are difficult to scale to models with more state variables. For Markov-switching models, Maih (2015) provides efficient perturbation methods.

[7]See Cuba-Borda et al. (2019).

[8]Indeed, Atkinson et al. (2019)[p.15] find "far larger gains in accuracy from estimating a richer, less misspecified piecewise-linear model."

[9]Campbell et al. (2019) provide a noteworthy application by estimating the strength of forward guidance policy in a rich DSGE model with noisy signals. In a comment, Bielecki et al. (2019) raise concerns about potentially biased results due to the linearized model approximation. They conclude that "estimating a richly specified DSGE model of a size similar to that used in policy institutions, in

ture more shocks than observables, too.[10] For example, in European Commission (2020) we extend a large-scale model (Albonico et al., 2019) to explore novel additional shocks to capture the effects of social distancing and virus containment measures on savings behavior during the COVID-19 pandemic.[11] Three, important model misspecification tests necessitate expanding the structure of exogenous innovations. For example, the agnostic structural disturbance approach (Den Haan and Drechsel, 2020) requires additional model shocks while keeping the number of observed variables unchanged.[12]

The second advantage of the PKF is its practicality in terms of speed and implementation. In workhorse applications, the PKF is several times faster than the IF and orders of magnitude faster than the particle filter. Besides the fast linear algebra, the PKF inherits many attractive features from standard Kalman filtering routines: It handles missing data and non-stationarity in the same way as standard diffuse Kalman filters. For example, the PKF makes it easy to use all available data, including nominal interest rates during the ZLB period. Similarly, the PKF allows estimating the initial states - an essential advantage in small samples or models with high endogenous persistence. Furthermore, we provide a freely available toolbox (to be released), which integrates the PKF seamlessly into Dynare (Adjemian et al., 2011), the most popular DSGE modeling software. Adapting the PKF to different models requires only a few additional lines to a Dynare code as we show in Section 6.

The paper proceeds as follows. Section 2 discusses the details of the filter. The first (prediction) step is a standard estimation of the current state variables. In the second (update) step, the algorithm updates state and covariance matrices and iterates one-step backward to estimate the shocks. Using the guess and verify strategy of OccBin, this step is tailored to models with OBCs. Given the initial condition, we use the shocks to update matrices of the piecewise-linear system satisfying the OBC. This fixed-point algorithm continues for each time t until convergence. The algorithm is very fast in practice and handles general shock structures. Filter initialization, likelihood evaluation, and parameter estimation with OBCs become straightforward.

Based on these theoretical considerations, two workhorse model applications show that the PKF provides fast and reliable parameter estimates. Section 3 benchmarks the PKF using a small consumption-savings model as in Cuba-Borda et al. (2019) (henceforth CGIZ). To capture essential features of a larger class of models, the model features a

a way that would explicitly account for the nonlinearity created by the [Z]LB, does not seem feasible at the moment" (p.136).

[10]The model in Smets and Wouters (2003) features ten structural shock to explain seven observable variables.

[11]See European Commission (2020)[p.61-65] and our discussion in Section 5.

[12]See also the shock selection algorithm proposed by Ferroni et al. (2019).

nonlinearity stemming from borrowing constraints. For artificial data sets generated with the piecewise-linear model, both the PKF and IF recover the true parameters. By contrast, using the fully nonlinear model as a data-generating process (DGP) implies an approximation error when using the PKF and IF. However, the advantages of the PKF are significant speed gains, a more predictable computing time and the possibility include more shocks than observables.

Section 4 turns to a workhorse New-Keynesian model with an endogenous ZLB constraint. Following Atkinson et al. (2019) (henceforth ART), we use a large number of datasets generated by fully nonlinear simulations of this model. The datasets contain either no ZLB event or a long-lasting ZLB of 30 quarters. The latter is comparable to the recent experience of advanced economies. For each dataset, we estimate multiple parameters and shocks using the PKF, the IF, and a standard linear Kalman filter (Lin-KF). In particular, for datasets with long-lasting ZLB events, the PKF improves upon the Lin-KF and the IF. The estimation accuracy, measured by the distance of the parameter estimates from their true value, also outperforms all methods and model specifications reported in ART, including small-scale (nested) model versions solved fully nonlinearly and estimated with the particle filter. Only the PKF includes the true value in the (5,95)-percentiles of all parameter estimates. In line with ART, we find that estimating a larger but less misspecified model with the PKF outweighs the costs of abstracting from high-order risk in this application.

Section 5 briefly discusses the suitability of the PKF for large models. Section 6 provides a practical guide on how to implement our method in Dynare.

Section 7 concludes by discussing an important caveat of the piecewise-linear solution built in the PKF. While the approximate model solution features a kink in the decision rules created by the OBC, it misses other nonlinearities and precautionary motives. Thus, some applications may require computationally more expensive methods to capture the full nonlinear effects. However, even in these cases, the PKF provides an easily implementable benchmark.

2. The Piecewise Kalman Filter

This section describes the PKF. Based on a general piecewise-linear model (Section 2.1), Section 2.2 and Section 2.3 describe the algorithm for filtering and fixed-interval smoothing. Finally, Section 2.4 discusses likelihood computations with the PKF and the IF and compares their respective filter initialization and computational efficiency.

2.1. The piecewise-linear model

The local linear approximation of the policy function of a DSGE model featuring OBC solved with the piecewise-linear approach can be represented as:

$$x_t = \mathbf{T}(x_{t-1}, \varepsilon_t)x_{t-1} + \mathbf{C}(x_{t-1}, \varepsilon_t) + \mathbf{R}(x_{t-1}, \varepsilon_t)\varepsilon_t, \tag{1}$$

where x_t is the vector of endogenous variables in deviation from the steady state of the 'baseline' regime [13] and ε_t is the vector of iid shocks, $\varepsilon_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$. The reduced form matrices $\mathbf{T}$, $\mathbf{C}$, and $\mathbf{R}$ are state-dependent functions of the lagged states and the current period shocks. Note that $\mathbf{C}(x_{t-1}, \varepsilon_t) \neq 0$ implies that given the states and the shocks, the constraint is expected to bind at least some time in the future.

A new shock in $t+1$ can require updating the state matrices. Therefore, the one-step recursion implies, in general, that:

$$\begin{aligned}
\mathbf{T}(x_t, \varepsilon_{t+1}) &\neq \mathbf{T}(x_{t-1}, \varepsilon_t) \\
\mathbf{C}(x_t, \varepsilon_{t+1}) &\neq \mathbf{C}(x_{t-1}, \varepsilon_t) \\
\mathbf{R}(x_t, \varepsilon_{t+1}) &\neq \mathbf{R}(x_{t-1}, \varepsilon_t).
\end{aligned}$$

To ease notation, let us re-define the state matrices as:

$$\begin{aligned}
\mathbf{T}_{t|t} &= \mathbf{T}(x_{t-1}, \varepsilon_t) \\
\mathbf{C}_{t|t} &= \mathbf{C}(x_{t-1}, \varepsilon_t) \\
\mathbf{R}_{t|t} &= \mathbf{R}(x_{t-1}, \varepsilon_t)
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{T}_{t|t-1} &= \mathbf{T}(x_{t-1}, 0) \\
\mathbf{C}_{t|t-1} &= \mathbf{C}(x_{t-1}, 0) \\
\mathbf{R}_{t|t-1} &= \mathbf{R}(x_{t-1}, 0).
\end{aligned}$$

[13] For expositional purposes and without loss of generality, we refer here to the baseline as the unconstrained regime, where the OBCs are slack. This setting is typical for ZLB constraints, as shown in Section 4. However, in some applications (such as the borrowing constraint model in Section 3), the constraint binds in the baseline regime. The PKF can flexibly handle either case.

2.2. State filtering and the likelihood

Assume we want to estimate the deep parameters of the model, given a set of observables y_t linked to x_t by the observation equation

$$y_t = \mathbf{H}x_t, \tag{2}$$

where, for simplicity and without loss of generality, we assume no observation error.[14]

Given the initial state mean and variance x_0 and $\mathbf{P}_0$, we denote their 'best' estimate at any time $t-1$ as

$$x_{t-1|t-1}, \mathbf{P}_{t-1|t-1}. \tag{3}$$

Algorithm outline. For each t, we guess and verify the state matrices in four steps. Given a guess, Steps 1 and 2 follow the prediction and update process of the standard Kalman Filter (e.g. Koopman and Durbin (2003)). For the Lin-KF, these two steps are sufficient. By contrast, the model with OBCs requires two additional steps because the time-varying state matrices, i.e. the OBC's timing and duration, depend on the shocks in t and states in $t-1$ *given* t. Incorporating this aspect is the PKF's main innovation: Using the shock and state estimates obtained in the third (smoother) step, Step 4 simulates the piecewise-linear model to find the regime sequence(s) and verify the guessed matrices. The algorithm then iterates until convergence.

Details. Following this outline, we now provide the details of this algorithm. We initialize the guess of the updated state matrices as:

$$\begin{aligned} \mathbf{T}(0)_{t|t} &= \mathbf{T}_{t|t-1} \\ \mathbf{R}(0)_{t|t} &= \mathbf{R}_{t|t-1} \\ \mathbf{C}(0)_{t|t} &= \mathbf{C}_{t|t-1}. \end{aligned}$$

Then, we iterate until convergence. Each iteration j follows the algorithm:

[14] We initialize the filter with the unconditional mean and variance of the 'baseline' regime with diffuse priors.

