FUNDAÇÃO GETULIO VARGAS
ESCOLA de PÓS-GRADUAÇÃO em ECONOMIA

Paulo Fernando Nericke Motula

Orientador: Tiago Couto Berriel

Estimation of DSGE Models: A Monte Carlo Analysis

Rio de Janeiro

2013

Paulo Fernando Nericke Motula

Estimation of DSGE Models: A Monte Carlo Analysis

Dissertação submetida à Escola de Pós-Graduação em Economia da Fundação Getúlio Vargas como requisito para obtenção do título de Mestre em Economia

Orientador: Tiago Couto Berriel

Rio de Janeiro

2013

Ficha catalográfica elaborada pela Biblioteca Mario Henrique Simonsen/FGV

Motula, Paulo Fernando Nericke
Estimation of DSGE models : a Monte Carlo analysis / Paulo Fernando Nericke Motula. - 2013.
83 f.

Dissertação (mestrado) - Fundação Getulio Vargas, Escola de Pós-Graduação em Economia.
Orientador: Tiago Couto Berriel.
Inclui bibliografia.

1. Método de Monte Carlo. 2. Modelos econométricos. I. Berriel, Tiago Couto. II. Fundação Getulio Vargas. Escola de Pós- Graduação em Economia. III. Título.

CDD – 518.282

FUNDAÇÃO GETULIO VARGAS

PAULO FERNANDO NERICKE MOTULA

ESTIMATION OF DSGE MODELS: A MONTE CARLO ANALYSIS .

Dissertação apresentada ao Curso de Mestrado em Economia da Escola de Pós-Graduação em Economia para obtenção do grau de Mestre em Economia.

Data da defesa: 18/06/2013

ASSINATURA DOS MEMBROS DA BANCA EXAMINADORA

Tiago Couto Berriel
Orientador (a)

Marco Antonio Cesar Bonomo

Carlos Viana de Carvalho

Resumo

Neste trabalho investigamos as propriedades em pequena amostra e a robustez das estimativas dos parâmetros de modelos DSGE. Tomamos o modelo de Smets and Wouters (2007) como base e avaliamos a performance de dois procedimentos de estimação: Método dos Momentos Simulados (MMS) e Máxima Verossimilhança (MV). Examinamos a distribuição empírica das estimativas dos parâmetros e sua implicação para as análises de impulso-resposta e decomposição de variância nos casos de especificação correta e má especificação. Nossos resultados apontam para um desempenho ruim de MMS e alguns padrões de viés nas análises de impulso-resposta e decomposição de variância com estimativas de MV nos casos de má especificação considerados.

Palavras-chave: DSGE, Monte Carlo, Má especificação

Abstract

We investigate the small sample properties and robustness of the parameter estimates of DSGE models. Our test ground is the Smets and Wouters (2007)'s model and the estimation procedures we evaluate are the Simulated Method of Moments (SMM) and Maximum Likelihood (ML). We look at the empirical distributions of the parameter estimates and their implications for impulse-response and variance decomposition in the cases of correct specification and two types of misspecification. Our results indicate an overall poor performance of SMM and some patterns of bias in impulse-response and variance decomposition for ML under the types of misspecification studied.

Keywords: DSGE, Monte Carlo, Misspecification

List of Figures

List of Tables

Contents

1 Introduction

Quantitative implications of DSGE models depend crucially on the parameter values used.[1] With that in mind, this work investigates the performance of two popular econometric procedures for the estimation of Dynamic Stochastic General Equilibrium (DSGE) models: Simulated Method of Moments (SMM) and Maximum Likelihood estimation (ML).

As a test ground we choose the Smets and Wouters (2007) (SW07) model. The SW07 model is widely cited and recently has been subject to some studies about its identifiability (see Komunjer and Ng (2011), Iskrev (2010b), Iskrev (2010a) and Caglar et al. (2012)).

We evaluate the small sample properties of the parameter estimates of the linearized model of both methods, considering the situations of *correct* specification (the estimated model is the Data Generating Process (DGP)) and 2 kinds of *misspecification* (the estimated model is different from the DGP).

Specifically, in the first case of misspecification we consider, the estimated model is linear while the DGP is a nonlinear version of the SW07 model, solved up to second order. In the second sort of misspecification, the DGP is the linear SW07 model and the estimated model is restricted to have i.i.d shocks.

We look at the distribution of the parameter estimates and their implications for impulse-response analysis and variance forecast decomposition.

Our results indicate poor performance of SMM relative to ML and that bias from the types of misspecifications considered can be big enough to have implications for business cycle accounting and impulse-response analysis.

2 Related Literature

In the literature, the work that comes closest to ours is Ruge-Murcia (2007). He presents different econometric methods in the context of estimation of a small *Real Business Cycle* (RBC) model with three observables and one shock[2]. The methods are ML (with and without priors[3]), GMM, SMM and Impulse-response Matching. Monte Carlo analysis is then used to compare the methods with respect to their small sample properties and behavior across different selection of observables and the presence of three types of misspecification. One of them is the estimation of a linear model with data generated from the nonlinear solution of the same model.

[1] See Ríos-Rull et al. (2012) for a recent example of their impact in the business cycle accounting.
[2] So that stochastic singularity is present.
[3] With the use of conjugate and diffuse priors under normality assumption, it was made possible to obtain the mean estimates through direct maximization, avoiding sampling methods as the Metropolis-Hastings algorithm.

Overall, the results favor the use of informative priors and measurement errors in the ML context, disfavor Impulse-response matching but show good performance of the other moment-based methods GMM and SMM. In special, results show some small evidence of greater robustness to misspecification in these two methods when compared to ML.

It is our opinion that some points in the Ruge-Murcia (2007)'s study deserve better attention. First, we think that it is worthwhile to explore the differences of the methods in a model with better empirical adherence and closer to the contemporary practice of DSGE estimation. Second, we think the evidence favoring the SMM/GMM may be underestimated by the use of correct priors and measurement errors only in the context of ML[4]. Third, estimating the model for different sample sizes we are able to analyze the rate of convergence of each method.

It seems the vast majority of econometric work with DSGE models can be recognized as following one of two approaches: a likelihood-based procedure (with a classical or a Bayesian perspective) or a moment-based procedure (as GMM, SMM or Impulse-response matching). As we take, respectively, ML and SMM as representatives of each approach, our choice of methods is set to investigate the performance differences that go along with that division.

The more natural choice of a Bayesian perspective over the classical one of ML is prevented due to the extra computational burden generated. One advantage of undertaking the study in a classical perspective is that we sidestep the problem of choosing and analyzing the effect of priors over the result. [5]

The SMM procedure seems to be the moment-based choice of recent literature.[6] Although used in important studies (Rotemberg and Woodford (1997), Christiano et al. (2005) and Christiano et al. (2010)), we choose to study SMM over Impulse-response matching considering the disappointing performance already show in Ruge-Murcia (2007) and Canova and Sala (2009).

