

Likelihood Evaluation of Models with Occasionally Binding Constraints*

Pablo Cuba-Borda Luca Guerrieri Matteo Iacoviello Molin Zhong[†]

April 29, 2019

Abstract

Applied researchers interested in estimating key parameters of DSGE models face an array of choices regarding numerical solution and estimation methods. We focus on the likelihood evaluation of models with occasionally binding constraints. We document how solution approximation errors and likelihood misspecification, related to the treatment of measurement errors, can interact and compound each other.

KEYWORDS: Measurement error, solution error, occasionally binding constraints, particle filter.

JEL CLASSIFICATION: C32, C53, C63.

*The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of anyone else associated with the Federal Reserve System. The authors have no conflict of interest related to the work presented in this article. We thank Fabio Canova and Marco Del Negro (our editors). We are also indebted to Arthur Lewbel and seminar participants at MWM Spring 2018, CEF 2018, LACEA/LAMES 2018, Dallas Fed, and NAWMES 2019. Replication codes are available from <http://www.lguerrieri.com/jae-replication.zip>.

[†]The authors are economists at the Federal Reserve Board.

They can be contacted at: PABLO.A.CUBABORDA@FRB.GOV, LUCA.GUERRIERI@FRB.GOV, MATTEO.IACOVIELLO@FRB.GOV, MOLIN.ZHONG@FRB.GOV.

1 Introduction

Consider the example of a researcher who wishes to interpret data on consumption through the lens of a DSGE model in order to estimate the coefficient of relative risk aversion. Except in a handful of special cases, both the solution and the estimation steps will require the use of numerical approximation techniques that introduce sources of error between the “true” value of the parameter and its estimate.

This paper offers a cautionary example of how solution approximation and estimation errors can interact to complicate inference regarding the parameters of a DSGE model. We examine the likelihood evaluation of a model with occasionally binding constraints. Several authors—see the Handbook of Macroeconomics chapter by [Fernández-Villaverde, Rubio-Ramírez, and Schorfheide \(2016\)](#), as well as the textbook treatment of [Herbst and Schorfheide \(2016\)](#)—have already analyzed estimation and inference issues for nonlinear DSGE models. However, their analyses focused on nonlinearities triggered either by Epstein-Zin preferences or by time-varying volatility of shocks. Occasionally binding constraints (or, equivalently, models with endogenous regime shifts) received less attention, at least until the Global Financial Crisis and the advent of the zero-lower-bound era.¹ In our application, solution approximation errors and errors in specifying the likelihood function can interact and compound each other.²

We base our analysis on a simple model for the choice of consumption and saving subject to a constraint that limits maximum borrowing to a fraction of current income. Our results are certainly conditional on this particular example. Nonetheless, the economic channels of this example apply to a variety of setups in which occasionally binding constraints hinder inter-temporal smoothing and amplify demand and financial shocks.

We consider three solution methods and three paths to specifying the likelihood function for

¹For models with nonstandard preferences, see for instance [van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramírez \(2012\)](#). For models with time-varying volatility of shocks, see for instance [Justiniano and Primiceri \(2008\)](#), [Amisano and Tristani \(2011\)](#), [Fernandez-Villaverde, Guerron-Quintana, and Rubio-Ramirez \(2015\)](#). For models with the zero lower bound, see for instance [Guerrieri and Iacoviello \(2017\)](#), [Gust, Herbst, López-Salido, and Smith \(2017\)](#), [Aruoba, Cuba-Borda, and Schorfheide \(2018\)](#), and [Atkinson, Richter, and Throckmorton \(2019\)](#).

²While we are not aware of other work that has analyzed the interaction between multiple sources of error on the ability to infer the parameters of a dynamic model, others considered the effect of approximation error. For instance, [Fernandez-Villaverde, Rubio-Ramirez, and Santos \(2006\)](#) consider the effects on statistical inference of using an approximated likelihood instead of the exact likelihood but abstracts from measurement error misspecification. [Fernandez-Villaverde and Rubio-Ramirez \(2005\)](#) simulate data with measurement error from a nonlinear real business cycle model and compare the estimation performance of using a Kalman filter with a linearized model versus a particle filter with a nonlinear model. [Canova, Ferroni, and Matthes \(2018\)](#) analyze the effects of estimating a constant parameter model when there is time-variation in the structural parameters.

the model. Throughout our analysis, we assume that the model’s most accurate solution is the data generating process for the observables and that the model only includes primitive shocks internalized by the agents. We use this setup to highlight how solution approximation errors and likelihood specification errors affect inference about structural parameters, and how their interaction is magnified in models with occasionally binding constraints.

The solution methods we consider fall on different points of the trade-off between speed and accuracy. In order of accuracy (and reverse order of speed), they include: (1) a global solution method, based on value function iteration; (2) the OccBin solution that relies on a shooting algorithm subject to linear constraints (Guerrieri & Iacoviello, 2015); and (3) a first-order perturbation method that disregards the occasionally binding constraint. We show that the less accurate a solution method is, the harder it gets to retrieve the parameter values that govern agents’ decisions, and as the quality of the solution deteriorates, some parameters become more difficult to identify.

The alternative approaches to forming the likelihood that we consider offer different degrees of generality and interact in different ways with the approximation errors for the alternative solution methods. Following the approach proposed by Fair and Taylor (1983), we showcase an “inversion filter” that relies on characterizing the likelihood function analytically by inverting the decision rule for the model.³ We also consider a particle filter, a standard approach to forming the likelihood for nonlinear models based on a sequential Monte Carlo approach (Fernández-Villaverde & Rubio-Ramírez, 2007). Finally, in conjunction with the first-order perturbation method, we consider a standard Kalman filter, for instance, see Hamilton (1994). As for the case of solution error, we show that as measurement error is calibrated to be more sizable, the likelihood misspecification becomes more problematic, making it harder to retrieve the parameter values that govern the data generating process (DGP).

In the implementation of the particle filter, it is common to posit that the DGP includes measurement error, and to fix the variance of this error to some constant value.⁴ This assumption may seem to be an innocuous way to get around degeneracy issues when choosing a computationally manageable number of particles. Indeed, if measurement error is part of the DGP and the

³The likelihood just involves a transformation of the probability density function chosen for the innovations to the shock processes. This transformation is given by the combination of a function that selects certain endogenous variables with the model-implied decision rules for those variables. See also Kollmann (2017) for a recent treatment of the inversion filter.

⁴See, for instance, Bocola (2016), van Binsbergen, Fernández-Villaverde, Kojen, and Rubio-Ramírez (2012), Gust, Herbst, López-Salido, and Smith (2017), and Cuba-Borda (2014).

variance of the measurement error is estimated alongside other parameters of interest, the particle filter delivers an unbiased estimate of the likelihood conditional on the model.⁵ However, in our setup—in which the true DGP does not contain measurement error—the misspecification error involved in the particle filter grows with the size of the assumed measurement error.⁶

In particular, we show that measurement error in estimation can amplify model approximation error and that the assumption of measurement error can be just as pernicious as an inaccurate model solution. Intuitively, measurement error makes it more difficult for the econometrician to distinguish between alternative regimes of a model based on limited observations. In turn, this difficulty in correctly identifying the regime may lead to a substantial deterioration in the inference about model parameters. In this sense, our paper complements the results of two strands of the literature. In the context of linearized DSGE models, [Canova, Ferroni, and Matthes \(2014\)](#) show that incorrectly assuming measurement error may distort parameter inference. In the context of nonlinear regression models, measurement error—regardless of whether it is introduced on the right or left-hand side variable—can lead to inconsistent parameter estimates. For instance, [Hausman \(2001\)](#) discusses how a mismeasured left-hand side variable can lead to biased and inconsistent estimators in a large class of nonlinear models, such as binary choice, quantile regression, or duration and hazard models. We show how this intuition applies to models in which occasionally binding constraints lead to regime changes that have to be inferred from the data.

There are many more approaches to forming the likelihood of a model than the ones considered here. Without attempting to offer a complete list, some additional alternatives include the extended Kalman filter, the unscented Kalman filter, and the central difference Kalman filter ([Andreasen, 2013](#)).⁷ Furthermore, it is certainly possible to deploy estimation methods that do not rely on the likelihood of the model.⁸ We have some simple justifications for the parsimonious choice of estimation methods considered here. We find the inversion filter appealing because it allows us to study models in which all stochastic innovations, including measurement error, if desired, are taken into account by the agents in the model. This characteristic is valuable, but we do

⁵When the particle filter is embedded in a Markov Chain Monte Carlo sampler, [Andrieu, Doucet, and Holenstein \(2010\)](#) show that one can sample from the correct posterior distribution of the parameters.

⁶[Canova \(2009\)](#) stresses that using measurement error for estimation can distort inference on otherwise properly identified structural parameters.

⁷[Kollmann \(2015\)](#) also discusses the likelihood evaluation of a model solved with a pruned second-order perturbation that does not rely on the particle filter.

⁸[Ruge-Murcia \(2012\)](#), for example, uses simulated method of moments. See [Fernández-Villaverde, Rubio-Ramírez, and Schorfheide \(2016\)](#) for an overview of alternative solution and estimation methods.

not see how to apply this method to overdetermined models.⁹ The particle filter is certainly more flexible in this last respect, and is a focal example that has received widespread attention for the estimation of nonlinear DSGE models. Finally, the ease of implementation of the Kalman filter and its resilience to the curse of dimensionality probably justify why this method is often applied in practice as a shortcut, even if it cannot encompass the occasionally binding constraints in the model of interest. Some pitfalls of linearization are already well documented.¹⁰ We highlight how this shortcut may come at nontrivial costs for inference purposes.

The rest of the paper proceeds as follows. Section 2 presents the conceptual framework for the analysis. Section 3 discusses the model that we take as the DGP for our Monte Carlo experiments and the model solution details. Section 4 describes the results of those experiments. Section 5 concludes.

2 Conceptual Framework

We consider dynamic stochastic models and their relationship to observed data through the lens of a nonlinear state-space representation. To fix notation, this general representation takes the form:

$$s_t = h(s_{t-1}, \eta_t; \theta), \tag{1}$$

$$y_t = g(s_t; \theta) + \zeta_t, \tag{2}$$

$$\eta_t \sim N(0, \Sigma). \tag{3}$$

$$\zeta_t \sim N(0, \Omega). \tag{4}$$

Equation (1) determines the evolution of the endogenous variables summarized in the vector s_t , η_t is a vector of exogenous stochastic innovations that are normally distributed with mean 0 and covariance matrix given by Σ ; the vector θ includes all other parameters. Equation (2) relates the observations summarized in the vector y_t to the endogenous variables in s_t , subject to white noise measurement error ζ_t with variance Ω . We are interested in characterizing the likelihood function

⁹In exactly determined models, the total number of innovations to exogenous processes, including measurement error, is the same as the number of observed variables. Overdetermined models have more sources of exogenous fluctuations than observed variables.

