Conditional Equivalence of Inversion Filter and Kalman filter in Estimating DSGE Models*

Elnura Baiaman kyzy† Hiroyuki Kubota‡ Takeki Sunakawa§

February 2024

Abstract

We show the condition under which the likelihood of a given set of structural parameters computed by the inversion filter (Fair and Taylor, 1983; Guerrieri and Iacoviello, 2017) is equivalent to the likelihood computed by the standard Kalman filter. This condition, however, does not hold when the number of observation variables is large in a medium-scale DSGE model. The difference of likelihoods also results in different estimates of structural parameters in DSGE models by using each of these filters.

*The authors would like to thank Yasuo Hirose and Taisuke Nakata for useful discussions and Shuma Sato for excellent research assistance. Kubota is supported by Grants-in-Aid for Scientific Research (22J12457). Sunakawa is supported by Grants-in-Aid for Scientific Research (21H00704 and 21K01480).

†Hitotsubashi University, E-mail: elnura.bayaman@gmail.com

‡University of California, Los Angeles, E-mail: hkubota@ucla.edu

§Hitotsubashi University, E-mail: takeki.sunakawa@gmail.com

1 Introduction

Estimating nonlinear dynamic stochastic general equilibrium models (DSGE) has been even more important than ever after the 2007-8 global financial crisis and also in the recent COVID-19 crisis, as non-linearities in the economy play an important role in these crises. Several papers have developed methods to estimate non-linear DSGE models (Aruoba et al., 2017; Gust et al., 2017; Plante et al., 2018). However, estimating more empirically relevant models, such as medium-scale New Keynesian DSGE models as in Smets and Wouters (2007), is usually difficult and time-consuming when such non-linearities are explicitly considered in the model. Among others, Guerrieri and Iacoviello (2017) developed a promising method to estimate piecewise-linear DSGE models with occasionally binding constraints by using an inversion filter (Fair and Taylor, 1983). The method can solve such models and evaluate the likelihood of a given set of parameters with less computation time. Therefore, the method can also estimate structural parameters in non-linear DSGE models with these constraints.

In this note, we will show the condition under which there is the equivalence of the inversion filter and the standard Kalman filter in terms of the likelihood of a given set of parameters in linear Gaussian DSGE models. If this condition is nearly satisfied, which is true in small- and medium-scale DSGE models with three observation variables, the inversion filter may provide a reliable estimate of the likelihood and can be very useful to estimate DSGE models even when considering non-linearities as shown in Atkinson et al. (2020); Cuba-Borda et al. (2019). We will demonstrate, however, this condition is not satisfied in a medium-scale DSGE model with five observation variables as in Gust et al. (2017). The likelihood computed by the inversion filter can be different from the likelihood computed by the standard Kalman filter even in linear Gaussian DSGE models. If this is the case, we need a better filtering method such as particle filters.

First we claim that the difference of likelihoods between the inversion filter and the Kalman filter stems from the variance-covariance matrix P of state variables. In particular, we prove that, if the P matrix in the Kalman filter is a zero matrix and the Kalman gain is constant, the likelihoods calculated by the two filtering methods are exactly the same. Otherwise, the likelihoods can be different to the extent how the Kalman gain influences estimates of state variables. Then we conduct Monte Carlo exercises to demonstrate that this can be an issue especially in estimating medium-scale DSGE models. The difference of likelihoods of a given set of parameters can lead to different estimates of the parameters.[1]

[1] Hirose and Sunakawa (2023) confirmed that the two filtering methods yield similar estimation results in a small-scale New Keynesian model with sticky wages using four observation variables (output gap, inflation, the policy rate, and hours worked).

Related Literature Gust et al. (2017) estimated a medium-scale nonlinear model with the ZLB constraint using a projection method and a particle filter to compute the likelihood. The particle filter may provide the most reliable estimation in nonlinear models, although it is very time-consuming and requires the use of supercomputers.[2] Hence, some alternative methods are developed to estimate models with occasionally binding constraints. These methods are potentially promising as they are less time-consuming and can estimate models with laptop computers.

Guerrieri and Iacoviello (2017) first proposed an estimation method using Occbin, a method to solve piecewise-linear models developed by Guerrieri and Iacoviello (2015), and the inversion filter to compute the likelihood. They estimated a medium-scale DSGE model with two occasionally binding constraints, the ZLB constraint and collateral constraint on borrowing, and showed that the both constraints are important to explain the business cycles after the financial crisis.

The most relevant paper to ours is Atkinson et al. (2020), which used a medium-scale nonlinear model with an occasionally binding ZLB constraint as the data generating process (DGP) to discuss usefulness of different filtering methods. Specifically, they solved the model by a projection method and generated a large sample of datasets from the DGP. Then, using such artificial datasets, they estimated the model using three different methods: The first method (NL-PF) solves a fully nonlinear model by a projection method and estimates the model by a particle filter. This method is the most time consuming but may provide the most reliable estimation. The second method (OB-IF) solves a piecewise linear model by Occbin and estimates the model by an inversion filter. The third method (Lin-KF) solves a linearized model ignoring the ZLB and estimates the model by the standard Kalman filter.

Then they evaluated these three methods by looking at if estimated parameters are close enough to their true values of the DGP. They found that the estimated parameters by NL-PF and OB-IF are close to each other. This result implies that estimating the medium-scale model by OB-IF is promising as estimating models by NL-PF is usually time consuming. However, they did not estimate some parameters in the medium scale model directly, as they use only three observation variables for the purpose of examining model misspecification.[3] They did not use some data such as consumption, investment, hours worked or wages that are usually used in estimating medium-scale DSGE models.