Our work also relates to a growing body of literature concerned with the reliability of the parameter estimates in DSGE models. Canova and Sala (2009) discuss identification problems to which DSGE models are commonly subject, focusing on the method of *Impulse-response Matching*. They use a sample free simulation procedure to measure the size of the weak identification region in a prototype medium-scale DSGE model and perform Monte Carlo analysis in the same model in order to gauge the distortions added to sample problems.

Iskrev (2010b) establishes sufficient conditions to local identification of linear DSGE models from first and second moments of the observables. Studying restrictions implied by observational

[4]In particular, the option for not working with measurement errors in the GMM and SMM context imply in a restriction to the number of observables.

[5]The choice of right priors would be an easy one, but probably not much relevant to empirical work.

[6]See Ruge-Murcia (2007, 2012) for a methodological perspective and Kim and Ruge-Murcia (2009, 2011) for applied work.

equivalence on the autocovariance matrix of the observables, Komunjer and Ng (2011) achieve similar results, being able to extend them for singular models as well.

Iskrev (2010a) deals with the problem of *weak identification* in the context of Maximum Likelihood (ML) estimation, establishing some measures of the problem.

An and Schorfheide (2007) review Bayesian methods for the estimation of DSGE models. In particular, for a small-scale model with Calvo pricing, they estimate the linear and the second-order accurate solution of the model on data generated from the last. Overall, the estimates are similar, with a few exceptions favoring the quadratic estimation [7].

Koop et al. (2011) study identification from a Bayesian perspective, but focusing on identification from the data. They show that the common practice of comparing posteriors and priors in order to assess identification can be misleading. Specifically, they show that the posterior of an (strictly) unidentified parameter will equal the prior only in the case of variational free parameter space[8] and prior independence.

They also develop two indicators of identification, the first for exact (which they call *Bayesian Comparison Indicator*) and the second for weak identification (which they call *Bayesian Learning Indicator*). Caglar et al. (2012) apply their second method to the SW07 model, suggesting that most of the parameters in the structural equations are not well identified.

Fernández-Villaverde et al. (2006) study conditions under which, as the approximated policy function converges to the exact policy, the approximated likelihood (computed from the approximated policy) also converges to the exact likelihood. As they investigate the speed of convergence, they find that second order approximation errors in the policy function have first order effects on the likelihood function, something, they suggest and show with an example, can compromise the ML estimates based on the approximated policy function.

Ackerberg et al. (2009) show (by counterexample) that Fernández-Villaverde et al. (2006)'s result is not general and that even if their result were correct one could not conclude the approximation error of the parameter estimates would have different order of magnitude than the approximation error of the policy function. Unfortunately, the authors were not able to provide a constructive result assuring they would have the same order of magnitude, unless for the restrictive case of static likelihood (independent of the realization of the history of the observables).

Adolfson and Lindé (2011) use Monte Carlo methods to study the small sample properties of the ML estimates of Adolfson et al. (2007) model. They document that the ML estimator

[7]They perform an illustration, rather than a Monte Carlo procedure, since they use a single sample.
[8]This is potentially important in the case of DSGE models, as the range of one parameter often depends on another due to determinacy conditions.

is consistent, unbiased for nearly all parameters and that the MSE is usually low, so that the implications of weak identifications for the MSE of the ML estimates is quite limited. However, the median Asymptotic Standard Error (ASE)[9] underestimates the uncertainty about the parameter estimates.

Our first misspecification experiment is inspired by Ruge-Murcia (2007), An and Schorfheide (2007) and stimulated by the not fully solved controversy of Fernández-Villaverde et al. (2006) and Ackerberg et al. (2009)[10]. In particular, we extend An and Schorfheide (2007)'s experiment (in a classical perspective and with a different model) in our first misspecification simulation to compare the performance of ML and SMM.

3 Estimation Methods for Linear DSGE Models

3.1 The Problem of Identification

Canova and Sala (2009) define the following concepts of identification:

1. **Observational equivalence**: The mapping between the structural parameters and the objective function does not have a unique minimum. In that case the estimator will be inconsistent.

2. **Under-identification**: The objective function is independent of certain parameters.

3. **Partial identification**: A group of parameters enters the objective function proportionally, making them separately unrecoverable.

4. **Weak identification**: The objective function has little curvature in the direction of some parameters. In this case the estimator is consistent but its convergence is slow, making small sample inference unreliable.

Both partial and under-identification are specific versions of observation equivalence. Iskrev (2010b) has a more precise definition (p. 192)[11]:

Definition 1. *Let $\theta \in \Theta \subset \mathbb{R}^n$ be the parameter vector of interest, and suppose that inference about θ is made on the basis of T observations of a random vector $\boldsymbol{x}$ with a known joint probability density function $f(X;\theta)$, where $\mathbf{X} = [x_{\mathbf{1}}, ..., x_{\mathbf{T}}]$. A point $\theta_0 \in \Theta$ is said to be* ***globally*** *identified if*

$$f(X;\tilde{\theta}) = f(X;\theta)\, with\, probability\, 1 \Rightarrow \tilde{\theta} = \theta_0 \tag{3.1}$$

[9]The one computed with the inverse Hessian of the likelihood function.

[10]As observed in section 3.3, in view of the study of Santos and Peralta-Alva (2005), one would not expect SMM to suffer from the same problem.

[11]This definition is traditional in econometrics, being already followed by Rothenberg (1971).

for any $\tilde{\theta} \in \Theta$. If 3.1 holds only for values $\tilde{\theta}$ in an open neighborhood of θ_0, then θ_0 is said to be ***locally*** *identified.*

While observation equivalence (in a strict sense) can be treated as a problem of the model, weak identification depends of the curvature of the objective function, thus varying with the data and the method used (see the discussion in Canova and Sala (2009)).

Let $\mathbf{z_t}$ be a m-dimensional vector of stationary variables, $\bar{\mathbf{z}}$ its steady-state and $\hat{\mathbf{z}}_\mathbf{t} = \mathbf{z_t} - \bar{\mathbf{z}}$. Let $\mathbf{u}_t$ be an n-dimensional random vector of structural shocks with $\mathbb{E}\mathbf{u}_t = 0$ and $\mathbb{E}\mathbf{u}_t\mathbf{u}_t' = I_n$. A linear (or linearized) DSGE model can be written in the form:

$$\Gamma_0(\theta)\hat{\mathbf{z}}_\mathbf{t} = \Gamma_1(\theta)\mathbb{E}_t\hat{\mathbf{z}}_\mathbf{t+1} + \Gamma_2(\theta)\hat{\mathbf{z}}_\mathbf{t-1} + \Gamma_3(\theta)\mathbf{u_t} \tag{3.2}$$

where θ is the $q \times 1$ vector of structural parameters in which we are interested.