¹⁰For instance, see [Kim and Kim \(2007\)](#).

of the model conditional on the matrix of observations through time T .

$$\mathcal{L} = \ell(\theta; y_{1:T}). \tag{5}$$

2.1 Model Solution

The function $h(\cdot)$ is determined by the economic model of interest.¹¹ An important class of models, including the example considered in this paper, does not support a closed-form solution. Accordingly, in practice, the function $h(\cdot)$ is also dependent on the numerical method chosen to approximate the solution of the model (and its approximation error). We consider three alternative solution methods:

1. **Value Function Iteration Solution**, as described in [Judd \(1998\)](#) or in [Ljungqvist and Sargent \(2004\)](#), which introduces bounds to and discretizes the support of the state variables; we denote the related solution function $h_{vfi}(\cdot)$.
2. **OccBin Solution**, as described in [Guerrieri and Iacoviello \(2015\)](#); we denote the related solution function $h_o(\cdot)$. This solution captures nonlinearities induced by occasionally binding constraints but ignores precautionary motives induced by the risk of future shocks.
3. **First-Order Perturbation Solution**, as described, for instance, in [Anderson and Moore \(1985\)](#), with the simplifying assumption that all of the constraints of the model always bind.

Section [A.1](#) of the online appendix provides details on our implementation of these methods. Results for the first-order perturbation solution are covered in the online appendix.

2.2 Likelihood Approximation

We consider three approaches to computing the likelihood function.

Inversion Filter. Under certain conditions that include the same number of observed variables in the vector y_t as innovations in the vector η_t plus the vector ζ_t , knowledge of the distributions of η_t and the measurement error ζ_t can be used to characterize the likelihood of y_t conditional on initial values of η_t and ζ_t , by substituting Equation (1) into Equation (2) and inverting the resulting

¹¹Without loss of generality, the function $g(\cdot)$ can be reduced to the role of selecting particular elements of the vector s_t .

combination function to back out the innovations. This approach to forming the conditional likelihood is familiar from the textbook treatment of the analogous problem for a simple ARMA process, as outlined, for instance in Chapter 5 of [Hamilton \(1994\)](#). [Fair and Taylor \(1983\)](#) showcase this approach to estimating nonlinear DSGE models and [Guerrieri and Iacoviello \(2017\)](#) implement it for the case of a medium-scale model with occasionally binding constraints and no measurement error.¹² Notice that this approach is compatible with all three solution methods considered. In particular, the OccBin solution can be represented as a VAR with time-varying coefficients, which is similar to the representation of the solution in [Farmer, Waggoner, and Zha \(2009\)](#) and [Kulish and Pagan \(2017\)](#).¹³ Notice also that in the case of the first-order perturbation solution, the inversion filter and the Kalman filter yield the same likelihood function, as long as the initialization schemes for the state variables coincide.¹⁴

While we do not see how to generalize the inversion filter for overdetermined models, it has the advantage of producing an exact value for the likelihood. Accordingly, we take the inversion filter as the benchmark against which we compare the alternative approaches that we consider.

Particle Filter. As pointed out by [Flury and Shephard \(2011\)](#), the particle filter can be thought of as a modern generalization of the Kalman filter, which is only applicable to linear, state space models. The particle filter is also applicable to nonlinear models. In particular, it does not require an analytical solution of the model of interest to form the likelihood of a given set of observations. It allows researchers that rely on numerical methods to find the solution of a model of interest to characterize the likelihood of a set of observations given the model by numerical simulation. Both the particle and Kalman filters can produce filtered estimates of unobserved states given data and estimates of the one-step-ahead density, which delivers the likelihood via the prediction error decomposition. In the Kalman case all these quantities are exact; in the particle filter case they are simulation-based estimates.

The particle filter applies more generally than the inversion filter. In particular, it can handle

¹²Section 4.3 of [Guerrieri and Iacoviello \(2017\)](#) describes this method in detail and spells out conditions for its application in conjunction with the OccBin solution method. [Amisano and Tristani \(2011\)](#) applied the same approach to form the likelihood of a model subject to exogenous regime switches.

¹³[Kulish and Pagan \(2017\)](#) can continue to use the Kalman filter to form the likelihood function since the variation in the coefficients is purely exogenous in their case. The Kalman filter would suffer from an endogeneity problem in our case, since the variation in the solution coefficients of the OccBin solution will depend more generally on the realization of the endogenous variables for the model.

¹⁴The online appendix offers additional details on the inversion filter.

not just models that are exactly determined, in the sense that the number of stochastic innovations in the vector η_t plus the number of errors in the vector ζ_t matches the number of variables in the vector y_t , but also models that are overdetermined (in which there are more stochastic innovations than observed variables). The typical configuration for the particle filter posits that the DGP includes measurement error for each variable observed. This choice avoids issues of degeneracy associated with a finite number of particles. Nonetheless, while this choice is expedient for statistical purposes, it divorces the information set of the econometrician from the information set of the agents in the model.

Given these considerations, we focus on a DGP that excludes measurement error from the observation equation. In this respect, any small amount of measurement error needed to avoid degeneracy of the particle filter introduces a misspecification error into the likelihood function. A typical approach, in practice, is to assume that measurement error covers a fixed fraction of the variance of the observed variables. We consider alternative values for this fraction, starting from small levels of 5%, nonetheless sufficient to avoid degeneracy for our example, up to a level of 20%, a value commonly chosen for other empirical applications.

Kalman Filter. In conjunction with the first-order perturbation solution, we form the likelihood function using a Kalman filter that, given our DGP, correctly excludes measurement error. In this case, the misspecification error is avoided. Our analysis focuses on the effect of model approximation error on the shape of the likelihood function. Results for this case are presented in the online appendix.

3 An Application: A Consumption Model with an Occasionally Binding Borrowing Constraint

We base our analysis on a simple model for the choice of consumption and saving subject to a constraint that limits maximum borrowing to a fraction of current income. We focus on this model for two reasons. First, the economic intuition for how this model works is remarkably simple. Versions of this model—with its emphasis on consumption smoothing—are the backbone of a large class of richer models in modern macroeconomics (see for instance the treatments in [Deaton 1992](#)

and [Ljungqvist and Sargent 2012](#)). Its economic intuition applies to a variety of setups in which occasionally binding constraints hinder intertemporal smoothing and amplify demand and financial shocks, such as [Mendoza \(2006\)](#), [Bocola \(2016\)](#), and [He and Krishnamurthy \(2011\)](#). Second, the model structure allows for a precautionary saving motive and for nonlinear, kinked decision rules that can be captured only by an extremely accurate global numerical solution.¹⁵ By contrast, the OccBin solution captures the nonlinearity of the consumption function but introduces a small solution error by ignoring precautionary saving motives, whereas a linearized decision rule that assumes that the constraint is always binding introduces even larger solution errors.

3.1 The Model

A consumer maximizes

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\gamma} - 1}{1-\gamma},$$

where γ is the coefficient of relative risk aversion, subject to the budget constraint and to an occasionally binding constraint stating that borrowing B_t cannot exceed a fraction m of income Y_t :

$$C_t + RB_{t-1} = Y_t + B_t, \tag{6}$$

$$B_t \leq mY_t. \tag{7}$$

Above, R denotes the gross interest rate. The discount factor β is assumed to satisfy the restriction that $\beta R < 1$, so that, in the deterministic steady state, the borrowing constraint is binding. Given initial conditions, the impatient household prefers a consumption path that is falling over time, and attains this path by borrowing today. If income is constant, the household will eventually be borrowing constrained and roll its debt over forever, and consumption will settle at a level given by income less the steady-state debt service cost.

¹⁵The model structure implies precautionary saving because of the combination of uncertainty, borrowing constraints, and a concave utility function. Specifically, there are two sources of precautionary saving in the model: the first source comes from the interaction of uncertainty with borrowing constraints, the second source comes from the interaction of uncertainty with concave utility. Both features make the value function concave in income and wealth. Either one would be sufficient to induce precautionary saving. See [Carroll \(2001\)](#) for a discussion.

The log of income follows an AR(1) stochastic process of the form

$$\ln Y_t = \rho \ln Y_{t-1} + \sigma \epsilon_t, \quad (8)$$

where ϵ_t is an exogenous innovation distributed as standard normal, and σ its standard deviation.

Denoting with λ_t the Lagrange multiplier on the borrowing constraint in Equation (7), the necessary conditions for an equilibrium can be expressed as a system of four equations in the four unknowns $\{C_t, B_t, \lambda_t, Y_t\}$. The system includes equation (6), equation (8), together with the consumption Euler equation and the Kuhn-Tucker conditions, given respectively by

$$C_t^{-\gamma} = \beta R E_t (C_{t+1}^{-\gamma}) + \lambda_t \quad (9)$$

$$\lambda_t (B_t - mY_t) = 0. \quad (10)$$

The transitional dynamics of this model depend on the gap between the discount rate and the interest rate, which can be measured as $g = 1/\beta - R$. When the gap is small, the economy can be characterized as switching between two regimes. In the first regime, more likely to apply when income and assets are relatively low, the borrowing constraint binds. In that case, borrowing moves in lockstep with income, and consumption is more volatile than income. In the second regime, more likely to apply when income and assets are relatively high, the borrowing constraint is slack, and current consumption can be high relative to future consumption even if borrowing is below the maximum amount allowed.

3.2 Calibration.

We set $\gamma = 1$, so that utility is logarithmic in consumption. We set the maximum borrowing at one year of income, so that $m = 1$. For the income process, we set $\rho = 0.90$ and $\sigma = 0.01$, so that the standard deviation of $\ln Y$ is 2.5 percent. Finally, we set $R = 1.05$ and $\beta = 0.945$. Under this calibration, the borrowing constraint binds 60 percent, and is slack about 40 percent of the time using the full nonlinear solution (whether the constraint binds and for how many consecutive periods depends on the shocks and on the predetermined values of the endogenous variables).

3.3 Model Solution

Figure 1 shows contours of the model’s policy functions for the different solution algorithms. We focus on contours for the optimal borrowing and optimal consumption chosen by the agent, expressed as a function of income, holding the level of debt at its deterministic steady-state value of 1. For lower-than-average realizations of income, the agent hits the borrowing constraint. In that case, the consumption function is relatively steep, the multiplier on the borrowing constraint is positive, and consumption is very sensitive to changes in income. For higher-than-average income, consumption is sufficiently high today relative to the future that it pays off to save for the future. In that case, the borrowing constraint becomes temporarily slack, the multiplier on the borrowing constraint is zero, and consumption becomes less sensitive to changes in income. Accordingly, the model-implied distribution of consumption is skewed even if the shocks are governed by a symmetric distribution.

We follow Judd (1992) and use the Euler equation residuals in units of consumption to quantify the error in the intertemporal allocation. The policy functions using value function iteration are minimally affected by approximation error, with Euler errors in the order of \$1 per \$100,000 of consumption.¹⁶ Accordingly, we take this solution method as the benchmark against which we compare the errors of the other methods. The errors for the OccBin solution are typically in the order of \$1 per \$1000 of consumption, a modest if nontrivial amount. Finally, in the class of models with occasionally binding constraints that is the focus of this paper, errors for the linear solution method can rise to the substantial amount of \$1 per \$10 of consumption. In turn, this kind of approximation error becomes problematic for estimation purposes.

4 Findings

In our Monte Carlo experiment, the DGP is the model of Section 3 solved with value function iteration. We simulate 100 samples of consumption data. We consider two alternative sample lengths T : 100 and 500 observations.¹⁷ Our DGP does not contain measurement error—or, equivalently, in Equation 4, the variance of measurement error Ω is 0. We are focused on inference regarding the parameter γ , the coefficient of relative risk aversion in the model. In the DGP, $\gamma = 1$. All

¹⁶See Section A.1.1 and Figure A.1 in the online appendix for further details.