1. Prediction step: Given the guessed matrices $\mathbf{T}(j-1)_{t|t}$,$\mathbf{R}(j-1)_{t|t}$, $\mathbf{C}(j-1)_{t|t}$,

$$\begin{aligned} x(j)_{t|t-1} &= \mathbf{T}(j-1)_{t|t} \cdot x_{t-1|t-1} + \mathbf{C}(j-1)_{t|t} \\ \mathbf{P}(j)_{t|t-1} &= \mathbf{T}(j-1)_{t|t} \cdot \mathbf{P}_{t-1|t-1} \cdot \mathbf{T}'(j-1)_{t|t} + \mathbf{R}(j-1)_{t|t} \cdot \mathbf{Q} \cdot \mathbf{R}'(j-1)_{t|t} \\ y(j)_{t|t-1} &= \mathbf{H}x(j)_{t|t-1} \\ \mathbf{F}(j)_t &= \mathbf{H} \cdot \mathbf{P}(j)_{t|t-1} \cdot \mathbf{H}' \\ v_(j)t &= z_t - y(j)_{t|t-1}, \end{aligned} \tag{4}$$

   where z_t are the observations of y_t.
2. Update state and covariance matrices given the guess

$$\begin{aligned} \mathbf{K}(j)_t &= \mathbf{P}(j)_{t|t-1} \cdot \mathbf{H}' \\ x(j)_{t|t} &= x(j)_{t|t-1} + \mathbf{K}(j)_t \mathbf{F}(j)_t^{-1} v(j)_t \\ \mathbf{P}(j)_{t|t} &= \mathbf{P}(j)_{t|t-1} - \mathbf{K}(j)_t \mathbf{F}(j)_t^{-1} \mathbf{K}(j)_t'. \end{aligned} \tag{5}$$

3. Compute the estimate of the shocks in t:

$$\varepsilon(j)_{t|t} = \mathbf{Q} \cdot \mathbf{R}'(j-1) \cdot \mathbf{H}' \mathbf{F}(j)_t^{-1} v(j)_t, \tag{6}$$

   and iterate one step backward (smoother step) to also update the state in $t-1$ given t, i.e. for $s = t, t-1$

$$\begin{aligned} \mathbf{L}_s &= \mathbf{I} - \mathbf{K}(j)_s \mathbf{F}(j)_s^{-1} \mathbf{H} \\ r(j)_s &= \mathbf{H}' \mathbf{F}(j)_s^{-1} v(j)_s + \mathbf{L}_s' \mathbf{T}'(j-1)_{s|s} r(j)_{s+1} \\ x(j)_{s|t} &= x(j)_{s|s-1} + \mathbf{P}(j)_{s|s-1} \cdot r(j)_s \end{aligned} \tag{7}$$

   where $\mathbf{I}$ denotes the identity matrix. Note that the one-step backward recursion is initialized by setting $r_{t+1} = 0$.
4. To find the regimes, i.e. the sequence of OBCs, implied by the states and shocks, project the piecewise-linear model given the initial condition $x(j)_{t-1|t}$ and shock $\varepsilon(j)_{t|t}$ for $s \in (t, \infty)$ and update the matrices $\mathbf{T}(j)_{t|t}$, $\mathbf{R}(j)_{t|t}$, $\mathbf{C}(j)_{t|t}$
   (a) if the update differs from the guess, set the guess to $\mathbf{T}(j)_{t|t}$,$\mathbf{R}(j)_{t|t}$, $\mathbf{C}(j)_{t|t}$ and restart from Step 1 with $j+1$.

(b) otherwise, proceed to $t+1$ and until T, by setting the updated state matrices

$$\begin{aligned}\mathbf{T}_{t|t} &= \mathbf{T}(j)_{t|t} = \mathbf{T}(j-1)_{t|t}\\ \mathbf{R}_{t|t} &= \mathbf{R}(j)_{t|t} = \mathbf{R}(j-1)_{t|t}\\ \mathbf{C}_{t|t} &= \mathbf{C}(j)_{t|t} = \mathbf{C}(j-1)_{t|t},\end{aligned}$$

as well as states and covariances

$$\begin{aligned}x_{t|t} &= x(j)_{t|t}\\ \mathbf{P}_{t|t} &= \mathbf{P}(j)_{t|t}.\end{aligned}$$

Each step of the algorithm is simple since it applies standard Kalman filter and smoother formulae, using the guessed state matrices and regimes. For a piecewise-linear model, this algorithm is optimal in the least-squares sense.[15]

Likelihood computation. Given the prediction error

$$v_t(j) = z_t - y(j)_{t|t-1} \tag{8}$$

we can compute the log-likelihood density of the data at time t:

$$\mathcal{L}(\theta)_t = -\tfrac{1}{2} n_t \log(2\pi) - \tfrac{1}{2}\log(\det(\mathbf{F}(j;\theta)_t)) - \tfrac{1}{2} v(j)_t' \cdot \mathbf{F}(j;\theta)_t^{-1} \cdot v(j)_t, \tag{9}$$

where n_t and θ denote the number of observables available in time t and the vector of deep parameters, respectively.

Convergence. At this stage, some discussion on convergence is in order. In the "pure" OccBin simulation, the state variables are known and fixed. By contrast, both the shocks (states) *and* OBC regimes are unknown in the filtering process. For each period t, the update step iterates until convergence of the latent shocks and the regime. Independently of the filter, multiple solutions may arise in this solution method.[16] It would be possible to search for all possible solutions and select the one with the highest likelihood density. However, this approach would significantly increase the filtering costs.

In the examples discussed below, we find that, in the case of multiple solutions, the most likely solution is the shortest OBC. To look for this minimum-duration solution, we apply a relaxation in Step 4. (a): The duration of the new (guessed) constrained

[15] The optimality of the time-varying system follows directly from Kalman (1960).

[16] In fact, multiple solutions, i.e. different possible regime sequences, can also arise for the simpler standard OccBin simulations with given initial states and the shocks. See Holden (2019).

regime is not allowed to deviate (increase) too much from the previous iteration. For this purpose, we apply the new guess matrices $\mathbf{T}(j^*)_{t|t}$, $\mathbf{R}(j^*)_{t|t}$, $\mathbf{C}(j^*)_{t|t}$, where the duration of the new binding regime may increase only up to the allowed maximum amount. With this relaxation, a few iterations are typically sufficient for convergence. If it fails, we give a penalty to the likelihood and try a new proposal for the deep parameters.

2.3. Fixed-interval smoothing

As a corollary of the PKF recursions described above, the fixed-interval smoothing proceeds as follows. We first guess sequence of the state matrices for $t \in (1, T)$, e.g. using the first forward path needed to compute the likelihood:

$$\begin{aligned} \mathbf{T}(0)_{t|T} &= \mathbf{T}_{t|t} \\ \mathbf{R}(0)_{t|T} &= \mathbf{R}_{t|t} \\ \mathbf{C}(0)_{t|T} &= \mathbf{C}_{t|t}. \end{aligned}$$

Then, we iterate until convergence with the following algorithm for each iteration j:

1. Run the standard filtering and smoothing recursions given the guess sequence of state matrices and the estimated regimes $\mathbf{T}(j-1)_{t|T}$,$\mathbf{R}(j-1)_{t|T}$, $\mathbf{C}(j-1)_{t|T}$.
2. From the obtained sequence of smoothed shocks $\varepsilon(j)_{t|T}$ and initial conditions $x(j)_{0|T}$, we simulate the piecewise-linear method, obtaining a new sequence of matrices/regimes $\mathbf{T}(j)_{t|T}$,$\mathbf{R}(j)_{t|T}$, $\mathbf{C}(j)_{t|T}$
   (a) if the entire sequence of matrices/regimes corresponds to the guess, convergence is achieved;
   (b) otherwise restart from 1) with j+1 using the new guess matrices $\mathbf{T}(j)_{t|T}$,$\mathbf{R}(j)_{t|T}$, $\mathbf{C}(j)_{t|T}$.

This fixed-interval smoothing is equivalent to other smoothing recursions in the literature (Koopman and Durbin, 2003). Similarly, it flexibly handles missing data and can accommodate more complex shock structures. As for the forward path, each step of the procedure is straightforward. Using the guessed sequence of matrices, it applies standard filtering and smoothing formulas, providing the best estimates in the least-square sense of the historical shocks, as well as latent and initial state variables.

2.4. Computing the likelihood with the IF and with the PKF

We now show under which conditions the likelihood of the PKF and the IF are identical. Recall first the IF's likelihood formula:

$$\mathcal{L}(\theta)_t^{IF} = -\tfrac{1}{2} n_t \log(2\pi) - \tfrac{1}{2}\log(\det(\mathbf{Q})) - \tfrac{1}{2}\varepsilon_t' \cdot \mathbf{Q}^{-1} \cdot \varepsilon_t + \log\left(|\det \frac{\partial \varepsilon_t}{\partial y_t}|\right), \qquad (10)$$

where $\frac{\partial \varepsilon_t}{\partial y_t}$ is the Jacobian matrix mapping shocks onto observables.[17] For a generic nonlinear model, and assuming for simplicity no observation error,

$$y_t = f(x_{t-1}, \varepsilon_t; \theta) \tag{11}$$

the Jacobian is only implicit and has to be approximated numerically. For the piecewise-linear model, it can be derived explicitly combining (1) and (2).

The likelihood computation with the IF requires two additional restrictions. First, the number of shocks must exactly equal the number of observables in each period t. Second, the shocks in the first period $t = 1$ are initialized assuming all state variables are at their steady-state values in $t = 0$, i.e. $x_0 = 0$. Then, the IF obtains the shocks by numerically solving the nonlinear system of equations (11) for $\varepsilon_t \, \forall t$.

Filter initialization. The filter initialization drives the difference in the likelihood between PKF and IF. To see this, note that in (7), the PKF also provides $\varepsilon_{t|t}$, i.e. the estimate of the shocks in time t. Hence, it can also compute the likelihood consistently with the IF. To do so, note that the IF is conditional to state variables being in steady state for the initial period $t = 0$, while in the PKF initial states are latent in $t = 0$.[18]

What happens if we relax the two additional requirements of the IF? First, with the PKF, we can handle problems with more shocks than observables, while this is impossible with IF. Second, if we let the state variables latent in $t = 0$, the likelihood value from (9) can differ from (10) even if the number of shocks is the same. Two main factors explain this difference:

1. the initial values of state variables and shocks;
2. the recursion of the covariance matrix $\mathbf{F}$ has one iteration less, the missing period $t = 0$.