Assuming that a unique solution exists[12], it has the form:

$$\hat{\mathbf{z}}_\mathbf{t} = A(\theta)\hat{\mathbf{z}}_\mathbf{t-1} + B(\theta)\mathbf{u_t} \tag{3.3}$$

In most applications, some of the variables in $\hat{\mathbf{z}}_\mathbf{t}$ are not observed and the model in (3.3) cannot be taken directly to data. Let $\mathbf{x_t}$ be the l-dimensional vector of observed variables and C be a selection matrix, so that:

$$\mathbf{x_t} = C\hat{\mathbf{z}}_\mathbf{t} \tag{3.4}$$

Equations (3.3)(*transition equation*) and (3.4)(*measurement equation*) form what it is called state-space representation of the model.

Identification is not a new concept in econometrics so it is important to understand why classical results do not apply to DSGE models. As pointed out by Komunjer and Ng (2011), the main reason is that the rank conditions of Rothenberg (1971) assume (Assumption VIII) identifiability of the reduced form parameters of the state-space representation, one condition Komunjer and Ng (2011) show not to be generally satisfied by DSGE models.

Iskrev (2010b) establishes sufficient conditions to local identification of linear DSGE models from first and second moments of the observables. Let τ be a vector with elements $A(\theta), B(\theta)$ e z^*. Define:

$$Ex_t \quad = \quad \mu_x \tag{3.5}$$

[12] A survey of solution methods to linear rational expectation models can be found in Anderson (2008).

$$\begin{aligned} \mathrm{cov}(x_{t+i}, x') &= \Sigma_x(i) && (3.6) \\ \sigma_T &= [vech(\Sigma_x(0)', ..., vech(\Sigma(T-1))']' && (3.7) \\ m_T &= [\mu', \sigma_T{}'] && (3.8) \end{aligned}$$

Then m_T is the vector $(T-1)l^2 + l(l+3)/2$-dimensional collecting the parameters that determine the first two moments of the data.

Assuming the model is determined (that is, has a unique solution), $\tau(\theta)$ is unique for each admissible value of θ. It follows that $m(\theta)$ is a function of θ and θ_0 is locally identified if for any $\tilde{\theta}$ in a open neighborhood of θ_0:

$$m_T(\tilde{\theta}) = m_T(\theta_0) \Leftrightarrow \tilde{\theta} = \theta_0 \tag{3.9}$$

If the structural shocks are normally distributed, this condition is also necessary. Otherwise, the condition is only sufficient for identification from the first and second moments of data. Put in a more useful form, the condition requires that:

$$J(q) = \frac{\partial m_T}{\partial \tau'} \frac{\partial \tau}{\partial \theta'} = J_1(T) J_2 \tag{3.10}$$

be of *full column rank* (k) for some $q < T$.

Iskrev (2008) shows how to compute analytically these derivatives for the linear case.

The first term regards the effect of the reduced form parameters in the moments, while the second capture the effects of the deep parameters in the reduced form solution of the model. If J_2 has rank less than k at θ_0, then some of the parameters in θ are not locally identified, regardless of which data series are used in the estimation.

From a definition of observational equivalence based on the spectral density matrix, Komunjer and Ng (2011) establish a necessary and sufficient condition for the local identification of linear DSGE models. One difference from the work of Iskrev (2010b) is that they are able to treat both singular (model with fewer shocks than observables) and non-singular cases.

Iskrev (2010a) study the problem of weak identification in the context of ML estimation, defining a measure for the problem. In its own words (Iskrev, 2010a, p. 20):

> *Identification is weak when a parameter is either nearly irrelevant, because it does not affect much the likelihood, or nearly redundant, because the effect of the parameter on the likelihood can be replicated by other parameters.*

In accordance, they define the following measure of weak identification:

$$s_i(\theta) = \sqrt{\Delta_i(1 - \delta^2{}_i)} \tag{3.11}$$

where:

$$\Delta_i = E\left(\frac{\partial l_T(\theta)}{\partial \theta_i}\right)^2$$

measures the impact of the parameter θ_i in the likelihood keeping other parameters constant and:

$$\delta_i = corr\left(\frac{\partial l_T(\theta^0)}{\partial \theta_i}, \frac{\partial l_T(\theta^0)}{\partial \theta_{-i}}\right)$$

gauges how much the effect on the likelihood of changing θ_i can be offset by changing other parameters of the vector $\boldsymbol{\theta}$.

He shows that $1/s_i(\theta)$ is equal to the square root of i-th element of the inverse Fisher information matrix $\mathcal{I}_T{}^{-1}(\theta)$. For a scale-independent measure, he suggests:

$$s_i^r(\theta) = \sqrt{\theta_i \Delta_i(1 - \delta^2{}_i)} \tag{3.12}$$

3.2 Maximum Likelihood

Denote the past observations of $\mathbf{x_t}$ by $\aleph_{t-1} = (\mathbf{x_0}, \mathbf{x_1}, ..., \mathbf{x_{t-1}})$ and consider the state-space representation of the DSGE model given by equations (3.3)(*transition equation*) and (3.4)(*measurement equation*). Using the Kalman Filter, we obtain the linear projection $\hat{\tilde{\mathbf{z}}}_{\mathbf{t|t-1}}$ of $\hat{\mathbf{z}}_{\mathbf{t}}$ onto $\aleph_{t-1}$ and the error $\mathbf{P_{t|t-1}}$ associated with this projection.

Assuming that the vector $\mathbf{u_t}$ is i.i.d. and normally distributed, the density of $\mathbf{x_t}$ conditional to $\aleph_{t-1}$ is given by:

$$f(\mathbf{x_t}|\aleph_{t-1}, \theta) = N(C\tilde{\mathbf{z}}_{t|t-1}, CP_{t|t-1}C') \tag{3.13}$$

and the density of the whole sample of T observations is given by:

$$f(\aleph_T|\theta) = f(\mathbf{x_0}|\theta)\prod_{t=1}^{T} f(\mathbf{x_t}|\aleph_{t-1}, \theta) \tag{3.14}$$

The ML estimator is then defined as:

$$\hat{\theta}_{ML} = \arg\max_{\theta \in \Theta} L(\theta) \tag{3.15}$$

where $L(\theta)$ denotes the log-likelihood function:

$$L(\theta) = \sum_{i=1}^{T} -(T/2)ln(2\pi) - (1/2)ln|CP_{t|t-1}C'| - (1/2)(\mathbf{x_t} - C\tilde{\mathbf{z}}_{t|t-1})'(CP_{t|t-1}C')^{-1}(\mathbf{x_t} - C\tilde{\mathbf{z}}_{t|t-1}) \tag{3.16}$$