¹⁷For each sample of data that we generate, we simulate 2000 observations starting from the deterministic steady state. We take either the last 100 or last 500 observations from this simulation.

other parameters are fixed at their true values as described in the calibration section above.

4.1 Overview

To build some intuition for the results across different draws of our Monte Carlo experiment, we focus first on a single simulated sample of 100 observations. However, the discussion below, and the results in Table 1 summarizing all of our Monte Carlo draws, confirm that the conclusions we illustrate with this single sample carry over systematically to the other samples.

The panels in Figure 2 depicts the likelihood function for this representative sample as the value of γ varies between 0 and 4.5, varying in each panel solution error and/or measurement error. The top left panel focuses on the value function iteration solution and forms the likelihood through the inversion filter. Of note, the likelihood peaks near 1, the true value in the DGP. Because we use enough nodes to render the Euler equation residuals negligible and because the inversion filter avoids the misspecification error in the measurement equation, we take the likelihood function for this case as the true likelihood against which we assess the alternative combinations of solution methods and filters shown in the figure—the solid blue line is replicated in every other panel.

The panels in Figure 2 are arranged so that the left column only considers the likelihood functions that rely on a value function iteration at the solution step, while the right column shows likelihood contours based on the OccBin solution method that introduces some solution error. The top row focuses on the inversion filter that avoids misspecification error. Moving down the figure, the rows below magnify the misspecification error—summarized by the variance of measurement error, expressed as percent of the variance of consumption. Accordingly, the differences across the likelihood contours presented in the left column stem only from misspecification error. Differences in the likelihood contours across each row highlight the effects of solution error. The top row only considers solution error. The other rows show the interaction of solution and misspecification errors.

Each panel allows some headline comparisons. The distance between the peak of each contour and 1, the true value of γ , is indicative of the bias in the point estimates. Moreover, the likelihood contours allow to visualize the extent to which likelihood misspecification and solution error affect the precision of the estimates.

4.2 Solution Approximation

Focusing on the top right panel, the OccBin solution biases the estimate of γ upwards. The solution method ignores precautionary motives and results in a consumption function that, over some regions, is more sensitive to variation in income relative to the consumption function from the accurate value function iteration method. Accordingly, one way to match the observations is through a higher level of risk aversion relative to the DGP, which results in the upward shift in the peak of the likelihood, but not in a substantial flattening of the likelihood contour.

4.3 Likelihood Misspecification

Moving to the first column of the second row, the panel labeled “Minimal Solution Error, 5% Measurement Error” focuses on the effects of measurement error. As expected, measurement error leads to a flattening of the likelihood contour, but in this case, it also leads to an upward bias in the point estimate of γ , unlike the effect of measurement error on the dependent variable in a linear regression framework. The bias is related to the kinked model decision rules used in the underlying DGP.

Figure 3 can be used to illustrate how this bias is related to the occasionally binding constraint in the model. The policy function underlying the DGP is represented by the blue, solid line. Notice that for realizations of the income lower than the average value of 1, the borrowing constraint binds and consumption is approximately linear in income. In turn, for higher than average realizations of income, the borrowing constraint is slack and consumption responds in a more muted way to high realizations of the income process. Accordingly, there is a kink in the consumption function right at the point where the borrowing constraint becomes slack. Notice also that the point where the constraint becomes slack is related to the underlying value of γ . Lower values of γ shift this point to the right.

Values of γ lower than its assumed value of 1 (coinciding with less risk aversion) would cause the consumption function to be too steep and could be reconciled with the observations for consumption only via skewed and, thus, less likely estimates of the income shocks. Values of γ higher than its assumed value of 1 would cause the consumption function to be too flat, and again call for less likely estimates of the income shocks. In sum, the position of the kink and subsequent shape of the consumption function inferred from observations on consumption can influence the

estimates of γ .

For the sake of argument, let's consider a case slightly different from the one we posited, in which the DGP includes measurement error, but the econometrician does not realize it. Adding normally distributed measurement error changes both features of the consumption function: it makes consumption more volatile, and dampens skewness. Accordingly, the econometrician would think that consumption function is consistent with a lower value of γ , such as the one represented by the dashed, red line. Figure A.2 in the online appendix illustrates this case, showing how “too little” assumed measurement error relative to what is embedded in the DGP biases the estimate of γ *downwards*, regardless of the solution error.

Our case is the mirror image of the one described above. The DGP does not include measurement error, but the econometrician assumes that it is part of the DGP. That is, the econometrician sees skewed and asymmetric consumption even after accounting for normally distributed, additive measurement error. Accordingly, the econometrician's estimates of γ are biased *upwards*, and this bias is greater, the greater the fraction of the observed variation incorrectly attributed to the measurement error.¹⁸

Going back to Figure 2, and moving down along the left-hand-side column, one can readily evince that the peak of the likelihood function shifts further to the right as the relative size of the measurement error increases. An additional effect of measurement error misspecification is to flatten out the likelihood function. One can see the extent that measurement error flattens the likelihood by comparing the two alternative likelihood functions, with and without measurement error, as the value of gamma moves away from the true value of 1. Consider, for instance, the bottom-left panel, labelled “Minimal Solution Error, 20% Measurement Error.” As γ moves to 3 along the true likelihood (the solid line labelled “VFI,IF”), the likelihood changes about 40 log points. As γ moves to 3 along the likelihood that is affected by measurement error (the dashed line labelled “VFI,PF 20%”), the change is only 2 log points, a substantial flattening.

¹⁸As a robustness exercise, Appendix A.5 shows that jointly estimating γ and the parameters governing the shock process lessens some of the misspecification of γ , but the bias persists.

4.4 Interaction of Solution Approximation and Likelihood Misspecification

Figure 2 illustrates the interaction between measurement error and misspecification error. The right-hand side panel in the second row, labeled “Some Solution Error, 5% Measurement Error,” can be compared against the panel to its left, which highlights misspecification error, and against the panel just above, which showcases approximation error. When the two sources of error interact, it is readily apparent that the bias is greater than the sum of the biases for each error in isolation. For instance, with 5% measurement error only, the modal estimate is 1.34, with solution error only, the estimate is 1.21, while with both sources of error, the estimate is 1.61. With larger measurement error, the biases compound each other even more.¹⁹

Another way to understand the bias from the approximation error of the OccBin solution is that the inability to capture precautionary motives moves the kink in the consumption function to the right of its true position. So, again, one way to match the observations would be to realign the consumption function produced by OccBin inferring a value of γ greater than the true value of 1 used in the experiment. But as we have seen above, additional shifts to the left for the kink in the consumption function require disproportionately larger increases in γ , which is why the cumulation of the two sources of error, which happen to go in the same direction, results in a sizable magnification of the bias.

4.5 Results for Multiple Monte Carlo Repetitions

The discussion of the Monte Carlo results has focused, so far, on one representative sample. Table 1 summarizes results across the 100 samples we considered. For ease of reference, the table is arranged in a way analogous to Figure 2, but reports results for two sample sizes. To help gauge the bias implied by solution and misspecification error, apart from the mean of the estimates across the Monte Carlo samples, the table reports the average of the percentiles for the true value of γ across the 100 repetitions. With 100 observations, using value function iteration (VFI in the table) with the inversion filter, the average of these percentiles is 51, indicating that the true value is on average near the median estimates of γ implied by the likelihoods across different

¹⁹The compounding effects of different sources of error discussed here apply systematically across the Monte Carlo draws, as can be seen from Table 1 below.

samples. By contrast, when introducing even modest observation error in conjunction with the particle filter (PF ME5 in the table), the true value of γ falls to almost the 20th percentile, on average (consistent with an *upward bias* in the estimates of γ). With measurement error and misspecification error in conjunction with the use of OccBin, the true value of γ falls to as low as the first percentile, on average. The lower half of the table shows that the bias persists with a longer sample of 500 observations.²⁰

Beyond quantifying the bias, the table also offers an assessment of the reliability of confidence intervals. In line with most practitioners, we focus on Bayesian posterior credible sets, constructed with a Uniform prior over the range from 0 through 4.5.²¹ The table reports the average lower bound of 90 percent sets across different samples, the average upper bound, and the frequency at which the true value of γ is included in the sets. While the effective coverage for IF/VFI, at 88, is close to the target value of 90, the effective coverage falls dramatically for other cases. The same upward bias that affects the point estimates of γ shifts the credible sets to the right and affect their coverage.

5 Conclusion

Occasionally binding constraints create challenges for standard numerical solution algorithms as well as for likelihood-based inference methods. We showed that model misspecification related to measurement error can, in a simple model of consumption, flatten the likelihood function and lead to biased estimates of the model parameters. We also showed that this misspecification error can interact with approximation error and that the bias resulting from this interaction is greater than the bias associated with each error in isolation.

Each model presents specific solution and estimation challenges, but some of the results for our simple example are bound to extend to other models. A challenge unique to models with occasionally binding constraints is that if the observations are subject to measurement error, this error will muddle the inference on whether the constraints bind or not. Imposing that the variance of measurement error is a fixed fraction of the variance of each observed variable is likely to result in biased estimates of the frequency at which the constraints bind, and through that channel, bias

²⁰Additionally, Figure A.4 in the online appendix shows sampling distributions of the mode of the likelihood.

²¹We take 20,000 draws from the posterior distribution, using a random walk metropolis algorithm. We use the first 10,000 draws as burn-in.

the estimates of other parameters that influence that frequency.

We find it challenging to explain the differential emphasis of recent empirical applications on reducing solution error to a minimum, while, at the estimation step, including measurement error that agents in the model do not consider. Our intent is to point out that, faced with finite time for solving and estimating a model, researchers might want to re-balance the attention devoted to solution and estimation issues.

We conclude with some practical options:

1. Rely on the particle filter, but use only modest amounts of measurement error in conjunction with a very large number of particles. The trade-off here is that increasing the number of particles may run into the curse of dimensionality even for a modest number of observed variables and parameters to be estimated. Ongoing research is exploring alternative options to address this trade-off. The paper by [Herbst and Schorfheide \(2019\)](#) is a recent example.
2. Use the decision rules from the model to form the likelihood analytically (what we are dubbing inversion filter). This approach comes at the cost of placing restrictions on the number of observed variables and the number of fundamental shocks in the model. Notice that, for linearized models, this restriction is standard following [Smets and Wouters \(2007\)](#).
3. Allow for more structural shocks than observable data series and use the particle filter without assuming measurement error. [Fernández-Villaverde and Rubio-Ramírez \(2007\)](#) provide the theoretical justification for this approach. In particular, [Fernandez-Villaverde et al. \(2015\)](#) apply this strategy to estimate a DSGE model (solved with a second-order perturbation method) with stochastic volatility. This approach, however, may lead to increases in the computational time and might expose the model evaluation to too many degrees of freedom.
4. Use the simulated method of moments to resolve the inference problem for structural parameters. However, this choice would still necessitate a choice of one of the methods in 1) through 3) above if estimates of the unobserved states are of interest.

Admittedly, none of our proposed solutions is universally applicable. Each solution implies different trade-offs, requiring a case-by-case assessment.