Using equation (10) with the shocks ε_t from the PKF may eliminate the second factor. However, initialization still matters. Unlike the IF, the PKF provides the optimal starting values for states and shocks in a least-square sense. While the initialization effects will vanish for large samples, in general, they matter for small samples. Furthermore, the differences likely increase in the estimation of shock variances. In case of the IF (eq. 10),

[17] See equation (23) in Guerrieri and Iacoviello (2017).

[18] To compute the filtered shocks with the PKF conditional to the state variables being at steady state, we can add an auxiliary data point at the beginning of the data sample. All observed variables and states are at their steady state, i.e. $x_0 = 0$, $y_0 = 0$. In this way, the PKF recovers the same sequence of shocks as the IF algorithm, and therefore equations (9) and (10) provide the same value of the likelihood. For both algorithms we enforce $\mathcal{L} = \sum_{t=1}^{T} \mathcal{L}_t$ and $\mathcal{L}^{IF} = \sum_{t=1}^{T} \mathcal{L}_t^{IF}$ by discarding the likelihood density for $t = 0$.

the variance $\mathbf{Q}$ does not influence the sequence of shocks of ε_t. Hence, it only 'scales' the shocks in the likelihood. By contrast, in the PKF (eq. 9), $\mathbf{Q}$ not only governs the covariance matrix $\mathbf{F}_t$. It also determines the initialization of latent states and shocks in $t = 1$ and, thereby, the prediction errors v_t in the likelihood function. Overall, we can expect the difference in the likelihood values between IF and PKF to decrease with sample size and increase with the strength of the internal propagation mechanism.

Computational efficiency. The linear recursive algebra of the PKF increases computational efficiency compared to the nonlinear solver incorporated in the IF. The PKF breaks one non-linear problem into two simpler ones, identified in Steps 3 and 4 of the PKF algorithm. In Step 3, for a given regime, the PKF exploits linear algebra and computes shocks as in the standard Kalman filter. In Step 4, given the shocks, the OccBin solver is used to update the regime. By contrast, the IF numerically solves a system of nonlinear equations for the regime *and* the shocks by calling the OccBin solver several times. Overall, the PKF requires fewer calls to the solver, resulting in significant speed gains as we show in the following sections.

3. Application 1: A small model of consumption choice

Our first application considers the small consumption-savings model of CGIZ. By embedding an occasionally binding borrowing limit, this prototypical model captures key nonlinear features of OBCs, such as shock amplification and precautionary savings. We use data generated from this model as a testbed for the PKF and other estimation methods.

3.1. The model

Cuba-Borda et al. (2019) provide the micro foundations of the model, in which a representative household maximizes utility subject to stochastic income shocks and an endogenous occasionally binding borrowing constraint. Here, we only state necessary conditions and refer for further model details to their paper. The equilibrium system consists of four equations in four unknowns $\{C_t, B_t, \lambda_t, Y_t\}$ with parameter vector $\theta = [\beta, \gamma, m, R, \rho, \sigma]$ and exogenous innovation ϵ_t:

$$C_t + RB_{t-1} = Y_t + B_t \tag{12}$$

$$\ln(Y_t) = \rho \ln(Y_{t-1}) + \sigma\epsilon_t \tag{13}$$

$$C_t^{-\gamma} = \beta R E_t(C_{t+1}^{-\gamma}) + \lambda_t \tag{14}$$

$$\lambda_t(B_t - mY_t) = 0, \tag{15}$$

where $E_t()$ denotes the expectation operator, C_t is consumption, B_t is debt, λ_t is the Lagrange multiplier associated with the borrowing constraint and Y_t is the exogenous stochastic income process. Equation (15) represents the complementary slackness condition associated to the borrowing constraint of the Kuhn-Tucker formulation. Our calibration is identical to CGIZ, i.e. $\theta = [0.945, 1, 1, 1.05, 0.9, 0.01]$.[19] Below, we infer the constant relative risk aversion parameter γ, which also governs intertemporal substitution in this model. In the DGP, $\gamma = 1$.

3.2. The two DGPs

The calibrated model forms the basis of two DGPs, which differ in the solution method used to simulate the artificial data. The first DGP is a piecewise-linear model approximation of the model. The solution underlying these datasets thus captures the occasionally binding borrowing limit. However, due to the piecewise-linear approximation, high-order risk vanishes, and there are no precautionary motives at work when generating the data. By contrast, the second DGP features precautionary savings as it is based on a fully nonlinear solution coming from value function iteration (VFI).

For both DGPs, we generate 100 replicas, each of length $T = 100$. Following CGIZ, we focus on the inference of γ, setting all other parameters to their true value. Consumption is the only observed variable.

3.3. Inference and computational efficiency

[Figure 1 about here.]

For each replica, we evaluate the likelihood with the PKF, the IF, and the Lin-KF.[20] The left panel of Figure 1 plots the average likelihood profiles for different values of γ. For the piecewise-linear DGP (top), the likelihood of the PKF and the IF peaks at the true value of ($\gamma = 1$). Thus, as expected in the absence of solution error, both methods successfully capture the borrowing constraint and identify the risk aversion parameter without bias.[21] By contrast, the standard Lin-KF cannot identify γ. The Lin-KF constructs the likelihood around the steady state where the constraint is always binding. In this case, the model solution does not feature intertemporal smoothing, and the corresponding likelihood function is flat.

[19] Note that $\beta < 1/R$ implies that in steady state the borrowing constraint is binding.

[20] CGIZ report estimations with the particle filter.

[21] Note, however, that the different filter initialization implies small differences in the likelihood as discussed in Section 2.4. Appendix B discusses this issue based on this small model. Section 4 will also return to it in the context of a larger model in which the initialization affects multiple parameter estimates.

Table 1: Computing times (seconds)

	Mean PKF	Mean IF	Std PKF	Std IF
OccBin DGP	101.38	409.21	27.71	122.01
VFI DGP	70.73	424.26	13.88	85.90

Notes: The this table reports the average computing time in seconds for the IF and the PKF across 100 replicas and 100 parameter values (ranging from 0 to 4.5 on an evenly distributed grid).

By abstracting from solution errors, the first DGP singles out the efficiency gains of the PKF. Table 1 shows that the PKF's likelihood evaluation is more than four times faster than the IF. The computing time is also less variable, increasing the predictability when planning long-lasting estimation exercises. Both PKF and IF are orders of magnitude faster than the particle filter (reported in CGIZ) and considerably easier to implement.

The second DGP based on VFI (bottom of panel (a) in Figure 1) highlights the role of approximation error in the filter. The likelihood evaluations with the PKF and IF miss the precautionary savings present in the DGP by relying on the piecewise-linear solution. This misspecification biases the estimate upwards. The likelihood function peaks above the true value, suggesting some limitations of the piecewise linear approach. CGIZ discusses these aspects in detail.

3.4. Multiple shocks

Besides the efficiency gains shown in Table 1, the PKF's flexibility is a further advantage. Unlike the IF, it allows for more shocks than observables. As an illustration, we use the same data as in the exercise above. In contrast, however, we now assume that the econometrician erroneously believes that the DGP includes measurement error corresponding to 5% and 20% of the standard deviation of consumption.

While this example mainly illustrates the PKF in a simple non-singular environment, it offers some economic insights. The right panel in Figure 1 shows that the likelihood profile flattens, and the upward bias increases in the (assumed) measurement error. Intuitively, by assuming measurement error, "the econometrician sees skewed and asymmetric consumption even after accounting for normally distributed, additive measurement error" (CGIZ p.9). For the particle filter, which requires measurement error, this assumption can thus lead to a biased inference.[22]

[22]See CGIZ for details on the particle filter. Note that unlike the particle filter, the PKF does not require any measurement error.

4. Application 2: A medium-scale model

We now turn to a medium-scale New-Keynesian model with an endogenous ZLB constraint, the same as in ART. Based on a large number of datasets generated by this model, we compare the accuracy of the PKF to existing methods.

4.1. The data generating process

The DGP is a workhorse New-Keynesian model featuring an occasionally binding zero lower bound constraint. To generate artificial time series, we use the fully nonlinear projection solution provided by ART. For the sake of brevity, we focus on maximization objectives and constraints of the different agents and the stochastic shocks. For additional details and optimality conditions, we refer to Appendix B.1 and ART.[23]

4.1.1. Firms

A continuum of monopolistically competitive firms supplies differentiated intermediate goods $y_t(j)$, $j \in [0,1]$. A perfectly competitive final goods firm purchases these goods to assemble a single final good, $y_t \equiv \left[\int_0^1 y_t(j)^{\frac{\theta_p-1}{\theta_p}} dj\right]^{\frac{\theta_p}{\theta_p-1}}$, where $\theta_p > 1$ is the elasticity of substitution. Profit maximization implies that the final good firm's demand function for good j follows $y_t(j) = \left(\frac{p_t(j)}{p_t}\right)^{-\theta_p} y_t \, \forall j$, where $p_t(j)$ and p_t are the price for intermediate good j and the aggregate price level, respectively.

The production function for intermediate good j is $y_t(j) = \left(\tilde{k}_{t-1}(j)\right)^{\alpha} (a_t n_t(j))^{1-\alpha}$, where α is a parameter. $\tilde{k}_{t-1}(j)$ and $n_t(j)$ denote capital and labor services hired at (real) rates r_t^k and w_t, respectively. Common technology evolves according to $a_t = g_t a_{t-1}$, where deviations from deterministic growth, $\bar{g}$, follow:

$$g_t = \bar{g} + \sigma_z \varepsilon_{g,t}, \quad \varepsilon_{g,t} \sim iid\,\mathcal{N}(0,1). \tag{16}$$

Intermediate firms set prices subject to quadratic price adjustment costs as in Rotemberg (1982), $adj_t^p(j) = \varphi_p/2(p_t(j)/(\bar{\pi} p_{t-1}(j)) - 1)^2 y_t$, where parameter φ_p governs the cost and $\bar{\pi}$ denotes the steady-state (gross) inflation rate. Firm j chooses $\tilde{k}_{t-1}(j), n_t(j)$, and $p_t(j)$ to maximize its expected present discount value of profits

$$E_0 \sum_{t=1}^{\infty} \beta^t \Lambda_t \left(\frac{p_t(j) y_t(j)}{p_t} - w_t n_t(j) - r_t^k \tilde{k}_{t-1}(j) - adj_t^p(j) \right), \tag{17}$$

where the marginal value of real profits in t, Λ_t, and the discount factor, β, reflect that households own all firms. E_0 is an expectation operator.