With some additional regularity conditions (see Theorem 3.3 of Newey and McFadden (1994) $\hat{\theta}_{ML}$ is consistent and:

$$\sqrt{T}(\hat{\theta}_{ML} - \theta) \to N\left(0, \mathcal{I}^{-1}\right) \tag{3.17}$$

where $\mathcal{I}$ is the information matrix given by:

$$\mathcal{I}(\theta) = -\operatorname{E}\left[\frac{\partial^2}{\partial\theta^2} L(\theta)\middle|\theta\right]$$

3.3 Simulated Method of Moments (SMM)

Let $\mathbf{m_t}$ be a p×1 vector with moments calculated using stationary and ergodic real data[13]. Let $\boldsymbol{m_i(\theta)}$ be the synthetic counterpart of $\mathbf{m_t}$, that is, a p×1 vector with moments calculated from data generated by the DSGE model with the $q \times 1$ parameter vector $\theta \in \Theta$, $\Theta \subset \Re^q$ a compact set. Given a sample of length T, we will denote τT the size of the artificial sample. Then the SMM estimator is defined as:

$$\hat{\theta}_{SMM} = \arg\min_{\theta\in\Theta} G(\theta)' W G(\theta) \tag{3.18}$$

where:

$$G(\theta) = (1/T)\sum_{i=1}^{T} \mathbf{m}_t - (1/\tau T)\sum_{i=1}^{\tau T} \boldsymbol{m}_i(\theta)$$

and W is the optimal weighting matrix, given by:

$$W = \lim_{T\to\infty} Var\left(\left(\frac{1}{\sqrt{T}}\right)\sum_{t=1}^{T}\mathbf{m}_t\right)^{-1}$$

Santos and Peralta-Alva (2005) show that even if $\boldsymbol{m_i(\theta)}$ is calculated from artificial data simulated from an approximation of the solution of the DSGE model, $E(\boldsymbol{m_i(\theta)}) \longrightarrow E(\mathbf{m_t})$ as the approximation errors of the computed solution go to zero and $\tau T \longrightarrow \infty$.

[13] The stationarity and ergodicity of x˙t may have been induced by a prior transformation of the raw data

Under further regularity conditions (see Duffie and Singleton (1993)) $\hat{\theta}_{SMM}$ is consistent and:

$$\sqrt{T}(\hat{\theta}_{SMM} - \theta) \rightarrow N\left(0, \left(1 + \frac{1}{\tau}\right)\left(D'W^{-1}D\right)^{-1}\right) \tag{3.19}$$

where:

$$D = E\left(\frac{\partial m_i(\theta)}{\partial \theta}\right)$$

with the derivatives numerically computed and the expectation approximated by the average over the simulated τT data points.

Note that when $\tau \rightarrow \infty$, the asymptotic variance of the SMM estimator converges to the asymptotic variance of the GMM estimator.

4 The Model

4.1 Smets and Wouters (2007)

Smets and Wouters (2007) (SW07) estimate a DSGE model for the US economy incorporating 4 nominal frictions (wage\price stickiness and indexation) and 5 real frictions (monopolistic competition in the labor and goods markets, investments adjustment costs, habit persistence and variable capacity utilization).

The economy features a deterministic growth path, 14 endogenous variables, 7 of which are observable, and 7 structural shocks. In all, the model has 41 structural parameters, 36 of which are estimated. The Appendix A describes the log-linearized model, its variables and parameters.

4.2 Known Identification problems

Originally, SW07 fix five parameters. They argue that two of them, the depreciation rate (δ) and the exogenous spending/GDP ratio (g_y) would be difficult to estimate, or weekly identified. The other three, the steady state mark-up in the labor market (ϕ_w) and the curvature parameters of the Kimball aggregators in both goods and labor markets (ϵ_p and ϵ_w) are fixed under the rationale of being unidentifiable.

Komunjer and Ng (2011) and Iskrev (2010b) study the identification of the parameters of the SW07 model from a classical perspective. They arrive at the same conclusion of linear dependence of the pairs ϵ_p and ζ_p and ϵ_w and ζ_w and consequently of the lack of their identification in the linearized model. The parameter ϕ_w is identified at the prior and posterior means, although

Iskrev (2010b) find 3 points (out of more than 900 thousand randomly draw) from the parameter space where it is not.

Iskrev (2010a) study the strength of identification of the parameters of SW07. He finds that, at the posterior mean, the worst identified parameters, the ones with the lowest values of his relative measure, are $\bar{l}, \beta, \rho_r, \sigma_l, \iota_p, \phi_w$, r_y, δ and ρ_b, with strong collinearity being the main culprit for ϕ_w[14] and low sensitivity of the likelihood for $\bar{l}, \beta, \rho_r, \sigma_l, \iota_p$. On the other hand, the best identified parameters, according to the same measure, are ρ_g, ρ_w, ρ_a, $\bar{\gamma}$, ρ, ρ_p, σ_g, μ_w and σ_r, that is, the deterministic trend and the parameters related to the dynamics of the shocks.

Caglar et al. (2012) analyze the identification of the parameters of the SW07 model using the Bayesian Learning Indicator, developed by Koop et al. (2011). The method explores the rate at which the posterior precision of a given parameter gets updated with the sample size, using simulated data. Koop et al. (2011) show that for identified parameters the posterior precision increases at the same rate as the sample size. But for parameters that are either not identified or weakly identified the posterior precision may be updated but its rate of update will be slower than the rate the sample size increases.

If Koop et al. (2011)'s test is taken in a strict sense, Caglar et al. (2012)'s results indicate that only 8 of the 36 parameters can said to be identified[15]. They are[16] $\bar{\gamma}, \rho_g, \rho_a, \rho_w, \mu_w, \rho_p, \sigma_I, \sigma_w$. On the other hand, the worst identified parameters[17] are $\bar{\pi}, r_\pi, r_{\Delta y}, \bar{l}, \beta, r_y, \sigma_l, \varphi$.

5 The Monte Carlo Experiments

In this chapter, we run Monte Carlo experiments to analyze the empirical distribution of the parameters' estimates of the SW07 model, under ML and SMM. In both procedures we follow SW07 and fix 5 of the 41 parameters at the same (true) values: $\delta = 0.025, g_y = 0.18, \phi_w = 1.5, \epsilon_p = 10, \epsilon_w = 10$. We then estimate the other 36 parameters arranged in the vector θ.