References

- Amisano, G., & Tristani, O. (2011). Exact likelihood computation for nonlinear DSGE models with heteroskedastic innovations. *Journal of Economic Dynamics and Control*, 35(12), 2167-2185. doi: 10.1016/j.jedc.2011.08.00
- Anderson, G., & Moore, G. (1985). A linear algebraic procedure for solving linear perfect foresight models. *Economics Letters*, 17(3), 247-252.
- Andreasen, M. M. (2013). Non-linear dsge models and the central difference kalman filter. *Journal of Applied Econometrics*, 28(6), 929–955. doi: 10.1002/jae.2282
- Andrieu, C., Doucet, A., & Holenstein, R. (2010). Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3), 269–342. doi: 10.1111/j.1467-9868.2009.00736.x
- Aruoba, B., Cuba-Borda, P., & Schorfheide, F. (2018). Macroeconomic dynamics near the zlb: A tale of two countries. *The Review of Economic Studies*, 85(1), 87-118. doi: 10.1093/restud/rdx027
- Atkinson, T., Richter, A., & Throckmorton, N. (2019). *The zero lower bound and estimation accuracy* (Tech. Rep.). Federal Reserve Bank of Dallas.
- Bocola, L. (2016). The pass-through of sovereign risk. *Journal of Political Economy*, 124(4), 879–926.
- Canova, F. (2009). How much structure in empirical models? In *Palgrave handbook of econometrics* (pp. 68–97). Springer.
- Canova, F., Ferroni, F., & Matthes, C. (2014). Choosing the variables to estimate singular dsge models. *Journal of Applied Econometrics*, 29(7), 1099–1117. doi: 10.1002/jae.2414
- Canova, F., Ferroni, F., & Matthes, C. (2018). *Detecting and Analyzing the Effects of Time-Varying Parameters in DSGE Models* (Tech. Rep.).
- Carroll, C. D. (2001). A theory of the consumption function, with and without liquidity constraints. *Journal of Economic perspectives*, 15(3), 23–45.
- Cuba-Borda, P. (2014). What explains the great recession and the slow recovery? *Manuscript, University of Maryland*.
- Deaton, A. (1992). *Understanding consumption*. Oxford University Press.
- Fair, R. C., & Taylor, J. B. (1983). Solution and Maximum Likelihood Estimation of Dynamic

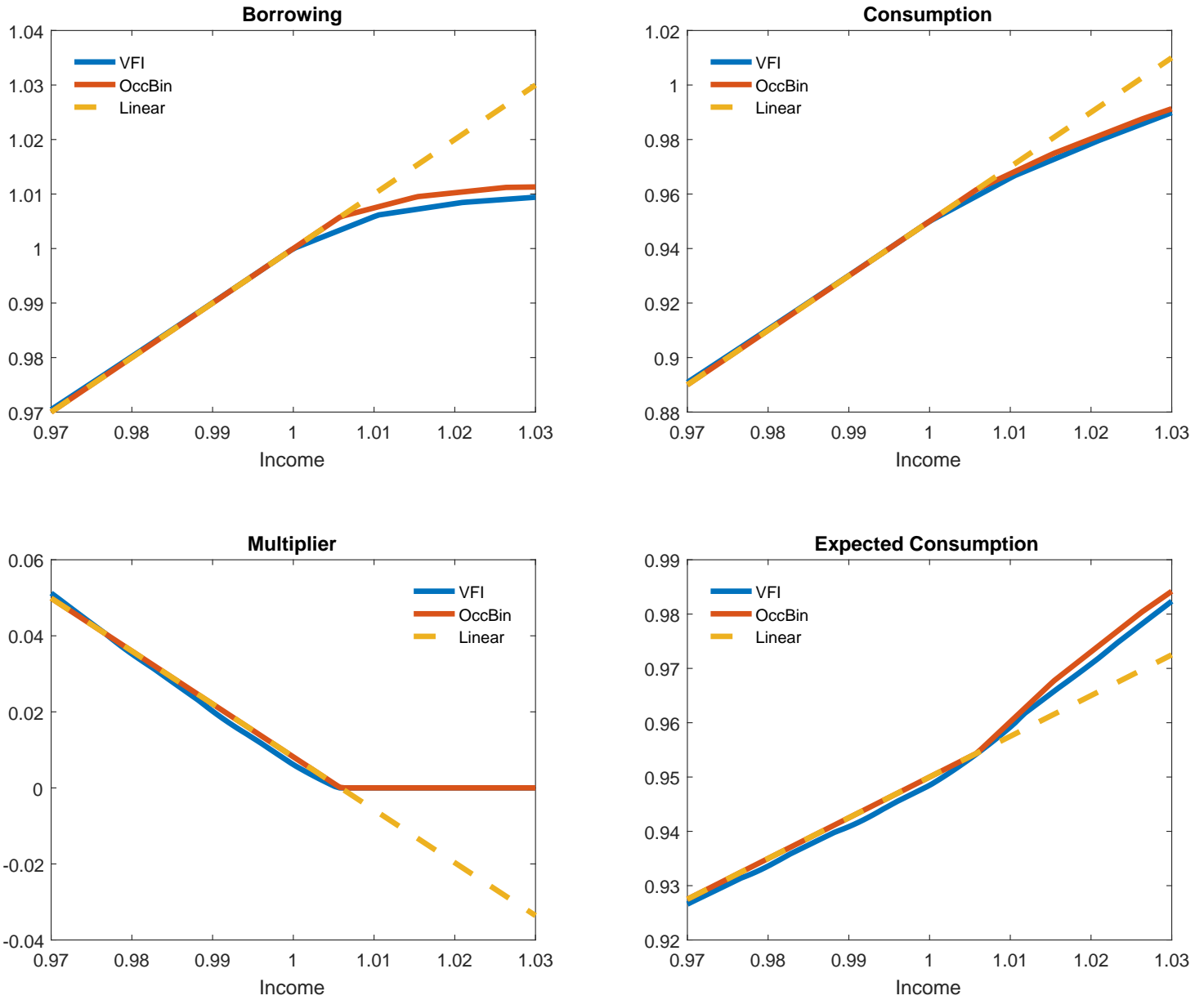
- Nonlinear Rational Expectations Models. *Econometrica*, 51(4), 1169-1185.
- Farmer, R. E., Waggoner, D. F., & Zha, T. (2009). Understanding Markov-switching rational expectations models. *Journal of Economic Theory*, 144(5), 1849-1867.
- Fernandez-Villaverde, J., Guerron-Quintana, P., & Rubio-Ramirez, J. F. (2015). Estimating dynamic equilibrium models with stochastic volatility. *Journal of Econometrics*, 185(1), 216-229. doi: 10.1016/j.jeconom.2014.08
- Fernandez-Villaverde, J., & Rubio-Ramirez, J. F. (2005). Estimating dynamic equilibrium economies: linear versus nonlinear likelihood. *Journal of Applied Econometrics*, 20(7), 891-910. doi: 10.1002/jae.814
- Fernández-Villaverde, J., & Rubio-Ramírez, J. F. (2007). Estimating macroeconomic models: A likelihood approach. *The Review of Economic Studies*, 74(4), 1059–1087.
- Fernandez-Villaverde, J., Rubio-Ramirez, J. F., & Santos, M. S. (2006, January). Convergence Properties of the Likelihood of Computed Dynamic Models. *Econometrica*, 74(1), 93-119.
- Fernández-Villaverde, J., Rubio-Ramírez, J. F., & Schorfheide, F. (2016). Chapter 9 - solution and estimation methods for dsge models. In J. B. Taylor & H. Uhlig (Eds.), (Vol. 2, p. 527 - 724). Elsevier.
- Flury, T., & Shephard, N. (2011). Bayesian inference based only on simulated likelihood: Particle filter analysis of dynamic economic models. *Econometric Theory*, 27(5), 933–956.
- Guerrieri, L., & Iacoviello, M. (2015). OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics*, 70(C), 22-38.
- Guerrieri, L., & Iacoviello, M. (2017). Collateral constraints and macroeconomic asymmetries. *Journal of Monetary Economics*, 90(C), 28-49.
- Gust, C., Herbst, E., López-Salido, D., & Smith, M. E. (2017). The empirical implications of the interest-rate lower bound. *American Economic Review*, 107(7).
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.
- Hausman, J. (2001, December). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives*, 15(4), 57-67. doi: 10.1257/jep.15.4.57
- He, Z., & Krishnamurthy, A. (2011). A model of capital and crises. *The Review of Economic Studies*, 79(2), 735–777.
- Herbst, E., & Schorfheide, F. (2016). *Bayesian Estimation of DSGE Models*. Princeton University

Press.

- Herbst, E., & Schorfheide, F. (2019). Tempered particle filtering. *Journal of Econometrics*, 210(1), 26 - 44. (Annals Issue in Honor of John Geweke “Complexity and Big Data in Economics and Finance: Recent Developments from a Bayesian Perspective”) doi: <https://doi.org/10.1016/j.jeconom.2018.11.003>
- Judd, K. L. (1992). Projection methods for solving aggregate growth models. *Journal of Economic Theory*, 58(2), 410-452.
- Judd, K. L. (1998). *Numerical Methods in Economics*. The MIT Press.
- Justiniano, A., & Primiceri, G. E. (2008). The Time-Varying Volatility of Macroeconomic Fluctuations. *American Economic Review*, 98(3), 604-641.
- Kim, J., & Kim, S. H. (2007). Two Pitfalls of Linearization Methods. *Journal of Money, Credit and Banking*, 39(4), 995-1001.
- Kollmann, R. (2015). Tractable latent state filtering for non-linear dsge models using a second-order approximation and pruning. *Computational Economics*, 45(2), 239–260.
- Kollmann, R. (2017). Tractable likelihood-based estimation of non-linear dsge models. *Economics Letters*, 161, 90 - 92. doi: 10.1016/j.econlet.2017.08.027
- Kulish, M., & Pagan, A. (2017). Estimation and Solution of Models with Expectations and Structural Changes. *Journal of Applied Econometrics*, 32(2), 255-274.
- Ljungqvist, L., & Sargent, T. J. (2004). *Recursive Macroeconomic Theory, 2nd Edition* (Vol. 1) (No. 026212274x). The MIT Press.
- Ljungqvist, L., & Sargent, T. J. (2012). *Recursive macroeconomic theory*. MIT press.
- Mendoza, E. G. (2006). Lessons from the debt-deflation theory of sudden stops. *American Economic Review*, 96(2), 411–416.
- Ruge-Murcia, F. (2012). Estimating nonlinear DSGE models by the simulated method of moments: With an application to business cycles. *Journal of Economic Dynamics and Control*, 36(6), 914-938.
- Smets, F., & Wouters, R. (2007, June). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3), 586-606. doi: 10.1257/aer.97.3.586
- Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters*, 20(2), 177–181.
- van Binsbergen, J. H., Fernández-Villaverde, J., Koijen, R. S., & Rubio-Ramírez, J. (2012).