More recently, Boehl and Strobel (2022) and Giovannini et al. (2021) developed new methods to estimate piecewise linear models with occationally binding constraints. However,

[2] Gust et al. (2017) suggested a way to parallelize a bootstrap particle filter with many cores of CPU.

[3] They also estimated a misspecified small-scale model and found that there is a bias due to the model misspecification.

to the best of our knowledge, there are no papers comparing the likelihood of a given set of parameters computed by the inversion filter with the one computed by the standard Kalman filter.[4]

The rest of the paper proceeds as follows. In Section 2, we prove that the conditional equivalence result of likelihoods between the inversion filter and the Kalman filter. In Section 3, we conduct Monte Carlo exercises to show that such an equivalence does not hold especially when the number of observation variables is five in a medium-scale DSGE model. We concludes in Section 4. The details of the log-linearized model is found in Appendix.

2 Relationship between Inversion Filter and Kalman Filter

2.1 Inversion Filter and Kalman Filter

Throughout the paper, we consider the following linear system of equations.

$$
\begin{aligned}
y_t &= A + Bs_t + u_t, \quad u_t \sim N(0, H) \\
s_t &= \Phi s_{t-1} + Re_t, \quad e_t \sim N(0, S_e)
\end{aligned}
$$

for $t = 1, \ldots, T$. In the above system, y_t is an $(l \times 1)$ vector of observable variables, s_t is an $(n \times 1)$ vector of state variables, e_t is an $(m \times 1)$ vector of exogenous structural shocks, and u_t is an $(l \times 1)$ vector of exogenous measurement errors. The first equation is the observation equation, which links the observation variables to the state variables. The second equation is the state equation, which can be obtained by solving the rational expectation equilibrium of a (log-)linearized DSGE model. We assume $H = 0$ and thus $u_t = 0$ for all $t > 0$ in this section.

In applying the inversion filter by Guerrieri and Iacoviello (2017), taking the initial state variables s_0 as given, we solve

$$y_t = A + B\Phi s_{t-1} + BRe_t$$

for $e_t = (BR)^{-1} (y_t - (A + B\Phi s_{t-1}))$, where we assume $l = m$ and BR is invertible. Then

[4] Herbst and Schorfheide (2015) does such a comparison between the particle filter and the Kalman filter and concludes that the likelihood computed by a particle filter can be very similar to the likelihood computed by the Kalman filter even in the medium-scale DSGE model of Smets and Wouters (2007) with seven observed variables.

we have

$$s_t = \Phi s_{t-1} + R(BR)^{-1}\left(y_t - (A + B\Phi s_{t-1})\right). \tag{1}$$

That is, taking (y_t, s_{t-1}) as given, we can solve for e_t and thus compute the *exact* values of s_t for $t = 1, \ldots, T$.

In applying the standard Kalman filter, taking the initial values of the mean and variance of state variables, $s_{0|0}$ and $P_{0|0}$, as given, we compute

$$\begin{aligned} s_{t|t-1} &= \Phi s_{t-1|t-1}, \\ P_{t|t-1} &= \Phi P_{t-1|t-1}\Phi' + RS_eR' \end{aligned}$$

for $t = 1, \ldots, T$, where $s_{t|t-1} = \mathbb{E}\left(s_t \mid y_1, \ldots, y_{t-1}\right)$ and $P_{t|t-1} = \mathbb{E}\left(P_t \mid y_1, \ldots, y_{t-1}\right)$ are conditional forecasts of s_t and P_t, taking the information set up to period $t-1$, $(y_1, \ldots, y_{t-1})$, as given. Then we update these forecasts using the information newly available at period t, y_t:

$$\begin{aligned} s_{t|t} &= s_{t|t-1} + K_t\nu_t \\ P_{t|t} &= P_{t|t-1} + K_tBP_{t|t-1} \end{aligned} \tag{2}$$

where

$$\nu_t = y_t - A - Bs_{t|t-1}$$

is a vector of ex-ante measurement errors, which is not necessarily equal to zero as we do not know the exact values of s_t, and

$$F_{t|t-1} = BP_{t|t-1}B'$$

is the variance and covariance matrix of ν_t and can be used to compute the Kalman gain, $K_t = P_{t|t-1}BF_{t|t-1}^{-1}$.

2.2 Conditional equivalence of Inversion filter and Kalman filter

Proposition 1 *If $P_{0|0}$ in the Kalman filter is the zero matrix, then the Kalman filter and the inversion filter are equivalent.*

Proof. Note that, if $P_{0|0} = 0$, we have

$$\begin{aligned} P_{1|0} &= RS_eR' \\ K_1 &= P_{1|0}B\left(BP_{t|t-1}B'\right)^{-1} \\ &= RS_eR'B\left(BRS_eR'B'\right)^{-1} \\ &= R(BR)^{-1} \\ P_{1|1} &= P_{1|0} + K_{1|0}BP_{1|0} \\ &= RS_eR' + R(BR)^{-1}BRS_eR' \\ &= 0 \end{aligned}$$

Therefore, we recursively have

$$\begin{aligned} &P_{t|t-1} = RS_eR' \\ &K_t = K = R(BR)^{-1} \\ &P_{t|t} = 0 \end{aligned}$$

for $t = 1, \ldots, T$. Then (2) can be written as

$$\begin{aligned} s_{t|t} &= s_{t|t-1} + K_t\nu_t \\ &= \Phi s_{t-1|t-1} + R(BR)^{-1}\left(y_t - A - B\Phi s_{t-1|t-1}\right) \end{aligned}$$

This is equivalent to (1) as $s_{t|t} = s_t$ holds when there is no uncertainty of s_t in the Kalman filter, i.e., $P_{t|t} = 0$. ■

This result also implies that, when the variance-covariance matrix $P_{t|t}$ is close to the zero matrix, the Kalman gain K_t is close to a constant K and the sequence of filtered variables $s_{t|t}$ by the Kalman filter is similar to the sequence of s_t of calculated by the inversion filter.