The following steps are conducted:

1. We solve the model with the fixed parameters and the vector θ at its posterior mean[18].

2. We then generate 500 artificial samples of size T from the solution of the model by simulating 1000+T periods initialized from the steady-state. The first 1000 observations

[14] Pairwise correlation analysis show a high (0.944) correlation with ζ_w.

[15] That is, only these parameters exhibit at least the expected reduction in the posterior variance, either when MCMC or the inverse hessian are used to obtain the posterior variances.

[16] They follow SW07 and fix the parameters $\delta, \phi_w, g_y, \epsilon_p$ and ϵ_w.

[17] Those that exhibit less than 1/4 of the expected reduction in the variance in both cases where MCMC and the inverse hessian are used to obtaining the posterior variances.

[18] Table A.1 in the appendix presents the bounds used and the posterior's means obtained in SW07 estimation.

are discarded as burn-ins. The innovations in the shock series are drawn from a normal distribution with mean zero and the posterior variance estimated by SW07.

3. Given the simulated data (and the calibrated parameters), we estimate θ by ML and SMM. The ML estimation is conducted entirely on Dynare[19]. The program for the SMM estimation is written in $Matlab^{®}$ and uses Dynare to solve the model.[20]. In both cases the optimization is constrained with the bound described in A.1 and, unless said otherwise, started from the true values[21]. For the SMM estimation,we use $\tau = 5$[22] and 70 moments: the means, covariances and autocorrelations up to fifth order of the 7 observable variables.

4. An estimation is conducted for each one of the 500 samples. The standard deviations are obtained from sample counterparts of (3.19) and (3.17) with the derivatives calculated numerically.

5.1 Correct Specification

In this experiment, we use data from the solution of the original (linear) SW07 model and estimate the same model. We consider sample sizes (T) of 100, 200 and 400 observations.[23]. We take into account only estimations that end satisfactorily. We define an estimation as satisfactory if $\boldsymbol{\mathcal{I}}$, in the ML case, or $\mathbf{D}'\mathbf{W}^{-1}\mathbf{D}$ in the SMM case, are positive definite matrices. This assures the standard deviations obtained will be real numbers. When that is not the case, we follow Adolfson and Lindé (2011) and drop the non-satisfactory optimization along with the sample associated with it[24].

Tables B.1 and B.2 in the Appendix B show statistics of the estimates' empirical distribution. Figure 5.1 below compare the density of the empirical distribution of ML and SMM estimates[25].

Visual inspection of the plots strongly suggests that the ML delivers better results than SMM. This is confirmed in the MSE metric: ML delivers smaller MSE for all parameters in all sample sizes. Weighting all parameters equally, the SMM estimates produce a MSE in average 22 (median 14.58) times larger than ML's for T=100. The biggest differences seems to come

[19] Adjemian et al. (2011). We opt for the CSMINWEL algorithm in the likelihood optimization.

[20] Programs available upon request. The optimization of the distance function is conducted with the function *fmincon* available in Matlab's Optimization Toolbox. In particular, after some tests evaluating both precision and speed we opt for the *Sequential Quadratic Programming* (SQP) algorithm.

[21] With this, we try to sidestep problems specific to the use of local optimizers, although their use can be justified based on the common practice of the profession.

[22] Results are unaffected when τ is set to 20.

[23] This sample sizes correspond to, respectively, 25, 50 and 100 years of quarterly data. In the original model, SW07 considered a sample of 39 years.

[24] As noted by Adolfson and Lindé (2011), in practice the econometrician would probably try to change the starting values or make other modifications in the model or estimation process. Our procedure circumvents the need of such arbitrary choices.

[25] We plot the kernel density estimates of the empirical distribution based on a normal kernel function, obtained with Matlab's *ksdensity* function.

from the estimates of the standard deviations of the shocks (all above the mean), very precisely estimated by ML. ML also does significantly better for ρ_g, ρ_w, $\bar{\gamma}$ and α.

The relative differences increase with sample size (except for $\bar{l}$). This happens due to faster convergence of ML estimates. While the average ratio of MSE for T=100 by the MSE for T=400 is 9.02 (median 4.38) for ML, it is 1.38 (median 0.98) for SMM. Considering only the standard deviation, the ratio decreases by a factor of average 2.52 (median 2.06) in the case of ML and 1.12 (median 0.99) in the case of SMM.

In the case of ML, biases decrease from T=100 to T=400 for all parameter but ψ, r_π and $r_{\Delta y}$. Standard-deviations decrease for all parameters, and for 25 of then the rate is greater than 2 (the theoretically expected). The same does not happen in the case of SMM. Biases do not decrease for 16 of the parameters and standard deviations does not decrease for 18 parameters.

Focusing in the case of T=100, the biggest biases in ML estimation occur in parameters $100(\beta^{-1} - 1)$, $\bar{\pi}$, $\bar{l}$ and σ_l. Most of SMM estimates have large biases. The few exception are: ι_p, ψ, r_π and $\bar{l}$.

The poor performance of SMM somewhat resembles the one obtained by Canova and Sala (2009) for Impulse-response matching. Specifically, in table 4 of Canova and Sala (2009) we see biases greater than 50% and some standard deviations increasing with the sample. On the other hand, it contrasts with Ruge-Murcia (2007)'s results, where SMM has a performance at least comparable to ML. One possible explanation is the deterioration of the SMM's performance with the increase in complexity of the model.

The parameters with the biggest coefficient of variation (with respect to the DGP) in the case of ML estimates are $\bar{l}$, $100(\beta^{-1} - 1)$, r_y and ρ_r . All four parameters pertain to the list of worst identified parameters of Iskrev (2010b) and three of them to the list of Caglar et al. (2012). In the case of SMM, the parameters with the greatest variability are $\bar{l}$, ρ_b, ι_p, σ_p and σ_l.

We now turn the analysis to asses whether the Asymptotic Standard Error (ASE), the ones obtained from (3.19) and (3.17) are correctly accounting for the sample uncertainty, judged by means of the Standard Deviation (SD) of the empirical distribution. Tables B.3 andB.4 in the appendix B display the results for each method.

For T=100, in the case of ML, the average ratio of SD/ASE is 1.12 (median, 1.03), with the main divergences being observed in the parameters ρ_a, ρ_g and ρ_w. The uncertainty is underestimated by the ASE for 24 out of the 36 parameters. In all cases, however, the discrepancies are substantially reduced when we increase the sample to T=400, with the average ratio SD/ASE dropping to 1.04 (median 1.02).

In the SMM case, the average SD/ASE ratio is 2.33 (median 1.2) for T=100 and 1.75 (1.01)

for T=400, strongly influenced by the results of the parameters $\bar{l}$, $\bar{\pi}$, $100(\beta^{-1}-1)$, $\bar{\gamma}$.