[23] The model is the same as in Gust et al. (2017), except for a few features such as government consumption and inflation indexation.

4.1.2. Households

Households choose choose consumption $c_t(\ell)$, labor, one period bonds, investment $x_t(\ell)$, the capital stock $k_t(\ell)$, and its utilization rate $\upsilon_t(\ell)$ to maximize expected life time utility

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\log(c_t(\ell) - hc_{t-1}^a) - \chi \frac{n_t(\ell)^{1+\eta} d\ell}{(1+\eta)} \right), \tag{18}$$

where h governs habits associated with aggregate consumption c^a. $1/\eta$ is the Frisch elasticity of labor supply. χ affects steady-state labor. A household's budget constraint in period t is $c_t(\ell) + x_t(\ell) + b_t/(s_t i_t) + u_t(\ell)k_{t-1}(\ell) + adj_t^w(\ell) = w_t(j)n_t(j) + r_t^k \tilde{k}_t(\ell) + b_{t-1}(\ell)/\pi_t + d_t$, where $b_t(\ell)$ is the real value of bonds and i_t the nominal interest rate. $adj_t^w(\ell)$ captures nominal wage adjustment costs (following an similar specification as price adjustment costs with scaling parameter φ_w). $\pi_t \equiv p_t/p_{t-1}$. Each household receives the same dividends, d_t, from intermediate firms anc income from capital services, $\tilde{k}_t(\ell) = k_t(\ell)\upsilon_t(\ell)$. The households' capital stock depreciates at rate $0 \geq \delta \geq 1$ and accumulates subject to adjustment costs (parametrized by ν): $k_t = (1-\delta)k_{t-1} + x_t(1 - \nu(x_t^g - 1)^2/2)$, where $x_t^g = x_t/(\bar{g}x_{t-1})$ is investment growth. The cost of capital utilization is $u_t = \bar{r}^k(\exp(\sigma_\upsilon(\upsilon_t - 1)) - 1)/\sigma_\upsilon$, where σ_υ is a parameter. s_t is a risk premium shock affecting the spread between the nominal rate and the return on risky assets following:

$$s_t = (1-\rho_s)\bar{s} + \rho_s s_{t-1} + \sigma_s \varepsilon_{s,t}, \quad \varepsilon_{s,t} \sim iid\mathcal{N}(0,1), \tag{19}$$

where ρ_s and $\bar{s}$ denotes the persistence parameter and the steady-state value, respectively.

4.1.3. Policy

The central bank sets the notional gross nominal interest rate, i_t according to a Taylor rule

$$i_t^n = \left(i_{t-1}^n\right)^{\rho_i} \left(\bar{\imath}(\pi_t/\bar{\pi})^{\phi_\pi}(y_t^g)^{\phi_y}\right)^{1-\rho_i} \exp(mp_t) \tag{20}$$

with parameters $0 < \rho_i < 1, \phi_\pi > 0, \phi_y > 0$. $y_t^g \equiv y_t^{gdp}/(\bar{g}y_{t-1}^{gdp})$ defines the output growth gap (see also below). $\bar{\pi}$ and $\bar{\imath}$ are targets of inflation and nominal interest rates, respectively. Monetary policy faces an occasionally binding lower bound on nominal interest rates, i.e. $i_t = \max\{1, i_t^n\}$. The interest rate shock process follows:

$$mp_t = \sigma_i \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim iid\mathcal{N}(0,1). \tag{21}$$

4.1.4. Market clearing

The aggregate resource constraint is $c_t + x_t = y_t^{gdp}$, where real GDP is defined as $y_t^{gdp} = [1 - \varphi_p(\pi_t/\bar{\pi} - 1)^2/2 - \varphi_w(w_t^g - 1)^2/2]y_t - u_t k_{t-1}$.

4.1.5. Calibration and fully nonlinear model solution.

The calibration and model solution follow ART. The calibration targets US data (1988Q1:2017Q4). Table B.1 in the Appendix shows the parameter values. The calibrated model is solved with a high-quality nonlinear algorithm, which accounts for the OBC. In contrast to the piecewise-linear approach, the solution method also captures the effects of uncertainty on the economic decisions of households and firms and other nonlinearities present in the model.

4.2. Data

The simulated data sets under consideration contain 120 quarters and feature either no ZLB event (0Q) or a single ZLB event lasting 30 quarters (30Q). The long ZLB episodes capture the recent experience of advanced economies (euro area, Japan, US) and provide a meaningful testbed for the different estimation methods. Fluctuations in the model are driven by three shocks ($\varepsilon_{z,t}, \varepsilon_{i,t}, \varepsilon_{s,t}$). To allow for applying the IF, the model estimations thus observe simulated data for three variables: output growth, the inflation rate, and the nominal interest rate.

Note that samples with 30 quarters ZLB events are rare in the DGP. As acknowledged in ART, this suggests potential sample selection bias.[24] Alternatively, we could randomly select samples. However, by using the same samples, our findings directly connect to ART (see below).

4.3. Estimation methods

We consider three estimation methods: PKF, IF, and the standard Lin-KF. All methods use the true model specification but differ in the filtering and model solution. The IF and the PKF estimate a piecewise-linear approximation of the true model. In contrast to the Lin-KF, the decision rules based on OccBin capture the kink of the OBC. However, all three methods miss other nonlinear features of the DGP. Therefore, it is useful to compare our findings to ART, who report results for the particle filter and fully nonlinear (but misspecified) model versions.

For all methods, we use the same 100 draws from this DGP. In total, we estimate nine parameters with the prior distributions centered around their true values (following ART). All other parameter values are calibrated to their true values.

[24] See Atkinson et al. (2019)[p.10]. Reaching 100 datasets with 30Q requires around 2 million simulations. See Table 5 in the Online Appendix of ART.

4.4. Accuracy of posterior parameter estimates

This section shows that the PKF delivers accurate parameter estimates.[25] For each method h, we measure the estimation accuracy the normalized root-mean square-error (NRMSE), i.e. the difference between the mean estimate of parameter j, $\hat{\theta}_{j,h,k}$ in dataset k and the true parameter $\bar{\theta}_j$.[26] Thus, the NRMSE for parameter j and method h follows

$$\text{NRMSE}_h^j = \frac{1}{\theta_j}\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\hat{\theta}_{j,h,k} - \bar{\theta}_j\right)^2}, \tag{22}$$

where $N = 100$ equals the number of datasets with a given ZLB duration. To measure total accuracy across all parameters, we summarize the normalized errors as $\Sigma_h = \sum_{j=1}^{J} \text{NRMSE}_h^j$. We also define the sum of normalized errors for all parameters except the shock standard deviations, denoted Σ_h^P.[27] In addition, the coverage ratio measures the probability that the posterior distribution contains the true value:

$$CR_h^j = \frac{1}{N}\sum_{k=1}^{N}\mathbb{I}\left(\hat{\theta}_{j,h,k}^{5\%} < \bar{\theta}_j < \hat{\theta}_{j,h,k}^{5\%}\right), \tag{23}$$

where $\mathbb{I}$ is an indicator function. $\hat{\theta}_{j,h,k}^{5\%}$ and $\hat{\theta}_{j,h,k}^{5\%}$ denote the (5,95)-percentiles of the posterior distribution, respectively. We define $\overline{CR}$ as the average CR across parameters.

Table 2 shows that the piecewise-linear approximation improves the accuracy of the estimation. We highlight three main points. First, only the PKF includes all true parameter values in the (5,95) percentiles of the estimates. In the datasets with a 30 quarter ZLB event, the PKF reduces the sum of the NRMSE from 1.42 to 1.20 compared to the Lin-KF. Most of the gain from accounting for the OBC comes from risk premium shocks (σ_s), pointing to missing amplification in the linear model: The weaker internal propagation mechanism in the absence of the ZLB constraint requires an overly large role of exogenous shocks. Second, the linear model's coverage ratio deteriorates substantially for the price adjustment cost (φ_p) and the Taylor rule coefficient on output (ϕ_y). Interestingly, the coverage ratio for the inflation coefficient (ϕ_π) is higher for the Lin-KF. This finding suggests that the DGP features nonlinearities not captured by the OccBin solution. In the linear model, both misspecification dimensions interact.[28] Finally, note

[25] Results are based on the mean draw. The mode estimates improve the fit across methods.

[26] Normalization by the true value, renders the NRMSE statistic invariant to the scale of the parameters.

[27] Thus, Σ_h^P sums the NRMSE for $\varphi_p, h, \rho_s, \rho_i, \phi_\pi$, and ϕ_y.

[28] Indeed, across methods, we find that the ZLB constraint deteriorates the estimates' accuracy even when accounting for the occasionally binding ZLB (PKF and IF), suggesting sample selection effects or interacting nonlinearities.

that the accuracy also improves for the 0Q datasets, where the ZLB constraint can bind in expectations.

Fully nonlinear estimation and model (mis)specification. ART report particle-filter estimates of fully nonlinear but misspecified small-scale versions of the model. The PKF's estimates of the medium-scale model are more accurate than those of a simpler but fully nonlinear version estimated with a particle filter. The latter accounts for the significantly higher computational costs associated with fully nonlinear filtering and also exemplifies the role of misspecification (inherent in all models but beyond our paper's scope). For datasets with 30 quarters of binding ZLB, the sum of the NRMSE of the PKF (1.20) is considerable than the corresponding NRMSE of the fully nonlinear smaller model (2.08).[29]

[29] See Table 4, p.8 in ART.