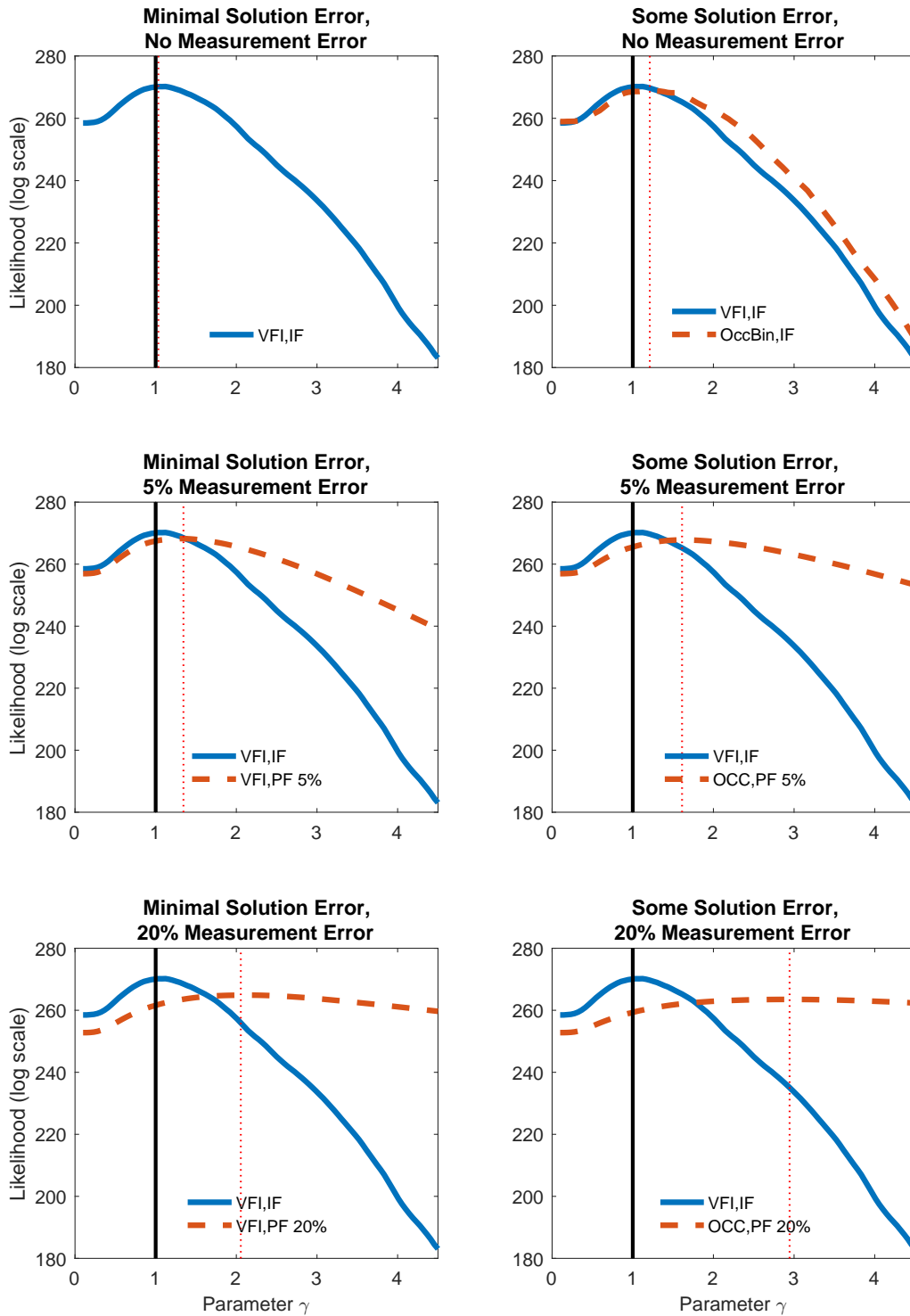
The term structure of interest rates in a dsge model with recursive preferences. *Journal of Monetary Economics*, 59(7), 634 - 648. doi: 10.1016/j.jmoneco.2012.09.002

Figure 1: POLICY FUNCTIONS OF THE CONSUMPTION-SAVINGS MODEL WITH OCCASIONALLY BINDING CONSTRAINTS



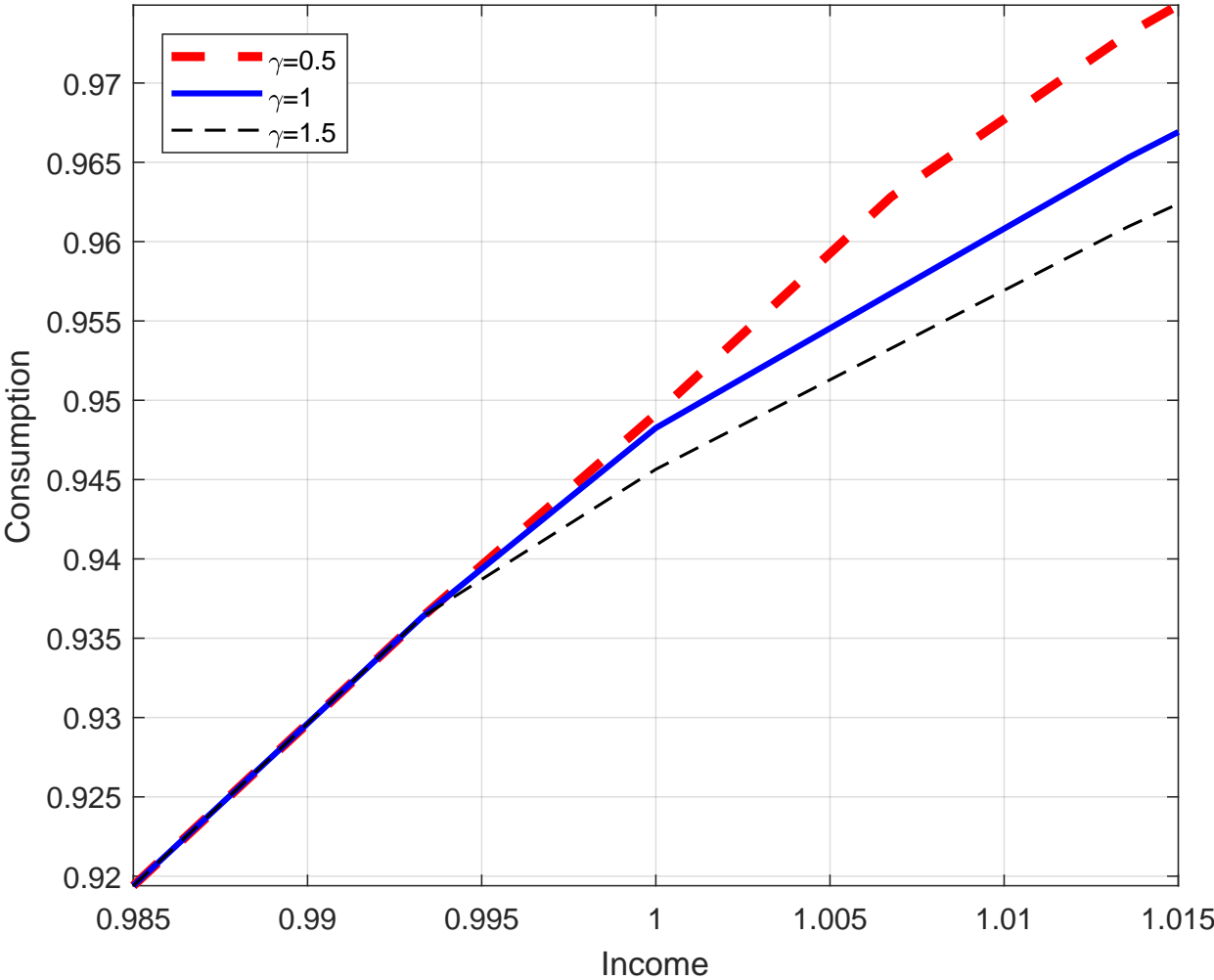
Note: “VFI” refers to the global solution method with value function iteration. “OccBin” refers to the OccBin solution in [Guerrieri and Iacoviello \(2015\)](#). “Linear” refers to a first-order perturbation solution. The policy functions shown are contours for an initial level of debt at its steady state value equal to 1.

Figure 2: LIKELIHOOD CONTOURS FOR ALTERNATIVE SOLUTION METHODS AND FILTERS. THE CASE OF NO MEASUREMENT ERROR IN THE DGP



Note: VFI and OccBin refer, respectively, to the global solution with value function iteration and to the OccBin solution in Guerrieri and Iacoviello (2015). IF and PF refer, respectively, to the inversion filter and to the particle filter for varying levels of the measurement error. The benchmark “VFI, IF” combination excludes misspecification error and is least affected by approximation error. The vertical lines in each panel denote the peaks of the likelihood contours shown. Under the true DGP, the value of the coefficient of relative risk aversion, γ , is 1.

Figure 3: CONSUMPTION FUNCTIONS FOR ALTERNATIVE VALUES OF THE COEFFICIENT OF RELATIVE RISK AVERSION, γ



Note: The policy functions shown are contours, computed using the value function iteration solution, for an initial value of debt at its non-stochastic steady state level equal to 1.

Table 1: Summary of Monte Carlo Results

T = 100										
	VFI					OccBin				
	Mean	Lower Bound	Upper Bound	Frequency True Value γ In 90% Credible Set	Average Percentile True Value γ	Mean	Lower Bound	Upper Bound	Frequency True Value γ In 90% Credible Set	Average Percentile True Value γ
IF	0.99	0.73	1.32	88	51	1.29	0.86	1.79	72	23
PF ME5	1.28	0.87	1.85	79	23	1.91	1.22	2.79	31	6
PF ME20	2.47	1.53	3.68	16	2	3.37	2.03	4.21	2	1
T = 500										
	VFI					OccBin				
	Mean	Lower Bound	Upper Bound	Frequency True Value γ In 90% Credible Set	Average Percentile True Value γ	Mean	Lower Bound	Upper Bound	Frequency True Value γ In 90% Credible Set	Average Percentile True Value γ
IF	0.99	0.89	1.11	88	52	1.23	1.06	1.40	32	9
PF ME5	1.28	1.09	1.50	26	5	1.78	1.49	2.11	2	0
PF ME20	2.67	2.08	3.38	0	0	3.55	2.78	4.17	0	0

Note: VFI and OccBin refer to, respectively, the global solution with value function iteration and the OccBin solution in [Guerrieri and Iacoviello \(2015\)](#). IF and PF refer, respectively, to the inversion filter and to the particle filter for varying levels of the measurement error. ME5 and ME20 refer to, respectively, the 5% and the 20% calibrations for the variance of measurement error relative to the variance of consumption in the model. The label $T = 100$ refers to results for Monte Carlo samples of 100 observations; the label $T = 500$ refers to the results for samples of 500 observations. To compute the credible sets, we take 20,000 draws from the posterior distribution using a random walk Metropolis algorithm. We use the first 10,000 draws as burn-in.

Appendix

A.1 Solution Methods

Value Function Iteration. We use dynamic programming to characterize a high-quality fully-nonlinear solution. The debt level B_t is the only state variable in the model. We seek a rule that will map the current state variable B_{t-1} and the realization of the stochastic process Y_t into a choice of B_t . We discretize and put boundaries on the support of the decision rule that we seek. We discretize the support of both B_{t-1} and Y_t . We consider a uniformly spaced set of nodes for B_{t-1} and B_t . The lower boundary for B_{t-1} is 25 percent below the non-stochastic steady state for borrowing. The upper boundary is 8 percent above the non-stochastic steady state for borrowing.²² We constrain Y to lie within three standard deviations of its process, i.e. $|\ln Y_t| \leq 3\sqrt{\frac{\sigma^2}{1-\rho^2}}$. We follow [Tauchen \(1986\)](#) to discretize the process $\ln Y_t$. Overall, the grid we consider involves 200 nodes for debt and 15 nodes for the income process.