It is important to note that, from the shape of figure B.2 it is hard to justify an asymptotic normal approximation for inference in the case of SMM.

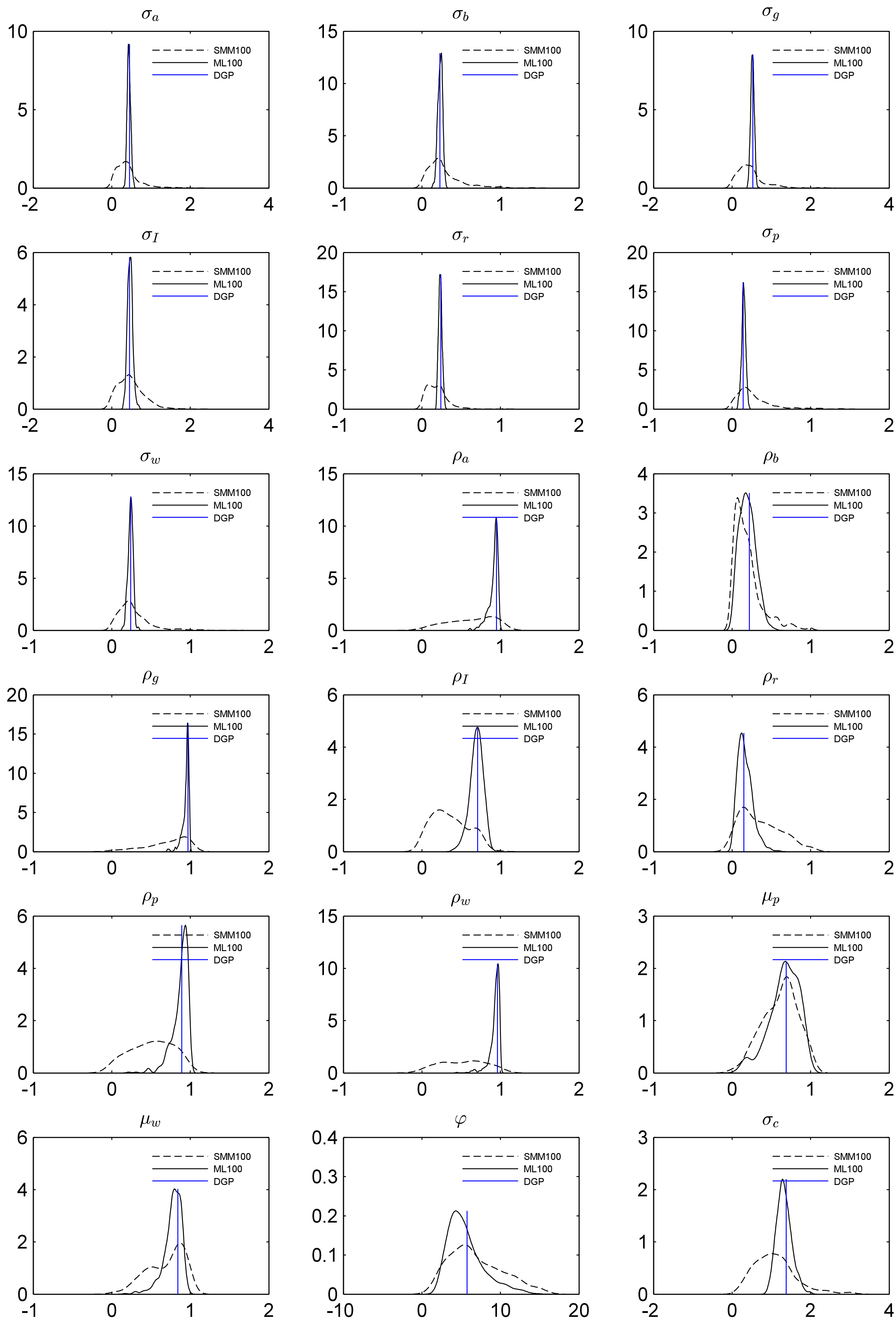

Figure 5.1: Kernel density estimates of the empirical distributions of the estimates for ML and SMM, T=100