Table 2: Accuracy of estimated structural parameters

Ptr	Truth	PKF		IF		Lin-KF	
		0Q	30Q	0Q	30Q	0Q	30Q
φ_p	100.00	99.92	125.75	103.74	126.98	99.70	128.47
	(25.00)	(79.47,116.72)	(99.91,145.53)	(82.20,118.40)	(104.36,147.96)	(84.21,119.69)	(113.28,152.55)
	Norm	{0.12,0.96}	{0.30,0.56}	{0.12,0.98}	{0.31,0.53}	{0.12,0.95}	{0.31,0.43}
h	0.80	0.81	0.78	0.81	0.78	0.81	0.80
	(0.10)	(0.75,0.88)	(0.71,0.84)	(0.74,0.86)	(0.72,0.84)	(0.74,0.87)	(0.73,0.86)
	Beta	{0.05,0.93}	{0.06,0.72}	{0.05,0.89}	{0.05,0.79}	{0.05,0.89}	{0.05,0.85}
ρ_s	0.80	0.80	0.82	0.81	0.82	0.80	0.81
	(0.10)	(0.76,0.84)	(0.78,0.86)	(0.76,0.84)	(0.79,0.85)	(0.76,0.84)	(0.78,0.85)
	Beta	{0.03,0.93}	{0.04,0.66}	{0.03,0.91}	{0.04,0.63}	{0.03,0.89}	{0.03,0.85}
ρ_i	0.80	0.79	0.80	0.79	0.80	0.79	0.80
	(0.10)	(0.77,0.83)	(0.76,0.84)	(0.77,0.83)	(0.76,0.84)	(0.77,0.82)	(0.76,0.83)
	Beta	{0.02,0.95}	{0.03,0.83}	{0.03,0.97}	{0.03,0.85}	{0.02,0.98}	{0.03,0.94}
σ_z	0.005	0.0045	0.0053	0.0047	0.0056	0.0045	0.0055
	(0.005)	(0.0039,0.0052)	(0.0044,0.0061)	(0.0040,0.0054)	(0.0043,0.0063)	(0.0037,0.0050)	(0.0047,0.0064)
	IGam	{0.13,0.64}	{0.12,0.81}	{0.11,0.77}	{0.18,0.63}	{0.14,0.65}	{0.17,0.69}
σ_s	0.005	0.0060	0.0053	0.0059	0.0053	0.0061	0.0066
	(0.005)	(0.0043,0.0078)	(0.0039,0.0071)	(0.0043,0.0076)	(0.0041,0.0070)	(0.0042,0.0076)	(0.0044,0.0081)
	IGam	{0.31,0.88}	{0.23,0.89}	{0.28,0.88}	{0.21,0.92}	{0.34,0.87}	{0.40,0.84}
σ_i	0.002	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020
	(0.002)	(0.0018,0.0022)	(0.0017,0.0022)	(0.0018,0.0022)	(0.0018,0.0022)	(0.0017,0.0022)	(0.0017,0.0022)
	IGam	{0.07,0.83}	{0.08,0.82}	{0.07,0.89}	{0.08,0.88}	{0.07,0.82}	{0.08,0.85}
ϕ_π	2.00	1.93	1.85	1.95	1.87	1.92	1.90
	(0.25)	(1.78,2.11)	(1.59,2.08)	(1.78,2.10)	(1.60,2.06)	(1.74,2.05)	(1.74,2.05)
	Norm	{0.07,0.99}	{0.11,0.71}	{0.06,0.99}	{0.10,0.72}	{0.07,1.00}	{0.07,0.95}
ϕ_y	0.50	0.42	0.50	0.42	0.48	0.42	0.42
	(0.25)	(0.27,0.63)	(0.34,0.66)	(0.22,0.62)	(0.31,0.68)	(0.24,0.62)	(0.25,0.58)
	Norm	{0.28,0.87}	{0.22,0.95}	{0.29,0.84}	{0.25,0.92}	{0.29,0.87}	{0.27,0.87}
Σ		**1.08**	**1.20**	**1.04**	**1.25**	**1.14**	**1.42**
Σ^P		**0.57**	**0.76**	**0.58**	**0.78**	**0.59**	**0.77**
$\overline{CR}$		**0.89**	**0.77**	**0.90**	**0.76**	**0.88**	**0.81**

Notes: This table shows the parameter estimates for each estimation method (first column header) and the length of the ZLB regime (second column header). For each parameter, the table reports the average posterior value (first row), (5,95) percentiles (second row) and the NRMSE and the CR (third row, in curly brackets). The second column reports the true parameter value (and prior mean), the prior standard deviation, and the prior distribution. The bottom rows (bold font) report the sum of the NRMSE for all parameters (Σ), the sum of the NRMSE all parameters excluding shock standard deviations (Σ^P), and the average coverage ratio $\overline{CR}$, respectively.

Filter initialization. We conclude this section with a closer look at filter initialization. As shown in Table 2, the IF features a smaller bias in the estimates of the technology shock for datasets without ZLB event. The NRMSE of the standard deviation (σ_z) is smaller than in the PKF. This result demonstrates the interplay of the filter initialization and model approximation. Namely, the IF's suboptimal initialization increases the estimated standard deviation, counterbalancing the downward bias of the piecewise-linear approximation present in the IF and the PKF.

Figure 2 provides an illustrative example of this interplay. Panel (a) shows the path of smoothed shocks obtained with the IF (red) and PKF (blue) for an example selected from the 200 datasets. In this example, the IF's initialization implies a large positive innovation of the technology shock in period one. Panel (b) highlights that the single outlier significantly biases the posterior estimates. It increases the shock's standard deviation (σ_z) and affects other estimates such as φ_p. The bias caused by the first period implies that a pre-sample is a simple remedy when using the IF (blue lines) for this example, as suggested in Section 2.4.

[Figure 2 about here.]

5. Large-scale model applications

To benchmark our approach, we have so far considered models accessible to other estimation methods. However, the PKF expands the reach of nonlinear estimation methods also to large models. Here, we briefly mention the estimation of the Global Multicountry (GM) model of the European Commission (Albonico et al., 2019) with an endogenous occasionally-binding ZLB constraint.[30] In terms of model complexity and specification, the GM model is comparable to other large models used at policy institutions.[31] However, despite the importance of the ZLB for policy prescriptions and observed macroeconomic dynamics, this class of models commonly abstracts from nonlinear estimation.

Estimating the GM model is challenging not only because of its size. It is written in a non-stationary form and features more shocks than observables.[32] In this setup, typical for policy institutions, computational efficiency is paramount. Here, both comparative advantages are crucial: Non-singularity rules out the application of the IF, while the

[30]The model features two regions: the euro area (EA) and the rest of the world (RoW). The EA economy features two types of households, multiple assets, several firm sectors, and a detailed fiscal policy block. The Taylor rule is subject to a ZLB constraint. A rich trade structure (including different commodities) and financial markets link the EA with a simplified RoW block. Albonico et al. (2019) provide more information.

[31]For example, the New Area-Wide Model (NAWM) of the ECB (Christoffel et al., 2008; Karadi et al., 2017) or the SIGMA model from the Federal Reserve Board (Erceg et al., 2006).

[32]A total of 36 shocks matches 33 time series.

model size makes the particle filter too costly.[33] A recent application to the COVID-19 crisis (European Commission, 2020)[p.61-65] exploits the PKF's flexibility. Integrating additional "lockdown" shocks allows the GM model to capture the rapid and enormous contraction in economic activity while maintaining the nonlinearity imposed by the ZLB. For brevity, we do not discuss the estimation results further and only illustrate some economic properties in Appendix D.[34]

6. Implementing the PKF in Dynare

This section provides a practical guide on how to implement the PKF in Dynare. As an example, consider the model outlined in Section 3. It consists of the following four equations, which we repeat here for convenience:

$$C_t + RB_{t-1} = Y_t + B_t \tag{24}$$

$$\ln(Y_t) = \rho \ln(Y_{t-1}) + \sigma\epsilon_t \tag{25}$$

$$C_t^{-\gamma} = \beta R E_t(C_{t+1}^{-\gamma}) + \lambda_t \tag{26}$$

$$\lambda_t(B_t - mY_t) = 0. \tag{27}$$

Implementing this model with the OBC into our Dynare toolkit requires only a few lines of code. First, we declare λ_t as an endogenous variable and introduce a parameter `BORRCON`. Second, the argument `occbin` in line 5 triggers the Occbin/PKF environment, including for the standard `estimation` command. In the last step, we state the conditions for entering and leaving the constrained regime with the keywords `bind` and `relax`. Following the OccBin toolkit (Guerrieri and Iacoviello, 2017), we express these conditions in their linearized form. The keyword `pswitch` declares the switching parameter `BORRCON`, which takes values 0 or 1 depending on whether the OBC is slack or binding.

```
1 var b c lb y ;
2 varexo eps_u ;
3 parameters RHO, BETA, M, R, STD_U, GAMMAC, BORRCON ;
4
5 model(occbin);
6 c = y + b - R*b(-1) ;
7 log(y) = RHO*log(y(-1)) + eps_u ;
8 lb  = 1/c^GAMMAC - BETA*R/c(+1)^GAMMAC ;
9 [pswitch = 'BORRCON', bind =  'lb<-0', relax = 'b>M*y']
```

[33] Indeed, the speed gains of the PKF are critical even if one would modify the shock structure according to the IF's requirements.

[34] In another application of the PKF, Cozzi et al. (2021) estimate an endogenous growth model featuring a rich macroeconomic dataset, stochastic trend productivity, and a ZLB constraint.

```
10 (b -  M*y)*(1 - BORRCON) + lb*BORRCON = 0 ;
11 end;
12
13 ...
14
15 estimation;
```

Listing 1: Sample for code for the borrowing constraint model

Appendix C provides a complete example for the model estimation.