The value function iteration algorithm that we use follows closely Chapter 12 of [Judd \(1998\)](#) and Chapter 3 of [Ljungqvist and Sargent \(2012\)](#). To accelerate the convergence of the dynamic programming algorithm, we alternate iterations on the value function and on the policy function, using the Howard improvement algorithm, as described in Chapter 2 of [Ljungqvist and Sargent \(2012\)](#).

OccBin Solution. The economy features two regimes: a regime in which the collateral constraint binds and a regime in which it does not (but is expected to bind in the future). The OccBin solution resolves the problem of computing decision rules that approximate the equilibrium adequately under both regimes. Essentially, this method linearizes the model at the constraint and away from the constraint and then joins the two systems of equations using a shooting algorithm that reduces the problem to only solving for the expected duration of each regime, rather than solving for paths of all endogenous variables. The implementation of this algorithm is the same as in [Guerrieri and Iacoviello \(2015\)](#), which can be consulted for further details.

First-order Perturbation. For this solution, we disregard the possibility that the constraint could ever be slack and linearize the model around the deterministic steady state. We use the

²²The asymmetric solution bounds allow us to economize on solution nodes given that the borrowing constraint implies an asymmetric ergodic distribution for borrowing.

solution computed by Dynare, a popular and convenient set of tools for solving and estimating DSGE models. See [Judd \(1998\)](#) for a general description and of this standard solution algorithm. A numerical implementation is described by [Anderson and Moore \(1985\)](#).

A.1.1 Euler Errors

To gauge the quality of the approximation of each solution method that we consider, we follow the standard bounded rationality approach of [Judd \(1998\)](#). Accordingly, starting from the Euler equation for consumption of Equation (10) and reproduced here,

$$C_t^{-\gamma} = \beta RE_t (C_{t+1}^{-\gamma}) + \lambda_t,$$

we take each side to the power of $-\frac{1}{\gamma}$ to obtain

$$C_t = [\beta RE_t (C_{t+1}^{-\gamma}) + \lambda_t]^{-\frac{1}{\gamma}}.$$

This transformation allows us to express the Euler equation in units of consumption. We use each numerical solution method to evaluate the left-hand side and the right-hand side of the equation above and call the difference between the two sides an Euler residual function. Figure [A.1](#) plots contours of the residual function for each of the solution method considered. The residuals are expressed in log 10 scale. Accordingly, in the figure, a level of -1 can be interpreted as an error of \$1 per \$10 of consumption and a level of -5 as an error of \$1 per \$100,000 of consumption. The figure confirms the high accuracy of the value function iteration solution, modest errors for the OccBin solution and large errors for the linear solution that disregards the occasional binding constraint on borrowing.

A.2 The likelihood function

A.2.1 Inversion Filter

In the absence of measurement error, we can combine Equations (1) and (2) to obtain:

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

Where $f \equiv (g \circ h)$ maps $\mathbb{R}^{n_s} \times \mathbb{R}^{n_\eta}$ into \mathbb{R}^{n_y} , with n_s , n_y , n_η , denoting the number of endogenous state variables, observed variables, and structural shocks, respectively. When the structural innovations η_t are drawn from a multivariate Normal distribution with covariance matrix Σ we can write the log-likelihood of the model as:

$$\log(p(y_{1:T})) = -\frac{Tn_y}{2} \log(2\pi) - \frac{T}{2} \log(\det(\Sigma)) - \frac{1}{2} \sum_{t=1}^T \eta_t' \Sigma^{-1} \eta_t + \sum_{t=1}^T \log \left(\left| \det \frac{\partial \eta_t}{\partial y_t} \right| \right) \quad (12)$$

In practical applications computing the log-likelihood function via the inversion filter (Equation 12) poses two challenges. First, the true structural shocks η_t are unobserved and inverting the function $f(\cdot)$ to recover them might be cumbersome. Second, and as we discussed in section 2, the function $h(\cdot)$ has to be approximated numerically with different degrees of accuracy and hence the Jacobian matrix $\frac{\partial \eta_t}{\partial y_t}$ —which is defined implicitly by Equation (11)—can be computed in closed form only in very special cases.

With respect to the first issue, Figure 1 shows that the policy function for consumption is monotonic despite the occasionally binding constraint of the model. Hence, in our application, we can invert the policy functions $h(\cdot)_o$ and $h(\cdot)_{vfi}$ to size the shocks given observations using standard algorithms for the solution of nonlinear equations.²³ We initialize the states at their true values, which, given observations is sufficient to recover all other shocks. For the second issue, in our benchmark specification which uses the most accurate solution approximation given by $h_{vfi}(\cdot)$, we construct the Jacobian matrix using a multi-point finite difference method. When using a OccBin solution, the Jacobian matrix is a by-product of the solution. As shown in Guerrieri and Iacoviello (2017), this is possible due to the local linearity in η_t of the matrices that define the solution $h_o(\cdot)$.

A.2.2 Particle Filter Description

Our implementation of the bootstrap particle filter follows Fernández-Villaverde and Rubio-Ramírez (2007). Given structural parameters summarized in the vector θ , the likelihood of the model can be written in its prediction error decomposition form, as in Equation (13). In turn, the prediction

²³In other applications, the nonlinearities in the model might be more severe and multiple sets of shocks may be consistent with the same observation. Standard results can be invoked to construct the likelihood function even for this general case of a correspondence. A root finding algorithm can investigate this possibility using alternative initial points chosen on a grid. If multiple solutions were detected, the likelihood would simply be modified by summing the probability associated with each of the solutions.

error decomposition is related to the distribution of the hidden states s_t of the model in Equation (14) and the estimate of s_t given time $t - 1$ data is related to the estimate of s_{t-1} given time $t - 1$ data via Equation 15.

$$p(y_{1:T}|\theta) = \prod_{t=1}^T p(y_t|y_{1:t-1}, \theta) \quad (13)$$

$$p(y_t|y_{1:t-1}, \theta) = \int p(y_t|s_t, y_{1:t-1})p(s_t|y_{1:t-1})ds_t \quad (14)$$

$$p(s_t|y_{1:t-1}) = \int p(s_t|s_{t-1}, y_{1:t-1})p(s_{t-1}|y_{1:t-1})ds_{t-1} \quad (15)$$

In the linear Gaussian case, the Kalman filter delivers exact formulas for $p(s_t|y_{1:t-1})$, $p(s_{t-1}|y_{1:t-1})$, and hence $p(y_t|y_{1:t-1})$. These exact formulas do not apply to nonlinear models. The idea of the particle filter is to simulate many particles representing the hidden states. By appropriately propagating these particles through time, they will then approximate the distributions of interest necessary to calculate the likelihood.

We rely here on a simple but widely used implementation of the particle filter, known as the bootstrap particle filter (Fernandez-Villaverde and Rubio-Ramirez (2005); Fernández-Villaverde and Rubio-Ramírez (2007); Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006)). The implementation steps for this filter are as follows:

0. **Initialization:** Simulate N particles from an initial distribution $p(s_0)$. Set the weight of each particle i : $w_0^{(i)} = 1$. This distribution comes from the steady state distribution of the states. We approximate it by taking a draw for each particle i following a simulation of the model for 50 periods. To obtain an accurate approximation of the likelihood, we set the number of particles in our simulation to $N = 2,000,000$.

Entering into time t , we have a swarm of N particles $\{s_{t-1}^{(i)}\}_{i=1}^N$ and associated weights $\{w_{t-1}^{(i)}\}_{i=1}^N$ that are distributed according to $p(s_{t-1}|y_{1:t-1})$.

1. **State forecast:** Take each particle $s_{t-1}^{(i)}$ simulate it forward via the state transition equation $h(s_{t-1}, \eta_t; \theta)$. This then gives us a new swarm of particles $\{s_t^{(i)}\}_{i=1}^N$ and associated weights $\{w_{t-1}^{(i)}\}_{i=1}^N$ that are distributed according to $p(s_t|y_{1:t-1})$.

2. **Observable forecast:** For each particle $s_t^{(i)}$, we can calculate $p(y_t|s_t^{(i)}, y_{1:t-1}) = p(y_t|s_t^{(i)})$. The weighted average of these particles is the discrete approximation to the integral that defines $p(y_t|y_{1:t-1}, \theta)$. Call this approximation $\hat{p}(y_t|y_{1:t-1}, \theta)$.

$$p(y_t|y_{1:t-1}, \theta) = \int p(y_t|s_t, y_{1:t-1})p(s_t|y_{1:t-1})ds_t \approx \frac{1}{N} \sum_{i=1}^N p(y_t|s_t^{(i)})w_{t-1}^{(i)} = \hat{p}(y_t|y_{1:t-1}, \theta) \quad (16)$$

3. **State update:** We now incorporate time t information y_t to have the particles be distributed according to $p(s_t|y_{1:t})$. We know the following identity holds:

$$p(s_t|y_{1:t}) = \frac{p(y_t|s_t, y_{1:t-1})p(s_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} \quad (17)$$

We already have particles distributed according to $p(s_t|y_{1:t-1})$. We can think about this as an importance sampling problem where we are using $p(s_t|y_{1:t-1})$ as the proposal density. We can use the expression $\frac{p(y_t|s_t, y_{1:t-1})}{p(y_t|y_{1:t-1})}$ to update each particle's importance weight. Therefore, the new importance weights are

$$w_t^{(i)} = w_{t-1}^{(i)} \frac{p(y_t|s_t^{(i)})}{\frac{1}{N} \sum_{i=1}^N p(y_t|s_t^{(i)})w_{t-1}^{(i)}} \quad (18)$$

4. **(Optional) Resampling step:** Every period, there is an optional step to replenish low weight particles with high weight ones. We draw N new particles from the set $\{s_t^{(i)}\}_{i=1}^N$ with replacement according to their weights $\{w_t^{(i)}\}_{i=1}^N$. Usually this step is done when certain measures of particle degeneracy (such as effective sample size) reach critical thresholds.

We iterate steps 1 – 4 until time T . This completes the recursion.

Importance of measurement error

The particle filter relies on the existence of a greater or equal number of elements of stochasticity than observables:

$$\dim(\eta_t) + \dim(\zeta_t) \geq \dim(y_t) \quad (19)$$

With a lower dimension of randomness than the number of observables, the researcher runs into well known problems of stochastic singularity. In fact, the bootstrap particle filter as imple-

mented in the algorithm actually relies on a more elements of randomness than observables. If the dimension of the shocks is exactly equal to the number of observables, only a finite number of states s_t can exactly match the observables y_t , which makes the implementation of the state and observable forecast steps of the filter as currently discussed inoperable. In practice, researchers assume a certain amount of measurement error in each observable y_t .