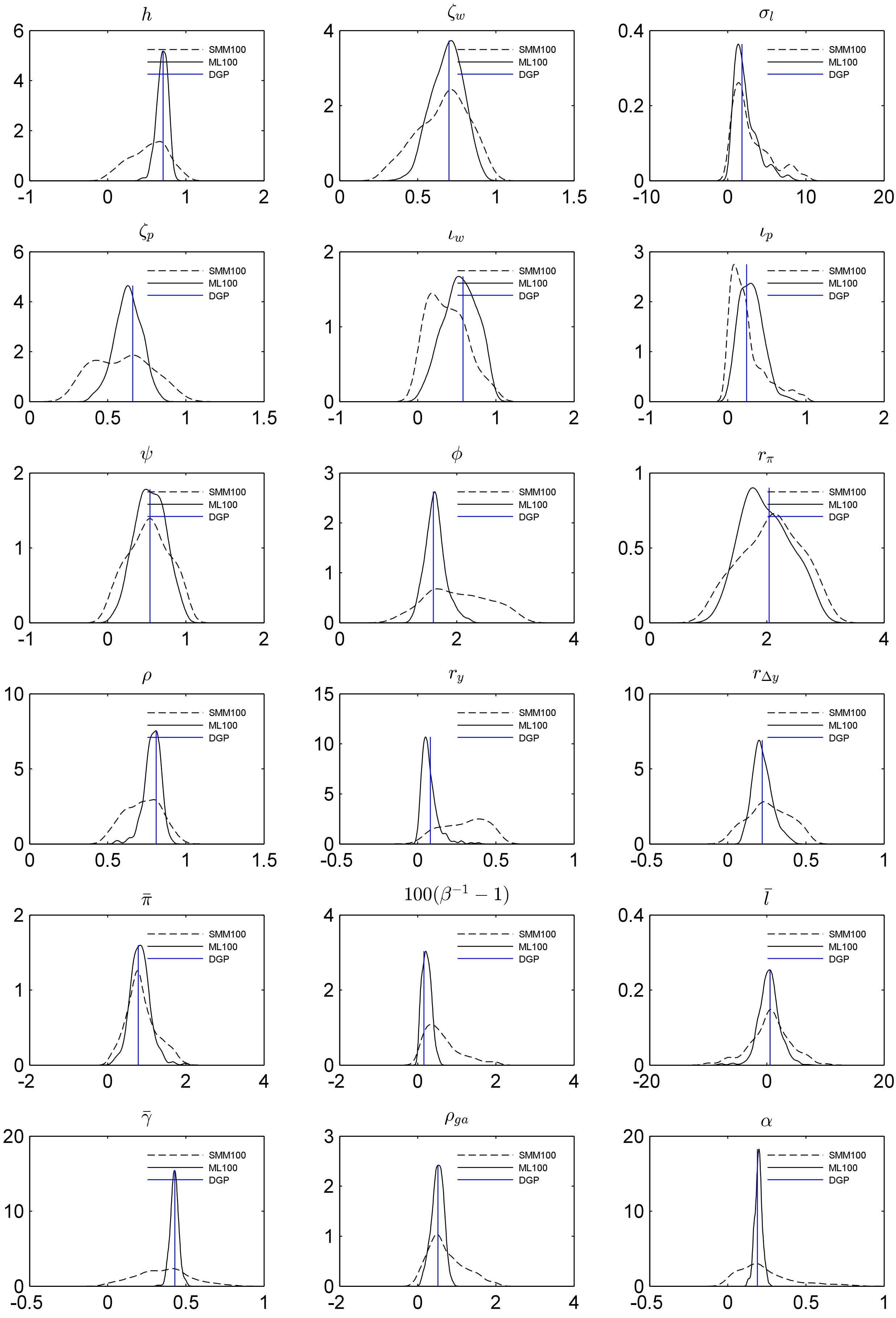

Figure 5.1: Kernel density estimates of the empirical distributions of the estimates for ML and SMM, T=100 (continued)

5.2 Misspecification: Second-order DGP

The question we are trying to answer in this experiment is: from a pure econometric sense - that is, ignoring possible new features that appear only in higher order terms - and considering the size of shocks usually estimated, what is the size of the error[26] incurred with linearization (or how does the order of approximation affects the parameter estimates)? And how does this error varies with the method used.

To that end, we generate 500 samples with T=100 from a nonlinear version of SW07 model, approximated up to the second order, and estimate a comparable linearized version of the model.

We begin describing the DGP. We try to follow SW07 as much as possible, but we are not able to replicate in the nonlinear setting two of its features: (i) the use of Kimball aggregators, (ii) and the particular structure of mark-up shocks. The nonlinear version thus is constructed with the Dixit-Stiglitz aggregator and with mark-up shocks in an *ad hoc* fashion as specified in equations (D.6) and (D.22).

While in the linear model we need only to specify the second derivatives of the cost functions for the capacity utilization and investment adjustment, in the nonlinear version we need to specify the whole functional form. We take these from Christiano et al. (2010).

Additionally, in the nonlinear model, the steady-state of labor is determined by the market clear condition in steady-state[27]. The appendix D describes the model equations.

Once, in the linear framework, the Dixit-Stiglitz aggregator is a special case of Kimball's when the Kimball's parameters are set to zero, for the estimated model we set $\epsilon_p = 0$ and $\epsilon_w = 0$ (their true values). We still estimate the steady-state labor, but with the optimization initialized in its new true value. The mark-up shocks are not changed[28].

Table B.5 and the figures B.3 and B.4 describe the estimates when the optimizer is started from the true value (the SW07's posterior estimates used in the DGP). We first focus on the ML estimates.

In the case of ML, we note that most of the shocks variances are overestimated, some (σ_b and σ_I) by more than 100%. Exceptions are σ_p and σ_w. The processes of the risk-premium (ϵ_b) and the investment-specific technology (ϵ_I) are estimated to be more persistent than they really are.

In the group of the structural parameters, stand out the biases of the response of the monetary policy to the output gap (r_y) and the share of the capital in production (α). The first one is overestimated and the second one (grossly) underestimated, with the median of the distribution

[26] In a study comparing the accuracy of different solution methods, (Aruoba et al., 2006, p. 2479) finds that "higher order perturbations display a much superior performance over linear methods (...)".

[27] See the appendix of (Del Negro and Schorfheide, 2012, p. 57).

[28] Thus providing an additional source of misspecification.

being its lower bound.

Turning to the SMM estimates, their main characteristic is that their median is exactly equal to the DGP. It seems that the initial point is in a *valley* of the misspecified distance function, so that the estimates remain in the initial point for more than half of the samples[29]. Even that being the case, the median of the ratios of MSE, ML over SMM is 0.25 (the mean is 1.01, strongly influenced by the parameters ρ_b and σ_g). That is, we cannot detect a superior performance of the moment based method in the case of misspecification.

In order to investigate the effect of the initial point in the estimates[30], we repeat the experiment initializing the optimizer from the prior means of SW07. The initial parameter vector is different from the true one, although not implausible or too distant. Table B.6 and figures B.5 and B.6 describe the empirical estimates.

The increased variability in ML and, especially, SMM is visible in the way less concentrated histograms of figures B.5 and B.6. Again, ML performs better than SMM. The median ratio of ML over SMM MSE is 0.55 (the average is 0.9) of the last. The SMM still performs much better in the cases of the parameters ρ_b and σ_g. Excluding these two, the average ratio falls to 0.67.

In the case of ML, most variance estimates continue to be overestimated. The process for the the risk-premium (ϵ_b) is now estimated to be even more persistent as is now the process for the monetary shock (ϵ_r). The reason seems to be the same: the estimator seems to be quite influenced by the starting value.

Some structural parameters are also strongly influenced by the starting value. The price indexation (ι_p), the parameters of the reaction of the monetary policy to the product gap (r_y and $r_{\Delta y}$) and the parameter related with the intertemporal discounting $100(\beta - 1)$, all move in the same direction as the starting value relative to the DGP.

Its important to note that when we conduct the estimation with the correct specification starting from the priors (results not reported), the same does not happen. On the other hand, despite the relatively large distance of the starting point from the DGP, the Calvo parameters for the prices and the wages seem to be very precisely estimated.

Turning to the SMM estimates, besides the above mentioned ρ_b and σ_g, the parameters r_π and $r_\Delta y$ are better estimated by SMM rather than by ML.

[29] A characteristic noticeable from the histograms in figure B.4.

[30] It is interesting to note that while starting from the DGP give us good excuse to use a local solver in the case of correct specification, the same does not happen in the case of misspecification. In this case, it is well possible that the global minimum be far from the DGP value. In our defense, we note that the common practice in DSGE estimation is the use of local solvers.

5.3 Misspecification: Estimated Model has i.i.d Shocks

As a reflection of the DSGE modeling strategy, the specification of the exogenous dynamics is usually arbitrary. The shocks are frequently assumed to follow an AR(1) process, but little is known about the consequences of this assumption in the estimation process.

In order to try to gain an intuition on the subject, we estimate a model where the dynamics of the shocks are suppressed, that is, the estimated model assumes that the shocks are i.i.d., when the DGP is the SW07 model [31].

With this exercise, we intend to gain insight for the general case of misspecification in the dynamics of the exogenous shocks, for example, restricting it to follow an AR(1) when its true dynamics is an AR(2), AR(3) or other ARMA process.