Likelihood The likelihood computed by the Kalman filter is given by

$$L = -\frac{nT}{2}\ln 2\pi - \frac{1}{2}\sum \ln\left|F_{t|t-1}\right| - \frac{1}{2}\sum \nu_t' F_{t|t-1}^{-1}\nu_t \tag{3}$$

If $P_{0|0} = 0$, we have

$$\begin{aligned} F_{t|t-1} &= BP_{t|t-1}B' \\ &= BRS_eR'B' \end{aligned}$$

and using $e_t = (BR)^{-1}\left(y_t - (A + B\Phi s_{t-1})\right) = (BR)^{-1}\nu_t$, we have

$$\begin{aligned}\nu_t' F_{t|t-1}^{-1}\nu_t &= e_t'(BR)'\left(BRS_eR'B'\right)^{-1}BRe_t \\ &= e_t'S_e^{-1}e_t\end{aligned}$$

Substituting these equations into (3), we have

$$\begin{aligned}L &= -\frac{nT}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}\ln|BRS_eR'B'| - \frac{1}{2}\sum_{t=1}^{T}e_t'S_e^{-1}e_t \\ &= -\frac{nT}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}e_t'S_e^{-1}e_t - \frac{1}{2}\sum_{t=1}^{T}\left(\ln|BR| + \ln|S_e| + \ln|R'B'|\right) \\ &= -\frac{nT}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}e_t'S_e^{-1}e_t - \frac{T}{2}\ln|S_e| - \sum_{t=1}^{T}\ln|BR|\end{aligned}$$

This is the formula for the likelihood function by the inversion filter, which is shown in Guerrieri and Iacoviello (2017), up to a constant.[5]

3 Monte Carlo exercises

In the previous section, we showed the condition under which the inversion filter is equivalent to the standard Kalman filter in terms of filtered variables and likelihoods. In this section, we show that the inversion filter and the Kalman filter yield very different values of the likelihood when the number of observable variables is five in a medium-scale New Keynesian DSGE model studied by Gust et al. (2017) and Atkinson et al. (2020). This difference results in very different posterior distributions inferred by a Bayesian method with each of these filters. However, the difference of likelihoods is negligible when the number of observable variables is three, and therefore the posterior distributions are very similar, which is consistent with the finding by Atkinson et al. (2020).

In conducting our Monte Carlo exercise, we do the following procedure: First we estimate the posterior distribution of parameters by a Bayesian method with the standard Kalman filter using actual data. Then we use the mean of the posterior distribution for the DGP to generate a set of simulated data. We estimate the posterior distribution with each of the Kalman filter and the inversion filter using the simulated data. We are particularly interested in whether the posterior means estimated with each of the filters using the simulated data recovers the true parameter values of the DGP (i.e., the posterior means estimated using the

[5]See Equation (25) on p. 36 in Guerrieri and Iacoviello (2017).

actual data) or not.

3.1 Data generating process

To obtain the parameter values of the DGP, we estimate a log-linearized medium-scale New Keynesian model. The details of the log-linearized model is in Appendix. In estimating the model, we use five time series of quarterly U.S. data, output growth, consumption growth, investment growth, inflation, and the interest rate, as in Gust et al. (2017). The sample period is from 1983:1 to 2007:4.[6]

We estimate the model by a random-walk Metropolis-Hastings (RWMH) algorithm to sample from the posterior distribution of parameters as a Markov chain. In doing so, we calculate the likelihood of a set of parameters at each draw using the standard Kalman filter.

Following Gust et al. (2017), for each of the observation variables, we set the variance of its measurement error to 25 percent of its total variance over the sample period. Note that we set the measurement errors to zero in Monte Carlo exercises. Our results are robust to setting the measurement errors to zero in estimating the DGP.[7]

Table 1 shows the prior distribution and the posterior distribution estimated with actual data. The choice of estimated parameters and their prior distributions basically follows Gust et al. (2017). As a whole, the posterior distribution is in line with those estimated in Gust et al. (2017) with their linearized model. The mean of the posterior distribution is used as the parameter values for the DGP in Monte Carlo exercises as explained next.

[6]While the sample period in Gust et al. (2017) is until 2014:1, we limit our sample to the non-ZLB period as the model does not explicitly take the ZLB constraint into account.

[7]Note that we use only five observation variables in the present paper. The non-equivalence result with five observation variables presented here holds with more observation variables as in standard medium-scale models such as Smets and Wouters (2007) among others.

	Prior Distribution			Posterior Distribution		
	Type	Mean	Std	Mean	90% CI	
$100z^{tr}$	N	0.500	0.030	0.517	0.473	0.561
$100\pi^{tr}$	N	0.620	0.100	0.621	0.531	0.719
$100\bar{r}$	G	0.250	0.100	0.187	0.092	0.277
h	B	0.600	0.100	0.629	0.552	0.712
σ_v	G	5.000	1.000	4.964	3.274	6.451
φ_p	N	100.000	25.000	105.859	68.882	144.163
φ_w	N	3000.000	5000.000	4697.919	1084.389	8643.029
α	N	0.300	0.050	0.243	0.205	0.278
ρ_i	B	0.600	0.200	0.798	0.747	0.849
ϕ_y	N	0.400	0.300	0.963	0.671	1.276
ϕ_π	N	1.700	0.300	2.090	1.670	2.504
η	G	2.000	0.750	1.941	0.797	2.970
ν	G	4.000	1.000	4.369	2.728	5.952
$1 - a_p$	B	0.500	0.150	0.594	0.359	0.824
$1 - a_w$	B	0.500	0.150	0.414	0.207	0.617
ρ_g	B	0.600	0.200	0.606	0.294	0.934
ρ_μ	B	0.600	0.200	0.591	0.409	0.777
ρ_s	B	0.600	0.200	0.876	0.812	0.936
$100\sigma_g$	IG	0.585	Inf	0.301	0.142	0.450
$100\sigma_i$	IG	0.585	Inf	0.169	0.127	0.207
$100\sigma_m$	IG	0.585	Inf	3.975	2.091	5.898
$100\sigma_s$	IG	0.585	Inf	0.201	0.126	0.277
$100\sigma_z$	IG	0.585	Inf	0.587	0.397	0.767