Suppose that $\zeta_t \sim N(0, \sigma_\zeta^2)$ and take a univariate y_t case for simplicity. The density $p(y_t | s_t^{(i)})$ is then given by:

$$p(y_t | s_t^{(i)}) \propto \frac{1}{\sqrt{\sigma_e^2}} e^{-\frac{(y_t - g(s_t^{(i)}; \theta))^2}{2\sigma_e^2}} \quad (20)$$

As one can see from Equation (20), a smaller assumed standard deviation of the measurement error penalizes poor proposals $s_t^{(i)}$ more. This increases the variance of $p(y_t | s_t^{(i)})$ across particles and there also that of the weights $w_t^{(i)}$.

Importance of the proposal density for $s_t^{(i)}$

An important factor in determining the accuracy of the particle filter approximation to the likelihood is the proposal density used to generate the time t particles $s_t^{(i)}$. Given a particle swarm $\{s_{t-1}^{(i)}, w_{t-1}^{(i)}\}_{i=1}^N$, we are proposing the time t particles by simulating from the transition density. Therefore, our proposal distribution is $p(s_t | s_{t-1})$.

A.3 Too little measurement error

For illustration, let's consider a case slightly different from the one we posited, in which the DGP includes measurement error, but the econometrician does not realize it. Adding normally distributed measurement error changes both features of the consumption function: it makes consumption more volatile, and dampens skewness. Accordingly, the econometrician would think that consumption function is consistent with a lower value of γ , such as the one represented by the dashed, red line. Figure A.2 in the online appendix illustrates this case, showing how “too little” assumed measurement error relative to what is embedded in the DGP biases the estimate of γ downwards, regardless of the solution error.

A.4 First-order perturbation solution and the Kalman filter

Figure A.3 shows results analogous to those presented in Figure 2. The solid, blue line denotes again the likelihood contour based on the global solution with value function iteration in conjunction with the inversion filter. The dashed, red line denotes the likelihood contour based on a first-order perturbation solution in conjunction with a Kalman filter that excludes measurement error, in line with the DGP. We focus here on approximation error. Since the model is linearized at the constraint and the consumption function is unaffected by γ at the constraint, as the figure makes apparent, γ is not identified in this case.

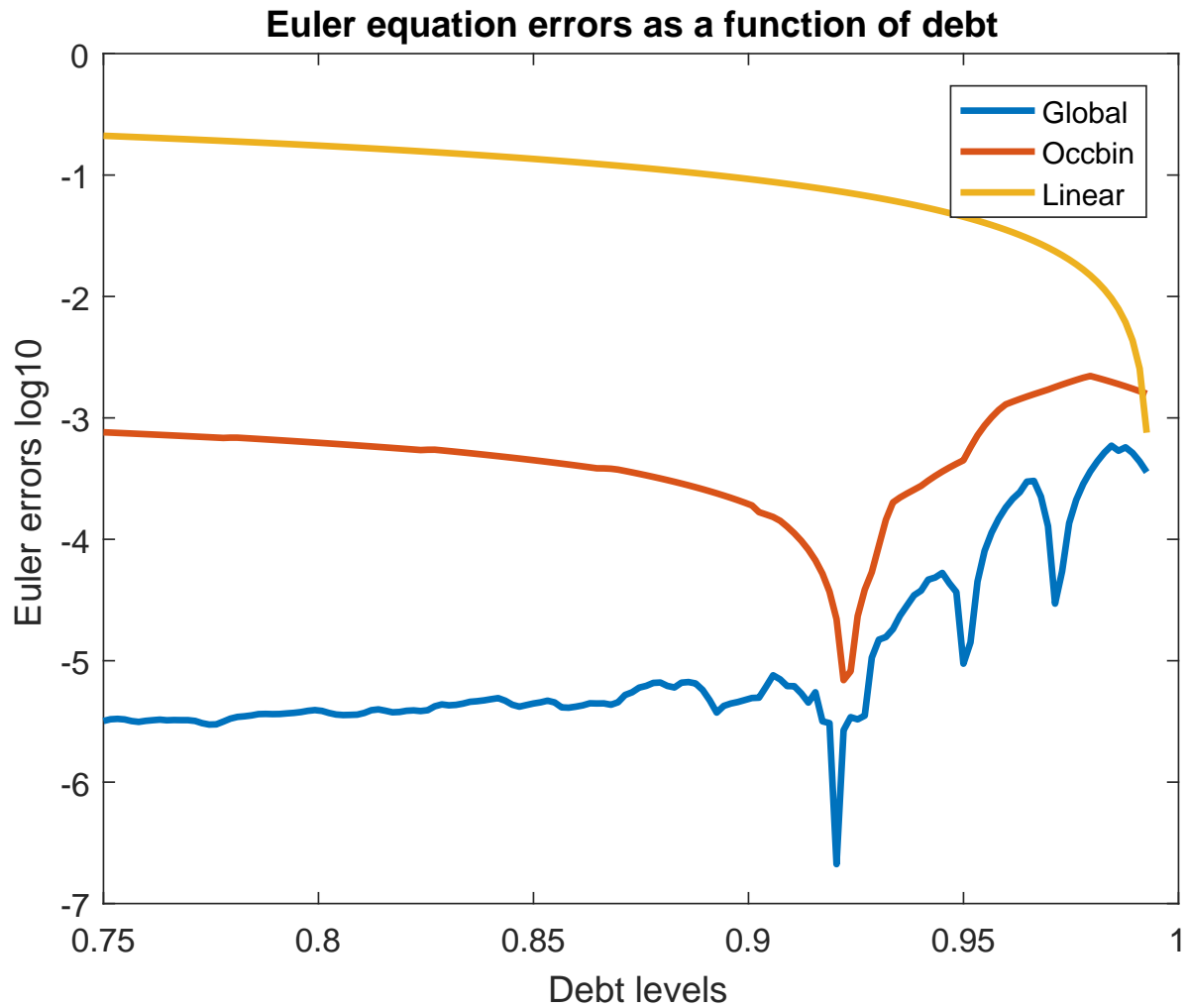
A.5 Estimation of additional parameters

In the main text, we considered the estimation of γ holding the parameters governing the shock process, ρ and σ , at their true values. As a robustness exercise, we consider the estimation of γ , ρ , and σ together. We implement our estimation by sampling the posterior distribution of the parameters of interest using a standard MCMC algorithm. We use uniform priors on all parameters, so that our estimate of the posterior mode coincides with the maximum likelihood estimate. We run the sampler to obtain 20,000 draws from the posterior distribution and discard the first 10,000 draws as a burn-in. To isolate the role of likelihood misspecification, we focus on the most accurate solution method and estimate the model with minimal solution error and 20% measurement error (VFI,PF 20%). We include the estimation results for the model with minimal solution error and no measurement error (VFI,IF) as a reference. We select a sample from the Monte Carlo exercise and use 500 observations in our estimation. The results are shown in Figure A.5.

We first discuss the estimation of the parameter of interest γ . The top-left panel (Only gamma), shows the results of our baseline estimation when the parameters ρ and σ are fixed to their true values. Consistent with the results in the main text, the results from this sample are indicative of a substantial bias in γ , with an estimated posterior mode of 3.96. The top-right panel (All parameters), shows the estimated value of γ when we also estimate the parameters that govern the shock process. The posterior density shifts closer to the true value, but there is still a noticeable

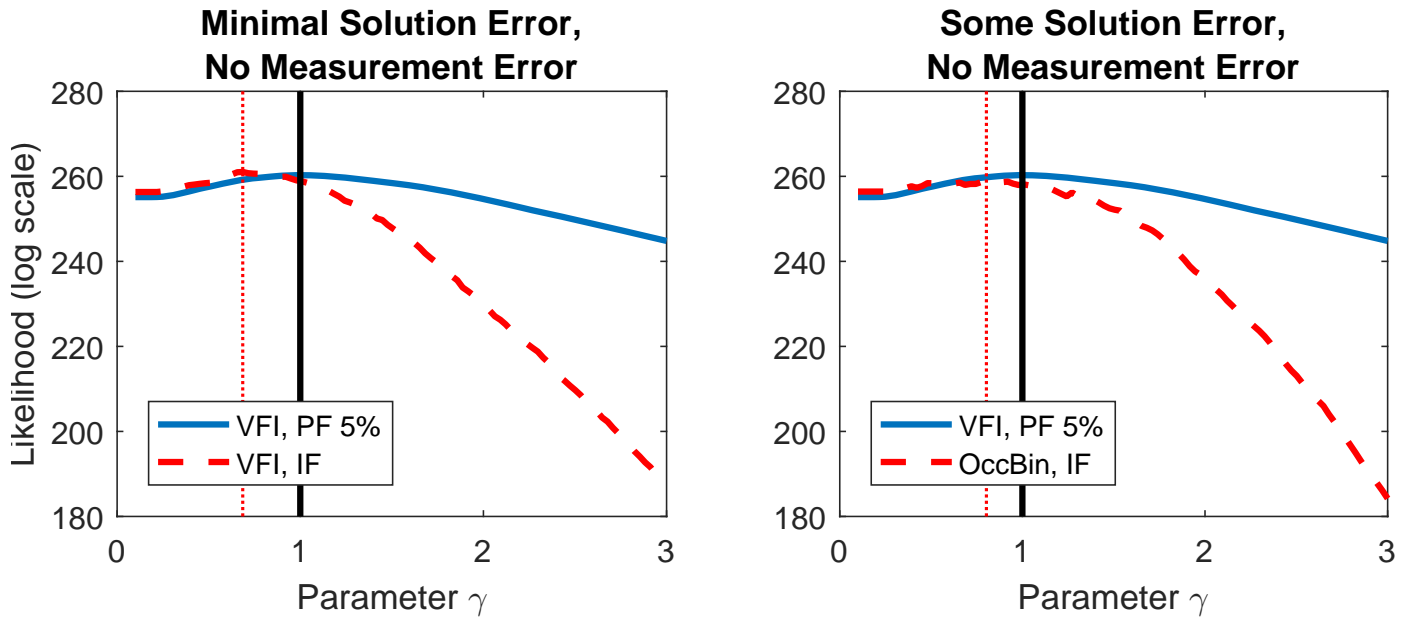
difference, with an estimated posterior mode of 1.59. The bottom panels show the estimation results for the parameters ρ and σ . In both cases, the posterior densities are shifted away from the true values, and in particular the posterior density of σ barely includes the true value. In our particular example, when estimating additional model parameters, the results are still suggestive of a substantial bias in the estimate of γ and also a bias in the estimates of the other parameters.