7. Conclusion

This paper has provided a fast, reliable, and widely applicable estimation procedure for DSGE models with occasionally binding constraints. The proposed methodology is conceptually simple. It embeds a fast nonlinear model solution into a diffuse Kalman filter. As a result, it alleviates and overcomes the restrictions of existing methods regarding ease of use, shock structure, model stationarity, missing data, and model size. We have presented several applications to support our conclusions.

We hope that the PKF will prove useful for applied researchers and policymakers interested in models with OBCs. Of course, there is not one method suitable for all purposes. Applications, where uncertainty and precautionary motives are central, may require computationally more expensive methods. However, even in these cases, the PKF provides a user-friendly benchmark approximation.

References

Adjemian, S., H. Bastani, M. Juillard, F. Karamé, J. Maih, F. Mihoubi, G. Perendia, J. Pfeifer, M. Ratto, and S. Villemot (2011). Dynare: Reference Manual Version 4. Dynare Working Papers 1, CEPREMAP.

Albonico, A., L. Calais, R. Cardani, O. Croitorov, F. Ferroni, M. Giovannini, S. Hohberger, B. Pataracchia, F. Pericoli, P. Pfeiffer, R. Raciborski, M. Ratto, W. Roeger, and L. Vogel (2019). The Global Multi-Country Model (GM): An estimated DSGE model for euro area countries. Working Papers 102, European Economy Discussion Papers.

Aruoba, S. B., P. Cuba-Borda, K. Higa-Flores, F. Schorfheide, and S. Villalvazo (2020). Piecewise-linear approximations and filtering for dsge models with occasionally binding constraints.

Atkinson, T., A. W. Richter, and N. A. Throckmorton (2019). The zero lower bound and estimation accuracy. *Journal of Monetary Economics*.

Bielecki, M., M. Brzoza-Brzezina, and M. Kolasa (2019). Comment on "the limits of forward guidance" by jeffrey r. campbell, filippo ferroni, jonas d. m. fisher and leonardo melosi. *Journal of Monetary Economics 108*, 135 – 139. "Central Bank Communications:From Mystery to Transparency"May 23-24, 2019Annual Research Conference ofthe National Bank of UkraineOrganized in cooperation withNarodowy Bank Polski.

Boehl, G. (2021). Solution, filtering and estimation of models with the ZLB. *mimeo*.

Burgert, M., P. Pfeiffer, and W. Roeger (2021). Fiscal policy in a monetary union with downward nominal wage rigidity. *mimeo*.

Campbell, J. R., F. Ferroni, J. D. Fisher, and L. Melosi (2019). The limits of forward guidance. *Journal of Monetary Economics 108*, 118 – 134. "Central Bank Communications:From Mystery to Transparency"May 23-24, 2019Annual Research Conference ofthe National Bank of UkraineOrganized in cooperation withNarodowy Bank Polski.

Christiano, L., M. Eichenbaum, and S. Rebelo (2011). When is the government spending multiplier large? *Journal of Political Economy 119*(1), 78–121.

Christoffel, K., G. Coenen, and A. Warne (2008). The new area-wide model of the euro area: a micro-founded open-economy model for forecasting and policy analysis. Working Paper Series 944, European Central Bank.

Cozzi, G., B. Patarrachia, P. Pfeiffer, and M. Ratto (2021). How much Keynes and how much Schumpeter? *European Economic Review 136*.

Cuba-Borda, P., L. Guerrieri, M. Iacoviello, and M. Zhong (2019). Likelihood evaluation of models with occasionally binding constraints. *Journal of Applied Econometrics n/a*(n/a).

Den Haan, W. and T. Drechsel (2020). Agnostic structural disturbances (asds): detecting and reducing misspecification in empirical macroeconomic models. *Journal of Monetary Economics*.

Erceg, C. J., L. Guerrieri, and C. Gust (2006). SIGMA: A New Open Economy Model for Policy Analysis. *International Journal of Central Banking 2*(1).

European Commission (2020). European economic forecast. autumn 2020. *Institutional Paper* (136).

Fair, R. and J. Taylor (1983). Solution and maximum likelihood estimation of dynamic nonlinear rational expectations models. *Econometrica 51*(4), 1169–85.

Farmer, L. E. (2021). The discretization filter: A simple way to estimate nonlinear state space models. *Quantitative Economics 12*(1), 41–76.

Ferroni, F., S. Grassi, and M. A. León-Ledesma (2019). Selecting structural innovations in dsge models. *Journal of Applied Econometrics 34*(2), 205–220.

Guerrieri, L. and M. Iacoviello (2015). Occbin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics 70*, 22–38.

Guerrieri, L. and M. Iacoviello (2017). Collateral constraints and macroeconomic asymmetries. *Journal of Monetary Economics 90*, 28–49.

Gust, C., E. Herbst, D. López-Salido, and M. E. Smith (2017). The empirical implications of the interest-rate lower bound. *American Economic Review 107*(7), 1971–2006.

Hirose, Y. and A. Inoue (2016). The zero lower bound and parameter bias in an estimated dsge model. *Journal of Applied Econometrics 31*(4), 630–651.

Holden, T. D. (2019). Existence and uniqueness of solutions to dynamic models with occasionally binding constraints. EconStor Preprints 144570, ZBW - Leibniz Information Centre for Economics.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *ASME Transactions Journal Basic Engineering* (83-D), 95–108.

Karadi, P., S. Schmidt, and A. Warne (2017). The new area-wide model ii: An updated version of the ecb's micro-founded model for forecasting and policy analysis with a financial sector. mimeo, European Central Bank.

Koopman, S. J. and J. Durbin (2003). Filtering and smoothing of state vector for diffuse state-space models. *Journal of Time Series Analysis 24*(1), 85–98.

Kulish, M., J. Morley, and T. Robinson (2017). Estimating DSGE models with zero interest rate policy. *Journal of Monetary Economics 88*(C), 35–49.

Maih, J. (2015). Efficient perturbation methods for solving regime-switching DSGE models. Working Paper 2015/01, Norges Bank.

OECD (2021). Short-term interest rates forecast (indicator). doi: 10.1787/9446e151-en. (Accessed on 23 January 2021).

Rotemberg, J. J. (1982). Monopolistic Price Adjustment and Aggregate Output. *Review of Economic Studies 49*(4), 517–531.

Smets, F. and R. Wouters (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association 1*(5), 1123–1175.

Appendix A. Model details

Appendix A.1. Equilibrium conditions in the consumption choice model

Cuba-Borda et al. (2019) describe the microfounded household maximization problem. Here, we state the necessary conditions for an equilibrium which are expressed as a system of four equations in four unknowns $\{C_t, B_t, \lambda_t, Y_t\}$ with parameter vector $\theta = [R, m, \sigma, \gamma, \beta]$ and exogenous innovation ϵ_t:

$$C_t + RB_{t-1} = Y_t + B_t \tag{1}$$

$$\ln Y_t = \rho \ln Y_{t-1} + \sigma \epsilon_t \tag{2}$$

$$C_t^{-\gamma} = \beta R E_t[C_{t+1}^{-\gamma}] + \lambda_t \tag{3}$$

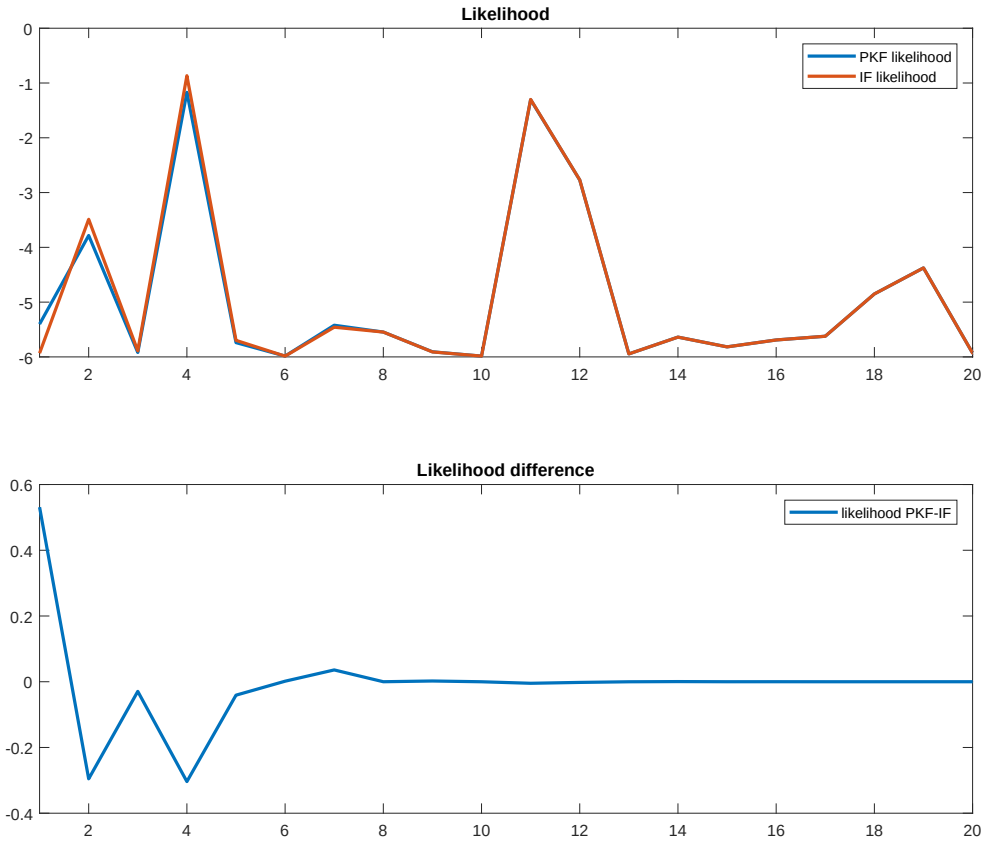
$$\lambda_t(B_t - mY_t) = 0, \tag{4}$$

where $E_t[]$ denotes the expectation operator.