Table 1. Estimation with the U.S. data from 1983:1 to 2007:1

Notes: N stands for the normal distribution, B stands for the Beta distribution, G stands for the Gamma distribution, and IG stands for the Inverse Gamma distribution. CI stands for the credible interval.

3.2 Five observation variables

Having the parameter values of the DGP at hand, now we are ready to conduct our Monte Carlo exercises, which consist of (i) generating simulated datasets by the DGP and (ii) estimating the model with each set of simulated data by a Bayesian method.

Simulation: We use 50 sets of simulated data for Monte Carlo exercise. For each set of simulated data, We generate 350 periods of simulated data: The initial 100 periods are discarded and we set 50 training periods to eliminate the effect of initial values of the state variables. Therefore, the likelihood we compute is based on 200 periods for each dataset.

Estimation: Then we estimate the model with each set of simulated data. We use the prior distribution in Table 1, which is the same as the one we use to estimate the DGP. We compute the likelihood of a given set of parameters by either the standard Kalman filter or the inversion filter. We estimate the model parameters by the standard Bayesian method as we do for estimating the DGP, except that we set the measurement errors to zero and we do not estimate $(z^{tr}, \pi^{tr}, \bar{r})$ but set their values to the true values of the DGP.[8]

Comparing likelihoods Before proceeding to the estimation result, we show the likelihoods computed by each of the Kalman filter and the inversion filter. Figure 1 shows the sequence of likelihood increments by the Kalman filter and that by the inversion filter with a set of simulated data. The left panel is for likelihood increments at each period (including the 50 training periods) of the posterior mean, whereas the right panel is for likelihood increments of the posterior mode. We see that the likelihoods are very different especially at the posterior mean.[9] The sum of likelihood increments (excluding the first 50 training periods) is -200.87 by the Kalman filter, whereas it is 555.28 by the inversion filter. The likelihoods are somewhat closer at the posterior mode. The sum of likelihood increments is -148.30 by the Kalman filter, whereas it is -201.85 by the inversion filter.

This result is due to the proposition in Section 2. The likelihood computed by the

[8]For the Kalman filter, we follow the standard procedure in Dynare. That is, we compute the Hessian at the mode and its inverse is used for the variance-covariance matrix of the proposal distribution. We run the RWMH algorithm with 50,000 draws from the posterior distribution, with the first 25,000 draws being burned off.

In contrast, we follow the following procedure of three stages for the inversion filter as in Atkinson et al. (2020). First, we search for the mode to create an initial variance-covariance matrix for the estimated parameters. The variance-covariance matrix is based on the parameters corresponding to the 90th percentile of the likelihoods from 1,000 draws. We follow this procedure as the objective function is not always differentiable and therefore it is difficult to compute the Hessian at the mode with the inversion filter. Second, we run the RWMH algorithm with 5,000 draws. The first 2,500 draws are burned off and the remaining draws of parameters are used to update the variance-covariance matrix from the mode search. Third, we conduct a final run of the RWMH algorithm with 50,000 draws, with the first 25,000 draws being burned off.

[9]The posterior distribution is based on the Bayesian inference with the inversion filter. [check]

Kalman filter is influenced by the variance-covariance matrix of state variables through the Kalman gain, whereas there is no such an adjustment in calculating the likelihood by the inversion filter. Note that such differences are mitigated when likelihoods are computed at the posterior mode. This is perhaps because the model misspecification is minimized at the mode and there is less uncertainty in inference of state variables in the Kalman filter. That is, the variance-covariance matrix P is closer to the zero matrix.

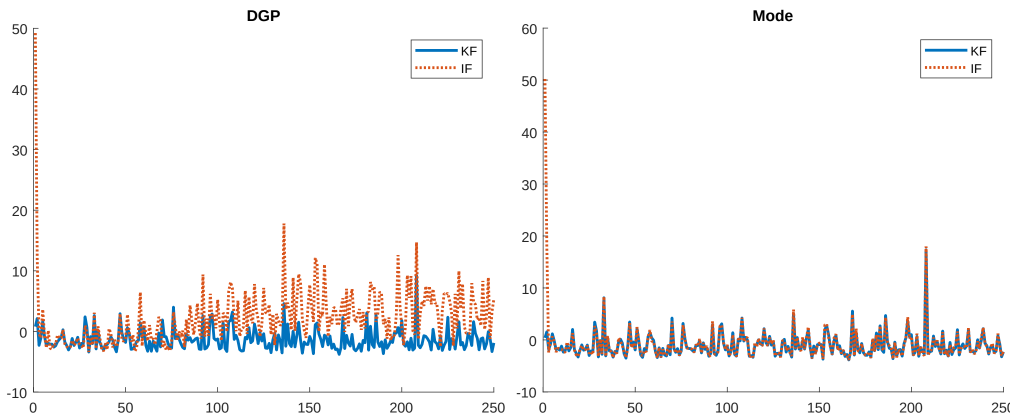

Figure 1. Likelihood increments: Five variables

Posterior distribution Table 2 shows the posterior distribution of the model parameters estimated by each of the standard Kalman filter and the inversion filter. The estimation result is the average of the posterior distributions with 50 simulation datasets. We have two results worth mentioning.