Figure A.1: Euler Errors



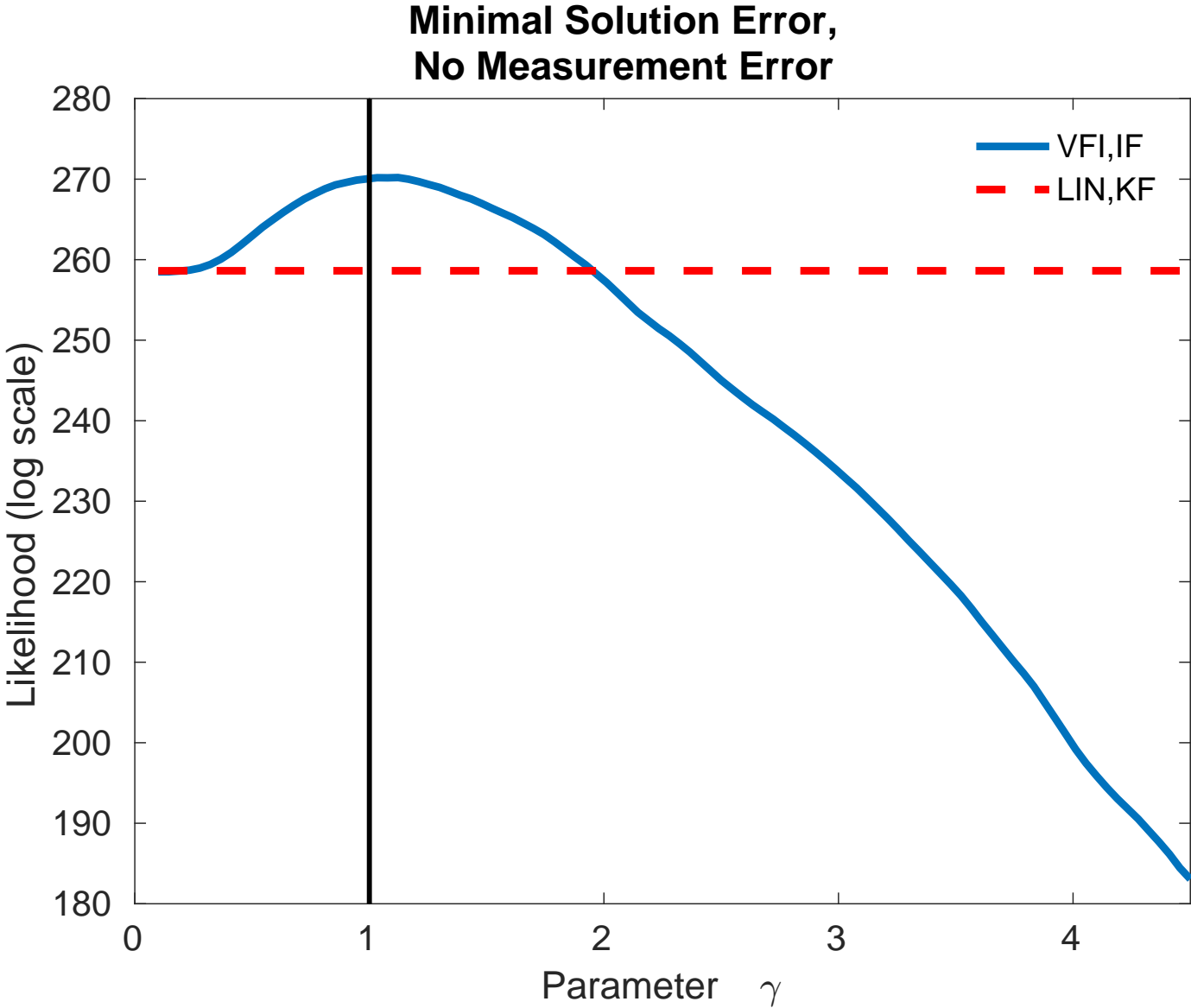
Note: Log_{10} of absolute value of Euler equation errors.

Figure A.2: CONTOURS OF THE LIKELIHOOD WHEN THE DGP INCORPORATES MEASUREMENT ERROR



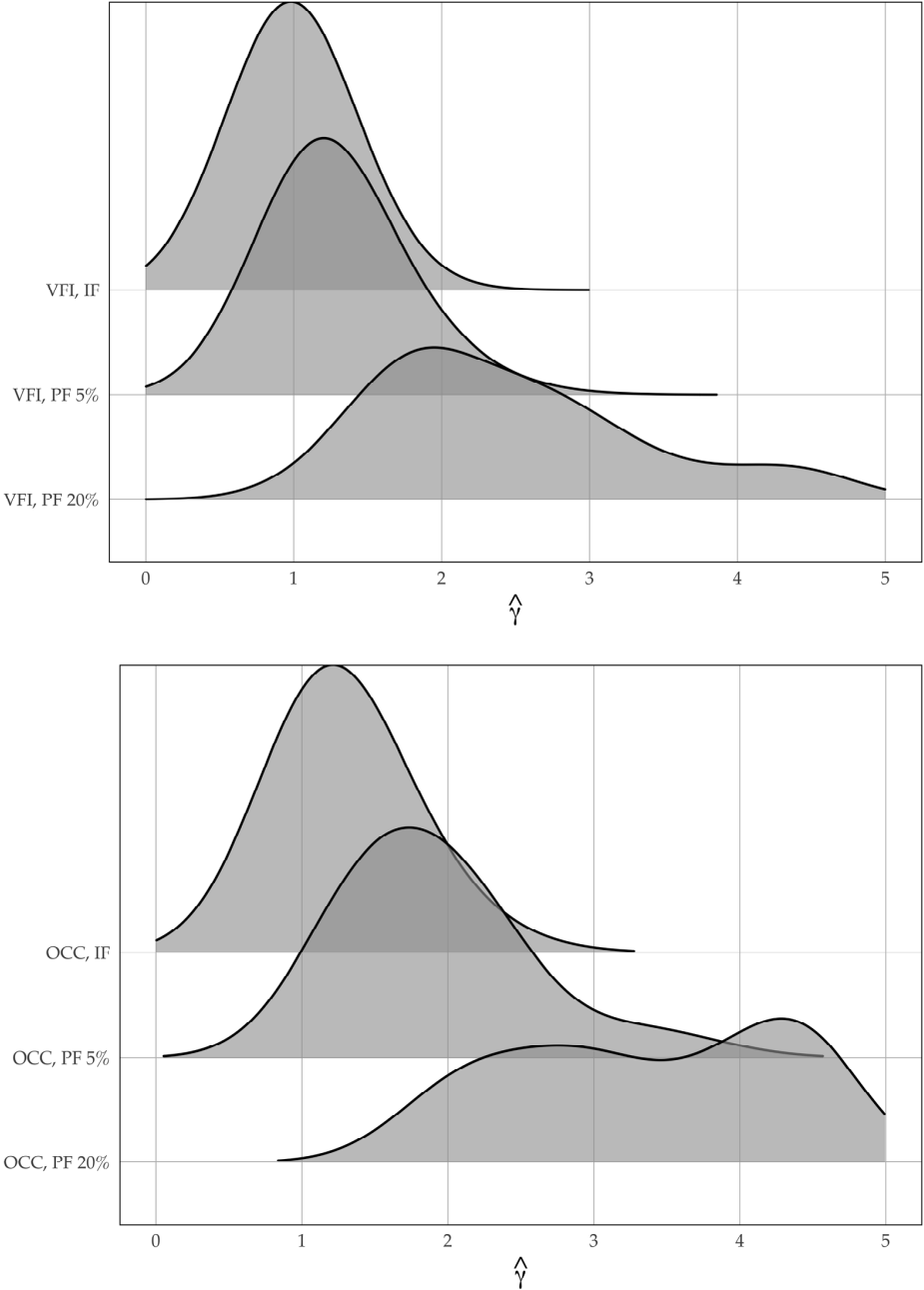
Note: In this case, the DGP is the model solved with the value function iteration method and it includes 5 percent measurement error in the observation equation for consumption. The particle filter is calibrated to recognize the measurement error in the DGP. The inversion filter incorrectly excludes measurement error.

Figure A.3: COMPARISON OF LIKELIHOOD FUNCTIONS FROM ALTERNATIVE APPROXIMATIONS



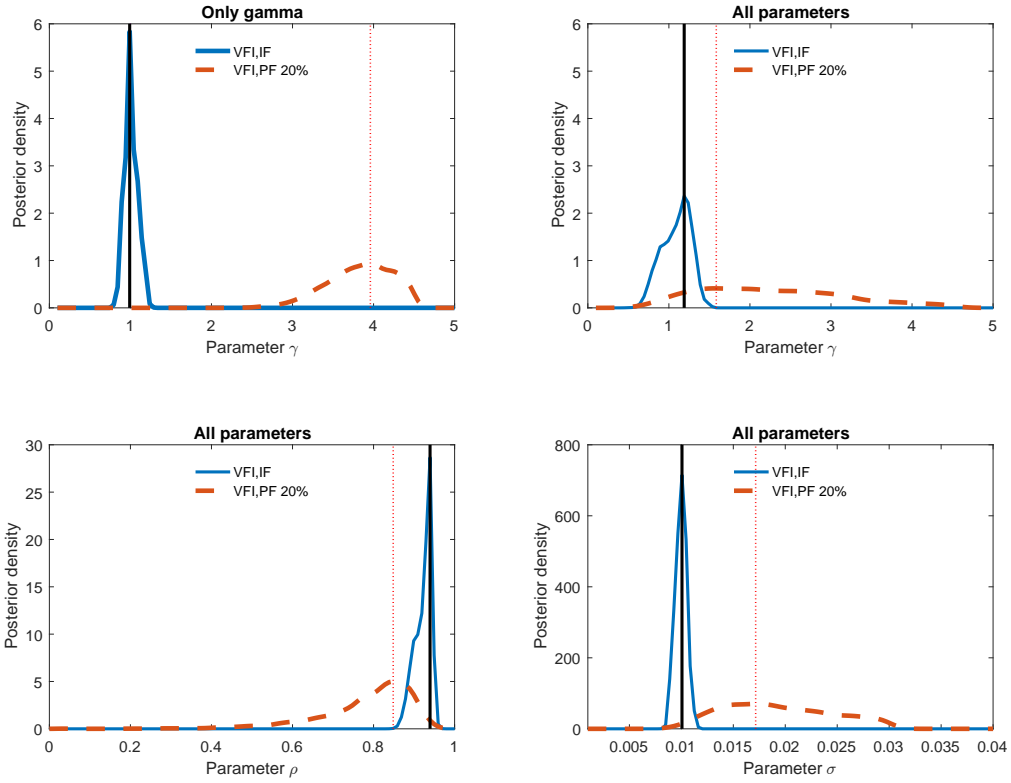
Note: “VFI” refers to the global solution method with value function iteration. “LIN” refers to a first-order perturbation solution. “IF” refers to the inversion filter and “KF” to the kalman filter. The solid vertical line denotes the peak of the likelihood contour for the benchmark VFI, IF combination.

Figure A.4: Sampling Distributions of the Mode of the Posterior Distribution for Alternative Solution Methods and Filters



Note: VFI, OccBin and Linear refer, respectively, to the global solution with value function iteration, the OccBin solution in [Guerrieri and Iacoviello \(2015\)](#), and first-order perturbation solution. IF and PF refer, respectively, to the inversion filter and to the particle filter for varying levels of the measurement error. The benchmark “VFI, IF” combination excludes misspecification error and is least affected by approximation error. The vertical lines in each panel denote the peaks of the likelihood contours shown. Under the DGP, the value coefficient of relative risk aversion, γ is 1.

Figure A.5: Posterior Distributions: Estimation of Additional Parameters



Note: VFI and OccBin refer, respectively, to the global solution with value function iteration and to the OccBin solution in [Guerrieri and Iacoviello \(2015\)](#). IF and PF refer, respectively, to the inversion filter and to the particle filter for varying levels of the measurement error. The benchmark “VFI, IF” combination excludes misspecification error and is least affected by approximation error. The vertical lines in each panel denote the peaks of the posterior density shown that coincides with the maximum likelihood estimate. Under the true DGP, the value of the coefficient of relative risk aversion, γ , is 1. The true value of the parameters governing the exogenous income process, ρ and σ , are 0.9 and 0.01, respectively.