Appendix B. Initialization of IF and PKF in the Application 1

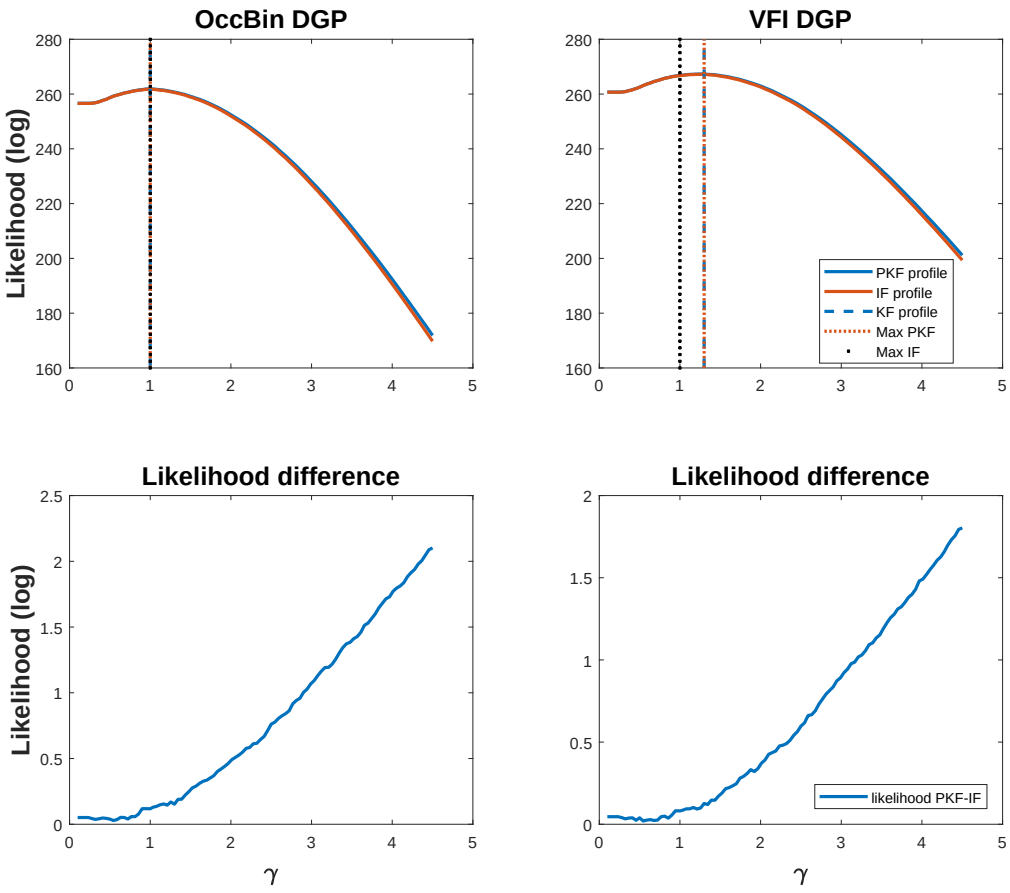
As an illustration, in Figure B.1 we show the likelihood values evaluated at the true value for γ by the two filters for the first 20 data points associated with a simulation with value function iteration as DGP. Indeed the difference converges to zero as we move away from the first observation. The likelihood profile of the PKF is always higher than the one resulting from the IF (see the bottom panels of Figure B.2). This finding suggests that allowing for latent initial states improves upon (particularly with small samples) the assumption that the initial state corresponds to the steady state as done in the IF.

Figure B.1: Likelihood differences due to initialization.



Notes: Top panel: Likelihood values at $\gamma = 1$ with PKF (blue) and IF (red) for $T = 1, .., 20$ using value function iteration DGP. Bottom panel: difference in likelihood values.

Figure B.2: Likelihood profile: PKF, IF. $T = 100$



Notes: Top panel: Likelihood values for different values of γ computed with PKF (blue) and IF (red) using OccBin DGP (left) and VFI DGP (right). Bottom panel: difference in likelihood values computed with PKF and IF using OccBin DGP (left) and VFI DGP (right).

Appendix B.1. Medium-scale New Keynesian Model

The de-trended equilibrium conditions are taken from Atkinson et al. (2019).

Setup. Preferences:

$$E_0 \sum_{t=0}^{\infty} \beta^t [\log(c_t - h c_{t-1}^a) - \chi \int_0^1 n_t(\ell)^{1+\eta} d\ell/(1+\eta)]$$

Constraints:

$$c_t + x_t + b_t/(s_t i_t) + u_t k_{t-1} + \tfrac{\varphi_w}{2} \int_0^1 (w_t^g(\ell) - 1)^2 y_t^f \, d\ell = \int_0^1 w_t(\ell) n_t(\ell) d\ell + r_t^k v_t k_{t-1} + b_{t-1}/\pi_t + d_t$$

$$k_t = (1-\delta) k_{t-1} + x_t (1 - \nu (x_t^g - 1)^2/2)$$

$$x_t^g = x_t/(\bar{g} x_{t-1})$$

$$w_t^g(\ell) = \pi_t w_t(\ell)/(\bar{\pi} \bar{g} w_{t-1}(\ell))$$

$$n_t(\ell) = (w_t(\ell)/w_t)^{-\theta_w} n_t$$

$$u_t = \bar{r}^k (\exp(\sigma_v (v_t - 1)) - 1)/\sigma_v$$

De-trended Equilibrium System.

$$r_t^k = \bar{r}^k \exp(\sigma_\upsilon(\upsilon_t - 1)) \tag{1}$$
$$\tilde{y}_t = (\upsilon_t \tilde{k}_{t-1}/g_t)^\alpha n_t^{1-\alpha} \tag{2}$$
$$u_t = \bar{r}^k(\exp(\sigma_\upsilon(\upsilon_t - 1)) - 1)/\sigma_\upsilon \tag{3}$$
$$r_t^k = \alpha m c_t g_t \tilde{y}_t/(\upsilon_t \tilde{k}_{t-1}) \tag{4}$$
$$\tilde{w}_t = (1-\alpha) m c_t \tilde{y}_t/n_t \tag{5}$$
$$w_t^g = \pi_t g_t \tilde{w}_t/(\bar{\pi}\bar{g}\tilde{w}_{t-1}) \tag{6}$$
$$\tilde{y}_t^{gdp} = [1 - \varphi_p(\pi_t/\bar{\pi} - 1)^2/2 - \varphi_w(w_t^g - 1)^2/2]\tilde{y}_t - u_t \tilde{k}_{t-1}/g_t \tag{7}$$
$$y_t^g = g_t \tilde{y}_t^{gdp}/(\bar{g}\tilde{y}_{t-1}^{gdp}) \tag{8}$$
$$i_t^n = (i_{t-1}^n)^{\rho_i}(\bar{\imath}(\pi_t/\bar{\pi})^{\phi_\pi}(y_t^g)^{\phi_y})^{1-\rho_i} \exp(\sigma_i \varepsilon_{i,t}) \tag{9}$$
$$i_t = \max\{1, i_t^n\} \tag{10}$$
$$\tilde{\lambda}_t = \tilde{c}_t - h\tilde{c}_{t-1}/g_t \tag{11}$$
$$\tilde{w}_t^f = \chi n_t^\eta \tilde{\lambda}_t \tag{12}$$
$$\tilde{c}_t + \tilde{x}_t = \tilde{y}_t^{gdp} \tag{13}$$
$$x_t^g = g_t \tilde{x}_t/(\bar{g}\tilde{x}_{t-1}) \tag{14}$$
$$\tilde{k}_t = (1-\delta)(\tilde{k}_{t-1}/g_t) + \tilde{x}_t(1 - \nu(x_t^g - 1)^2/2) \tag{15}$$
$$1 = \beta E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(s_t i_t/(g_{t+1}\pi_{t+1}))] \tag{16}$$
$$q_t = \beta E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(r_{t+1}^k \upsilon_{t+1} - u_{t+1} + (1-\delta)q_{t+1})/g_{t+1}] \tag{17}$$
$$1 = q_t[1 - \nu(x_t^g - 1)^2/2 - \nu(x_t^g - 1)x_t^g] + \beta\nu\bar{g}E_t[q_{t+1}(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(x_{t+1}^g)^2(x_{t+1}^g - 1)/g_{t+1}] \tag{18}$$
$$\varphi_p(\pi_t/\bar{\pi} - 1)(\pi_t/\bar{\pi}) = 1 - \theta_p + \theta_p m c_t + \beta\varphi_p E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(\pi_{t+1}/\bar{\pi} - 1)(\pi_{t+1}/\bar{\pi})(\tilde{y}_{t+1}/\tilde{y}_t)] \tag{19}$$
$$\varphi_w(w_t^g - 1)w_t^g = [(1-\theta_w)\tilde{w}_t + \theta_w \tilde{w}_t^f]n_t/\tilde{y}_t + \beta\varphi_w E_t[(\tilde{\lambda}_t/\tilde{\lambda}_{t+1})(w_{t+1}^g - 1)w_{t+1}^g(\tilde{y}_{t+1}/\tilde{y}_t)] \tag{20}$$
$$g_t = \bar{g} + \sigma_g \varepsilon_{g,t} \tag{21}$$
$$s_t = (1-\rho_s)\bar{s} + \rho_s s_{t-1} + \sigma_s \varepsilon_{s,t} \tag{22}$$
$$mp_t = \sigma_i \varepsilon_{i,t} \tag{23}$$

This is a system of 23 equations in 23 unknowns:
$\{\tilde{c}, n, \tilde{x}, \tilde{k}, \tilde{y}^{gdp}, \tilde{y}, u, \upsilon, w^g, x^g, y^g, \tilde{w}^f, \tilde{w}, r^k, \pi, i, i^n, q, mc, \tilde{\lambda}, g, s, mp,\ \tilde{x}_t \equiv x_t/z_t\}$

Calibrated parameters. Table B.1 presents parameter calibration. Since our parameter choices are identical, we refer to the discussion provided in ART.