First, we see that the standard Kalman filter can replicate the DGP. The mean of the posterior distribution with the Kalman filter almost coincides with the true parameter values of the DGP. Second, the mean of the posterior distribution with the inversion filter is considerably different from the DGP. Even the 90% confidence interval of the posterior distribution also does not include the DGP for some parameters $(\varphi_p, \varphi_w, \alpha, \phi_i, \phi_y, \phi_\pi, \nu, \rho_g, \rho_\mu, \sigma_m)$. Again, this difference comes from the fact that the likelihoods computed by each of the filters are different with a given set of parameters.

[We evaluate the difference of the parameter values by some methods. See ART.]

3.3 Three observation variables

Next, we conduct the same Monte Carlo exercise with three observation variables, output growth, inflation, and the interest rate, as in Atkinson et al. (2020). In simulation, we use the same parameter values for the DGP as in the case of five observation variables, while

		Kalman Filter			Inversion Filter		
	DGP	Mean	90% CI		Mean	90% CI	
h	0.629	0.613	0.570	0.655	0.611	0.571	0.648
σ_v	4.964	5.040	3.505	6.495	3.969	2.808	5.368
φ_p	105.859	105.116	76.437	133.756	83.703	69.991	98.423
φ_w	4697.919	4608.249	2590.438	6637.061	2834.703	2193.325	3390.309
α	0.243	0.242	0.237	0.247	0.224	0.217	0.231
ϕ_i	0.798	0.763	0.726	0.799	0.729	0.693	0.761
ϕ_y	0.963	0.876	0.673	1.075	0.548	0.486	0.599
ϕ_π	2.090	1.920	1.613	2.213	1.760	1.591	1.927
η	1.941	2.083	1.059	3.115	2.225	1.580	2.921
ν	4.369	4.203	3.015	5.352	2.418	2.040	2.804
$1-a_p$	0.594	0.568	0.498	0.640	0.605	0.539	0.676
$1-a_w$	0.414	0.404	0.315	0.494	0.440	0.362	0.510
ρ_g	0.606	0.594	0.474	0.711	0.991	0.985	0.997
ρ_μ	0.591	0.575	0.504	0.646	0.708	0.656	0.761
ρ_s	0.876	0.855	0.822	0.889	0.850	0.813	0.881
$100\sigma_g$	0.301	0.303	0.273	0.331	0.321	0.294	0.352
$100\sigma_i$	0.169	0.171	0.154	0.188	0.168	0.155	0.183
$100\sigma_m$	3.975	3.924	2.904	4.913	2.125	1.821	2.406
$100\sigma_s$	0.201	0.201	0.154	0.248	0.249	0.205	0.325
$100\sigma_z$	0.587	0.564	0.500	0.628	0.569	0.517	0.626

Table 2. Posterior distribution: Five variables

we set standard deviations of the government spending shock and the marginal efficiency of investment (MEI) shock, σ_g and σ_m, to zero. We estimate a smaller set of parameters $\{h, \varphi_p, \rho_i, \phi_y, \phi_\pi, \rho_s, \sigma_i, \sigma_s, \sigma_z\}$. The prior distributions for these parameters are the same as before. In estimation, the other parameters that are not estimated are fixed at the true parameter values of the DGP.

Figure 2 shows the likelihoods by each of the filters with a set of simulated data as before. In this case of three observation variables, the likelihoods are almost indistinguishable. The sum of likelihood increments at the posterior mean is -330.26 by the Kalman filter, whereas it is -330.22 by the inversion filter. The sum of likelihood increments at the posterior mode is -332.91 by the Kalman filter, whereas it is -333.09 by the inversion filter. As a result, the posterior distributions displayed in Table 3 are also very similar to each other. The mean of the posterior distributions by each of the filters mostly replicate the true parameter values of the DGP. The only exception is found in estimates of ϕ_y in this case. The 90% CI estimated with the Kalman filter $[0.499, 1.014]$ covers the true parameter value 0.963, whereas that estimated with the inversion filter $[0.516, 0.933]$ does not. This result is in line with the one in Atkinson et al. (2020).

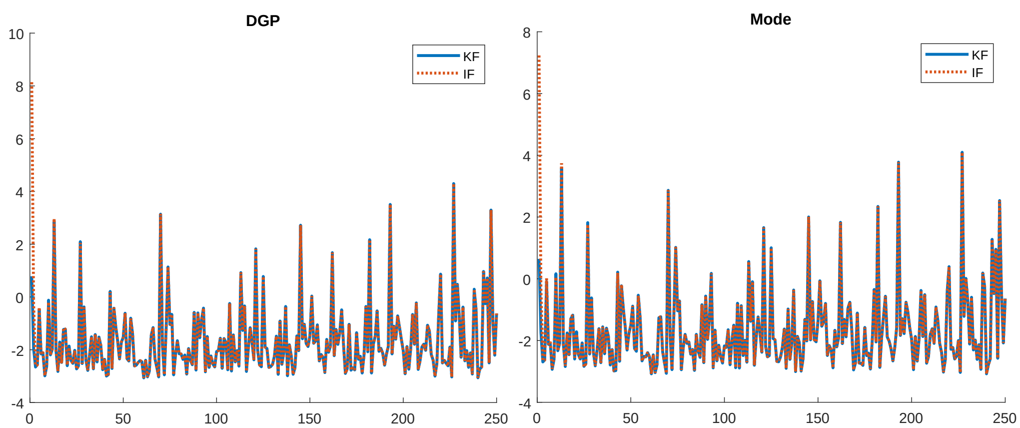

Figure 2. Likelihood increments: Three variables

small-scale model Further, we also conduct the Monte Carlo exercise with a small-scale model with three observation variables. The medium-scale model nests the small-scale model here by setting $\alpha = 0$, $\varphi_w = 0$, and $\theta_w \to \infty$, as in Atkinson et al. (2020). For the common parameters with the medium-scale model, we use the same parameter values of the DGP for simulation and the same prior distribution for estimation as in the case of three observation variables.

Figure 3 shows the likelihoods and Table 4 displays the posterior distributions by each of the filters. The equivalence result with three variables in the medium-scale model also holds

	DGP	KF			IF		
		Mean	90% CI		Mean	90% CI	
h	0.629	0.674	0.586	0.762	0.677	0.592	0.762
φ_p	105.859	92.900	65.082	120.660	94.231	70.823	120.259
ϕ_i	0.798	0.780	0.742	0.820	0.777	0.740	0.812
ϕ_y	0.963	0.760	0.499	1.014	0.722	0.516	0.933
ϕ_π	2.090	1.890	1.571	2.206	1.868	1.568	2.189
ρ_s	0.876	0.848	0.798	0.899	0.853	0.805	0.896
$100\sigma_i$	0.169	0.165	0.150	0.180	0.210	0.152	0.276
$100\sigma_s$	0.201	0.244	0.155	0.334	0.310	0.182	0.477
$100\sigma_z$	0.587	0.555	0.479	0.629	0.519	0.425	0.637

Table 3. Posterior distribution: Three variables

in the small-scale model. The sum of likelihood increments at the posterior mean is -159.38 by the Kalman filter, whereas it is -159.37 by the inversion filter. The sum of likelihood increments at the posterior mode is -162.05 by the Kalman filter, whereas it is -163.17 by the inversion filter. In this case, however, the 90% CI estimated with neither the Kalman filter nor the invertion filter covers the true parameter value of ϕ_y.

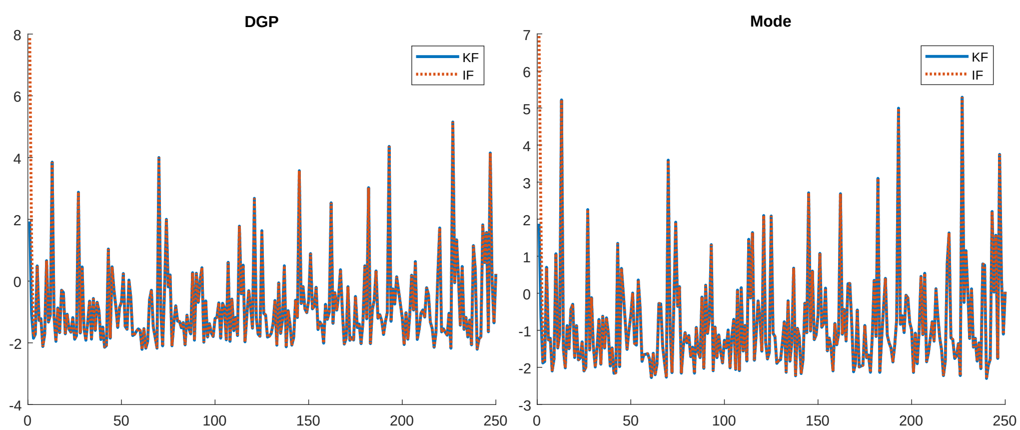

Figure 3. Likelihood increments: Three variables in a small-scale model

4 Conclusion

In this note, we demonstrated that, if the variance-covariance matrix of state variables is a zero matrix, the likelihood computed by the inversion filter of Guerrieri and Iacoviello (2017) is exactly the same as the one computed by the standard Kalman filter. Such an equivalence result, however, does not hold when the number of observation variables is large. We showed

		KF			IF		
	DGP	Mean	90% CI		Mean	90% CI	
h	0.629	0.672	0.611	0.737	0.665	0.600	0.726
φ_p	105.859	98.507	72.527	124.357	93.178	69.267	119.777
ϕ_i	0.798	0.785	0.748	0.822	0.776	0.739	0.810
ϕ_y	0.963	0.659	0.389	0.931	0.587	0.355	0.828
ϕ_π	2.090	1.908	1.605	2.205	1.863	1.628	2.113
ρ_s	0.876	0.837	0.779	0.895	0.836	0.778	0.890
$100\sigma_i$	0.169	0.165	0.150	0.180	0.168	0.153	0.183
$100\sigma_s$	0.201	0.249	0.162	0.331	0.239	0.170	0.329
$100\sigma_z$	0.587	0.588	0.521	0.653	0.591	0.526	0.661

Table 4. Posterior distribution: Three variables in a small-scale model

that this is typically the case in estimating a medium-scale DSGE model used in Gust et al. (2017). Our results cast doubt on using the inversion filter especially for larger models with many observation variables.

References

Aruoba, S. B., Bocola, L., and Schorfheide, F. (2017). Assessing DSGE Model Nonlinearities. *Journal of Economic Dynamics and Control*, 83:34–54.

Atkinson, T., Richter, A. W., and Throckmorton, N. A. (2020). The Zero Lower Bound and Estimation Accuracy. *Journal of Monetary Economics*, 115:249–264.

Boehl, G. and Strobel, F. (2022). Estimation of DSGE Models with the Effective Lower Bound. *Available at SSRN 4138532*.

Cuba-Borda, P., Guerrieri, L., Iacoviello, M., and Zhong, M. (2019). Likelihood evaluation of models with occasionally binding constraints. *Journal of Applied Econometrics*, 34(7):1073–1085.