Table B.1: Calibrated parameter values

Description	parameter	value
Subjective discount factor	β	0.9949
Frisch Labor Supply Elasticity	$1/\eta$	3
Price Elasticity of Substitution	θ_p	6
Wage Elasticity of Substitution	θ_w	6
Steady-state Labor Hours	$\bar{n}$	1/3
Steady-state Risk Premium	$\bar{s}$	1.0058
Steady-state Growth Rate	$\bar{z}$	1.0034
Steady-state Inflation Rate	$\bar{\pi}$	1.0053
Capital Share of Income	α	0.35
Capital Depreciation Rate	δ	0.025
Investment Adjustment Cost	ν	4

Appendix C. Sample code

The new occbin integration into DYNARE features the PKF when the estimation command is triggered. Setting up the occbin environment requires only minor modifications to a standard Dynare code.

```
// MODEL CODE (Standard Dynare)
var b bnot c ec lb maxlev y ;
varexo eps_u ;
parameters RHO, BETA, M, R, STD_U, GAMMAC, BORRCON ;

BORRCON = 0;
STD_U = 0.01;

// OCCBIN environment
model(occbin);
ec = c(1);
c = y + b - R*b(-1) ;
// define OBC -------------------------
[pswitch = 'BORRCON', bind =  'lb<-lb_ss', relax =  'b>bnot']
// END OBC ----------------------------
(b -  M*y)*(1 - BORRCON) + lb*BORRCON = 0 ;  // complementarity lb = 0;
bnot =  M*y ; //
lb  = 1/c^GAMMAC - BETA*R/c(+1)^GAMMAC ;
log(y) = RHO*log(y(-1)) + eps_u ;
maxlev = b-bnot;
end;

shocks;
var eps_u;
stderr STD_U;
end;

steady;
check;
// ESTIMATION (Standard Dynare commands)
varobs c;
estimated_params;
GAMMAC,, 0.01000, 10.0000, NORMAL_PDF, 3, 1;
end;

estimation;
```

Listing 2: Full sample for code for the borrowing constraint model

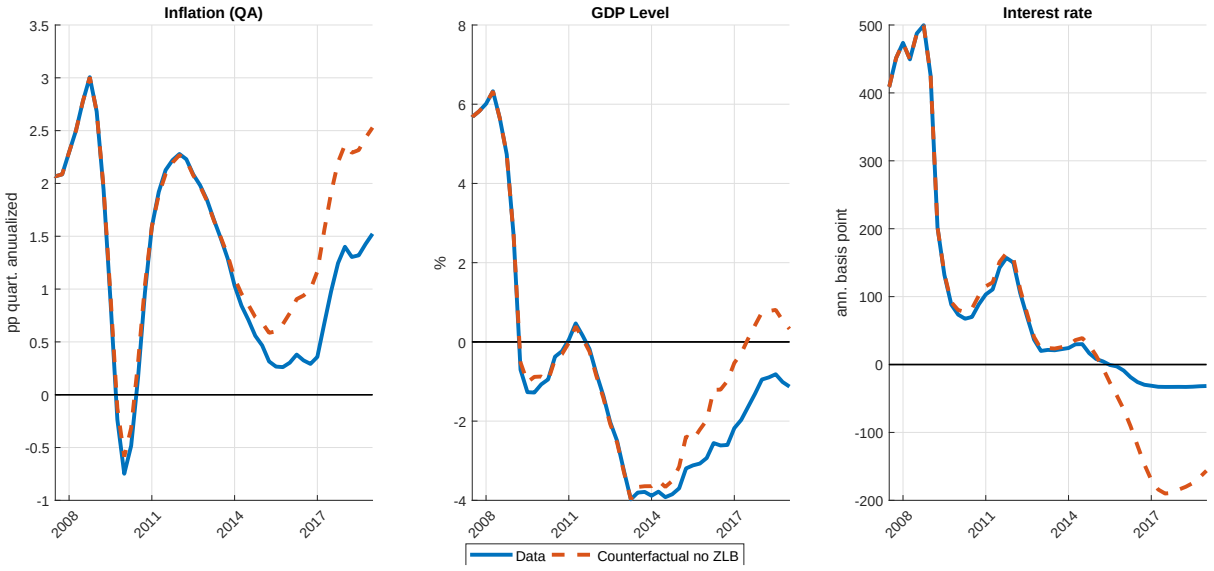
Appendix D. A large-scale model with ZLB

This Appendix illustrates some economic properties of the large-scale GM model (Albonico et al., 2019) with an occasionally binding ZLB, which we estimate using the PKF. Here, we present simulations on the macroeconomic costs of the ZLB. For this purpose, we run two simulations. The first feeds the estimated shocks into the nonlinear model, given our parameter estimates. By construction, these shocks recover the observed time series. The second simulation feeds the same set of shocks into a model without ZLB constraint. This procedure provides counterfactual trajectories of inflation, output, and the interest rate.

Figure D.3 shows that the ZLB entails substantial macroeconomic costs. The estimated model identifies adverse demand shocks as the primary mechanism for low inflation and the ZLB environment.[35] Under ZLB, the central bank cannot accommodate the deflationary shocks. The missing interest rate response amplifies the shock transmission, and the shocks lead to a persistent decline in output and inflation. By contrast, without the ZLB constraint, the monetary policy rule would imply an interest rate of -200 basis points. Without the ZLB constraint, the estimates suggest that inflation would have reached 2.5 percent annually, and output would grow around its long-run trend (before 2020). While abstracting from unconventional monetary policy, this exercise underlines the importance of the ZLB for the current (and future) macroeconomic environment in the EA and other advanced economies.

[35] This finding is in line with those of other estimated models such as Gust et al. (2017).

Figure D.3: Counterfactual experiment

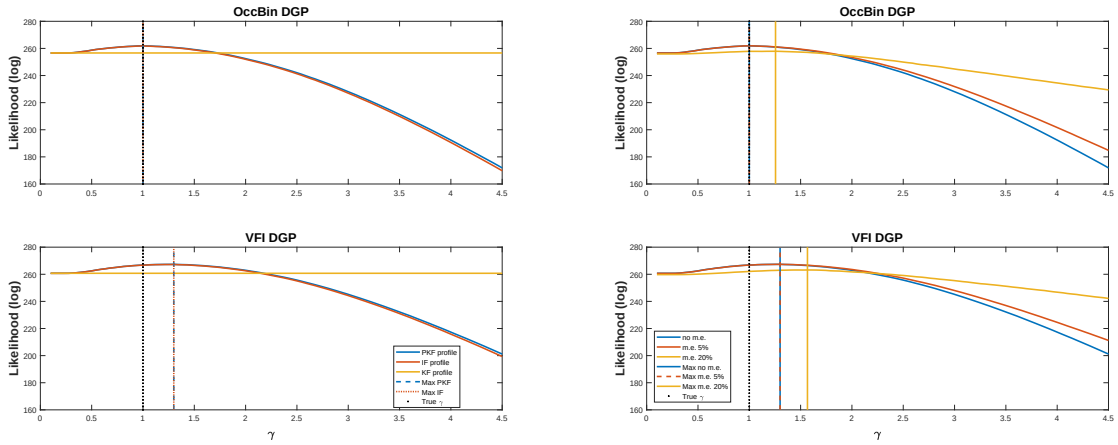


Notes: This figure displays data (blue, solid) and counterfactual simulations without the ZLB constraint (dotted, red). Inflation is measured as the quarterly annualized growth of the consumption deflator. GDP level is reported in percentage deviation from trend, the interest rate is measured in annualized basis points.

Declaration(s) of interest

All authors: None.

Figure 1: Likelihood profiles and estimation methods

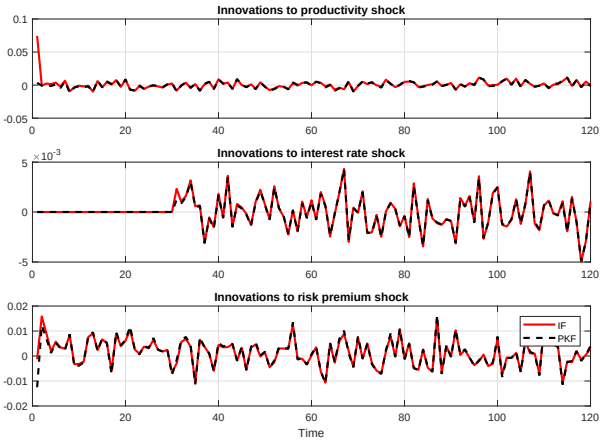


(a) Likelihood profiles for PKF, IF and Lin-KF.

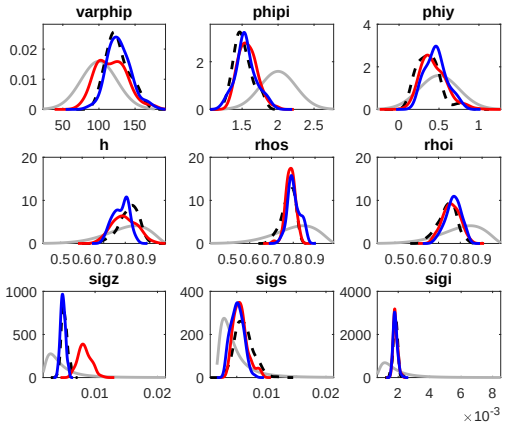
(b) Likelihood profile of the PKF with measurement error.

Notes: The left panel (a) shows likelihood values for different values of γ computed with PKF (blue), IF (red) and linear KF (yellow). The right panel (b) shows likelihood values for different values of γ with no measurement error (blue) and with measurement error equal to 5% (red) and 20% (yellow) of the standard deviation of consumption. Vertical lines indicate the maximum. Top (bottom) panels show estimations based on the OccBin (VFI) DGP.

Figure 2: Initialization effects and estimation results



(a) Smoothed shocks



(b) Posterior densities of different estimators

Notes: Black (red) lines show posterior densities from the PKF (IF) using a selected dataset (98 out of 100). Blue (gray) lines depict the IF estimates with a one period pre-sample (prior distributions).