Fair, R. C. and Taylor, J. B. (1983). Estimation of solution and maximum likelihood dynamic nonlinear rational expectation models. *Econometrica*, 51(4).

Giovannini, M., Pfeiffer, P., and Ratto, M. (2021). Efficient and robust inference of models with occasionally binding constraints. Technical report, JRC Working Papers in Economics and Finance.

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

Guerrieri, L. and Iacoviello, M. (2017). Collateral Constraints and Macroeconomic Asymmetries. *Journal of Monetary Economics*, 90:28–49.

Gust, C., Herbst, E., López-Salido, D., and Smith, M. E. (2017). The Empirical Implications of the Interest-Rate Lower Bound. *American Economic Review*, 107(7):1971–2006.

Herbst, E. P. and Schorfheide, F. (2015). Bayesian Estimation of DSGE Models. In *Bayesian Estimation of DSGE Models*. Princeton University Press.

Hirose, Y. and Sunakawa, T. (2023). The natural rate of interest in a nonlinear dsge model. *Forthcoming in Journal of International Central Banking*.

Plante, M., Richter, A., and Throckmorton, N. (2018). The zero lower bound and endogenous uncertainty. *Economic Journal*, 128(611):1730–1757.

Smets, F. and Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, 97(3):586–606.

Appendix A Log-linearized model

$$\widehat{r}_t^k = \sigma_\nu \widehat{\nu}_t$$
$$\widehat{y}_t = \alpha\left(\widehat{\nu}_t + \widehat{k}_{t-1} - \widehat{z}_t\right) + (1-\alpha)\,\widehat{n}_t$$
$$\widehat{u}_t = \overline{r}^k \widehat{\nu}_t$$
$$\widehat{r}_t^k = \widehat{\mathrm{mc}}_t + \widehat{y}_t + \widehat{z}_t - \nu_t - \widehat{k}_{t-1}$$
$$\widehat{w}_t = \widehat{\mathrm{mc}}_t + \widehat{y}_t - \widehat{n}_t$$
$$\widehat{w}_t^g = \widehat{\pi}_t + \widehat{w}_t + \widehat{z}_t - \widehat{\pi}_t^w - \widehat{w}_{t-1}$$
$$\widehat{y}_t^{\mathrm{gdp}} = \widehat{y}_t - k_y \bar{z}^{-1}\widehat{u}_t$$
$$\widehat{y}_t^g = \widehat{y}_t^{\mathrm{gdp}} - \widehat{y}_{t-1}^{\mathrm{gdp}} + \widehat{z}_t$$
$$\widehat{i}_t = \rho_i \widehat{i}_{t-1} + (1-\rho_i)\left(\phi_\pi \widehat{\pi}_t + \phi_y \widehat{y}_t^g\right) + \sigma_i \varepsilon_{i,t}$$
$$\left(1 - h\bar{z}^{-1}\right)\left(1 - \beta h \bar{z}^{-1}\right)\widehat{\lambda}_t = -\left(\widehat{c}_t - h\bar{z}^{-1}\left(\widehat{c}_{t-1} - \widehat{z}_t\right)\right) + \beta h \bar{z}^{-1}\left(\widehat{c}_{t+1} - \widehat{z}_{t+1} - h\bar{z}^{-1}\widehat{c}_t\right)$$
$$\widehat{w}_t^f = \eta \widehat{n}_t + \widehat{\lambda}_t$$
$$c_y \widehat{c}_t + x_y \widehat{x}_t + g_y \widehat{g}_t = \widehat{y}_t^{\mathrm{gdp}}$$
$$\widehat{x}_t^g = \widehat{x}_t - \widehat{x}_{t-1} - \widehat{z}_t$$
$$\widehat{k}_t = (1-\delta)\bar{z}^{-1}\left(\widehat{k}_{t-1} - \widehat{z}_t\right) + \left(1 - (1-\delta)\bar{z}^{-1}\right)\left(\widehat{\mu}_t + \widehat{x}_t\right)$$
$$\widehat{\lambda}_t - E_t \widehat{\lambda}_{t+1} + \widehat{s}_t + \widehat{i}_t - E_t \widehat{\pi}_{t+1} - E_t \widehat{z}_{t+1} = 0$$
$$\widehat{q}_t = \widehat{\lambda}_t - E_t \widehat{\lambda}_{t+1} - E_t \widehat{z}_{t+1} + \beta \bar{z}^{-1}\left((\beta^{-1}\bar{z} - (1-\delta))E_t \widehat{r}_{t+1}^k + (1-\delta)\,E_t \widehat{q}_{t+1}\right)$$
$$\nu \widehat{x}_t^g = \widehat{q}_t + \widehat{\mu}_t + \beta \bar{z}^{-1} \nu E_t \widehat{x}_{t+1}^g$$
$$\widehat{\pi}_t - \widehat{\pi}_t^p = \beta\left(E_t \widehat{\pi}_{t+1} - E_t \widehat{\pi}_{t+1}^p\right) + \varphi_p^{-1}\left(\theta_p - 1\right)\widehat{\mathrm{mc}}_t$$
$$\widehat{w}_t^g = \beta E_t \widehat{w}_{t+1}^g + \varphi_w^{-1}\left(\theta_w - 1\right)\omega\left(\widehat{w}_t^f - \widehat{w}_t\right)$$
$$\widehat{z}_t = \sigma_z \varepsilon_{z,t}$$
$$\widehat{s}_t = \rho_s \widehat{s}_{t-1} + \sigma_s \varepsilon_{s,t}$$
$$\widehat{g}_t = \rho_g \widehat{g}_{t-1} + \sigma_g \varepsilon_{g,t}$$
$$\widehat{\mu}_t = \rho_m \widehat{\mu}_{t-1} + \sigma_m \varepsilon_{m,t}$$
$$\widehat{\pi}_t^p = (1-a_p)\,\widehat{\pi}_{t-1}$$
$$\widehat{\pi}_t^w = (1-a_w)\left(\widehat{\pi}_{t-1} + \widehat{z}_t\right)$$

Variables: $\widehat{c}_t, \widehat{n}_t, \widehat{x}_t, \widehat{k}_t, \widehat{y}_t^{\mathrm{gdp}}, \widehat{y}_t, \widehat{u}_t, \widehat{\nu}_t, \widehat{w}_t^g, \widehat{x}_t^g, \widehat{y}_t^g, \widehat{w}_t, \widehat{w}_t^f, \widehat{r}_t^k, \widehat{\pi}_t, \widehat{i}_t, \widehat{q}_t, \widehat{\mathrm{mc}}_t, \widehat{\lambda}_t, \widehat{z}_t, \widehat{s}_t, \widehat{g}_t, \widehat{\mu}_t, \widehat{\pi}_t^p, \widehat{\pi}_t^w$
Shocks: $\varepsilon_{z,t}, \varepsilon_{s,t}, \varepsilon_{i,t}, \varepsilon_{g,t}, \varepsilon_{m,t}$
Calibrated parameters: We calibrate $\delta = 0.025, g_y = 0.2, \theta_p = 6, \theta_w = 6, \chi = 1$ following

Gust et al. (2017). We also set $\bar{r} = 0.1870, z^{tr} = 0.5168, \pi^{tr} = 0.6211$ at the true parameter values of DGP for Monte Carlo exercises.

Steady state: Steady state values $\bar{z}, k_y, c_y, x_y, g_y, \omega = \bar{w}\bar{n}\bar{\lambda}$ in the log-linearized equilibrium conditions are computed by the following equations:

$$
\begin{aligned}
\beta &= \left(1 + \frac{\bar{r}}{100}\right)^{-1} \\
\bar{z} &= 1 + \frac{z^{tr}}{100} \\
\bar{\pi} &= 1 + \frac{\pi^{tr}}{100} \\
\bar{i} &= \beta^{-1}\bar{z}\bar{\pi} \\
\bar{r}^k &= \beta^{-1}\bar{z} - (1-\delta) \\
\text{mc} &= \theta_p^{-1}(\theta_p - 1) \\
k_y &= \alpha \text{mc} z / \bar{r}^k \\
x_y &= k_y \\
c_y &= 1 - x_y - g_y \\
\lambda^y &= \frac{1 - \beta h \bar{z}^{-1}}{c_y(1 - h\bar{z}^{-1})} \\
n_y &= (k_y \bar{z}^{-1})^{\frac{\alpha}{\alpha-1}} \\
\bar{w} &= (1-\alpha)\text{mc} n_y \\
\bar{g} &= (1 - g_y) \\
\bar{w}^f &= \theta_p^{-1}(\theta_p - 1)\bar{w} \\
\bar{y} &= (\chi^{-1} w^f \lambda^y n_y)^{\frac{1}{1+\eta}} \\
\omega &= \bar{w}\bar{n}\bar{\lambda} = \bar{w} n_y \lambda^y
\end{aligned}
$$

where $k_y = \bar{k}/\bar{y}, x_y = \bar{x}/\bar{y}, c_y = \bar{c}/\bar{y}, \lambda^y = \bar{\lambda}\bar{y}, n_y = \bar{n}/\bar{y}$. The variables with bars denote steady state values. The notations are made consistent with the ones in Atkinson et al. (2020) as much as possible.

We have two additional shocks, the government spending shock $\widehat{g}_t$ and the MEI shock $\widehat{\mu}_t$, to the original model specification by Atkinson et al. (2020) as we estimate the model with five observation variables. When we estimate the model with three observation variables, we drop these shocks. We also have internal habit, price indexation, and wage indexation on top of the original specification. Regarding the slope of wage New Keynesian Phillips curve (WNKPC), we adapt the specification by Gust et al. (2017), whereas the slope of WNKPC is given by $\tilde{\omega} = \bar{w}\bar{n}/\bar{y}$ in Atkinson et al. (2020)[10]. We log-linearize the original nonlinear equations, while Atkinson et al. (2020) linearize these equations.

The model is also very similar to the one used in Gust et al. (2017). Gust et al. (2017) has the output gap instead of the output growth in a Taylor-type monetary policy rule.

The observation variables, $\widehat{y}_t^{obs}, \widehat{c}_t^{obs}, \widehat{x}_t^{obs}, \widehat{\pi}_t^{obs}, \widehat{i}_t^{obs}$ are linked to the state variables of the model via the observation equation as follows.

$$\begin{aligned}
\widehat{y}_t^{obs} &= z^{tr} + \widehat{y}_t - \widehat{y}_{t-1} + \widehat{z}_t \\
\widehat{c}_t^{obs} &= z^{tr} + \widehat{c}_t - \widehat{c}_{t-1} + \widehat{z}_t \\
\widehat{x}_t^{obs} &= z^{tr} + \widehat{x}_t - \widehat{x}_{t-1} + \widehat{z}_t \\
\widehat{\pi}_t^{obs} &= \pi^{tr} + \widehat{\pi}_t \\
\widehat{i}_t^{obs} &= z^{tr} + \pi^{tr} + \overline{r} + \widehat{i}_t
\end{aligned}$$

[10]The difference is minor, however, as $\omega/\tilde{\omega} = \bar{\lambda}\bar{y} \approx 1$ holds.