Histograms of the parameters are shown in figures B.7 and B.8. Table B.7 collects the statistics of the parameter estimates. We start with the ML estimates.

As expected, the variances of the shocks are all overestimated, with the exception of the one for the monetary policy shock (which has relatively little persistency in the DGP).

Turning to structural parameters, we see the price/wage stickiness and indexation parameter estimates being pressured upward, and the parameters σ_c and ψ being underestimated.

Contrary to the ML estimates, the effect on the SMM estimates for the variance of the shocks is ambiguous. The pattern for σ_c and ψ repeats.

Compared to ML estimates, the SMM ones present greater dispersion. The median ratio of ML over SMM MSE is 0.24, again indication of poor relative performance of SMM. The few exceptions (ratios above 1) are the standard deviation parameters other than σ_m and ξ_w.

[31] In the SW07 model, the shocks are endowed with a rich dynamic, with 5 of them following an AR(1) processes and the other 2 (the mark-up shocks) an ARMA (1,1).

5.4 Does it make a difference?

Economic models will always be incomplete. The hope of the researcher is that the features omitted in its simplification of reality are not relevant for the analysis of the problem at hand.

In this section we assess to what degree this hope is justified when the simplification is attained by means of (1) linearization of the model and (2) incomplete modeling of the dynamics of the shocks.

In the first case, the researcher is not interested in the properties of higher order terms of its model and thus conscientiously solves its model up to the first order.

As models are usually thought in nonlinear terms, in his or her own view of the DGP there are high order forces driving the dynamics. The estimations of the linearized form hence imbeds the expectations that the high order terms ignored will not bias the estimation of the parameters present in the linear approximation.

Although this expectation is reasonable in an arbitrarily small neighborhood of the steady-state[32], its validity is generally taken as given for the size of shocks usually estimated.

In the second case of misspecifications, the researcher is careless with the specification of the dynamics of the exogenous shocks, hoping it does not affect the parameters associated with the endogenous propagation mechanisms.

Our experiment evaluates the consequences of using parameters estimated with misspecified models in two of the most common exercises in the DSGE literature: impulse-response analysis and forecast error variance decomposition. Due to the poor performance of SMM, we focus only on ML.

In both cases of misspecification, the model used for the computation of impulse-responses and variance decomposition is the same used for the estimation of the parameters. With that, we try to replicate the situation where the researcher estimates a misspecified model and uses it in the analysis of an economic problem. As we aim at the problem arising from the bias of the estimates, our benchmark is the misspecified model with the DGP (correct) parameters.

In subsection C.1 in the appendix, we measure the bias in using the median parameter estimates of subsection 5.2[33] in impulse-response and variance decomposition analysis. The model used is the linear version of SW07's model, restricted to use the Dixit-Stiglitz aggregator.

Looking at the impulse-response functions, we see a pattern of more intense responses, as a reflection of the generally overestimated variances of the shocks. Other than that, only the responses to ε_i seem to be qualitatively different from the benchmark. This behavior could be

[32] Although not yet rigorously confirmed, as show in the controversy of Fernández-Villaverde et al. (2006) and Ackerberg et al. (2009)

[33] We use the estimates obtained starting from the SW07 posteriors (the DGP).

attributed to the underestimated α.

Turning to the decomposition of the errors variance, what we see is the distortion caused by the overestimated variances of ε_g and ε_b. The effect of the overestimated variance ε_i is inhibited by the underestimated α, that prevents the propagation of ε_i to variables other than di.

In subsection C.2 in the appendix, we measure the bias of using the median parameter estimates of subsection 5.3 in impulse-response and variance decomposition analysis. The model used is the linear version of SW07's model restricted to have i.i.d. shocks.

Compared to approximation order, the effects of misspecification in the dynamics of the shocks seem to have a greater impact in the responses of the system.

When compared to the DGP parametrization, the log-difference of investment has a stronger reaction to ε_a and ε_g, and exhibits a different dynamics in response to ε_w. These results could be related to the underestimation of the capital utilization adjustment cost function parameter (ψ).

The same is true for the responses of log hour worked to the shocks in $\varepsilon_w, \varepsilon_\pi$ and ε_m, a result most probably driven be the overestimation of the wage stickiness parameter ξ_w from 0.7 to 0.95.

The responses of π and dw are generally much weaker than in the DGP parametrization (the exceptions being the response to ε_π and ε_w), with the initial response to ε_i in the case of π and ε_a in the case of dw even changing sign. This result can be attributed to the positive bias in the stickiness/indexation parameters.

The dynamic of r significantly changes only in response to ε_w, being much stronger with the biased parameters.

In the variance decomposition exercise, we observe the wage shock gaining dramatic importance in the explanation of l, the importance of the monetary policy being substantially reduced in the composition of inflation and the importance of the inflation shock gaining participation in the dynamics of dw. The interest rate is now majorly driven by shocks other than ε_m, with greater importance acquired by ε_w and ε_p.

6 Conclusion

In this paper we conducted a Monte Carlo study of the Smets and Wouters (2007) (SW07) model, analyzing the small sample properties of two different procedures: Maximum Likelihood (ML) and Simulated Method of Moments (SMM). We paid special attention to the implications of the two types of misspecification in impulse-response and variance forecast decomposition analyzes.

In the first case of misspecification, the estimated model is a linear version while the DGP is a second order approximation of the SW07 model. In the second case, the DGP is the SW07 model whereas the estimated model is the SW07 model restricted to have i.i.d. shocks.

Our first result is the poor performance of SMM relative to ML in both cases correct specification and misspecification. A good portion of SMM estimates have relatively large biases, with dispersed distributions that do not seem to collapse as the sample increase. This result resembles the one of Canova and Sala (2009) in the context of Impulse-response matching, and suggest that Ruge-Murcia (2007)'s results does not extend to larger models.

The ML estimates on the other hand confirm the results of Adolfson and Lindé (2011) in that biases are small and the MSE decrease with sample size in the case of correct specification.

Focusing in the ML estimates, we examine the effects of using estimates from the misspecified models in impulse-response analysis and variance decomposition.

In both cases of misspecification, the variances of the shocks are overestimated. In the impulse-responses, this translates in a shift in scale of the responses.

In the case of misspecification where the DGP is nonlinear, the most striking result is the underestimation of the parameter α, which results in large bias in the responses following the investment technology shock ε_i. Together with the overestimated variances for the shocks, the final effect in the business cycle accounting is a substantial increase in the participation of the risk premium (ε_b) and government spending shocks (ε_g).

In the case of misspecification where the estimated model is restricted to have i.i.d shocks, the bias in the estimation of the price/wage stickiness/indexation ($\zeta_p, \zeta_w, \iota_p$ and ι_w) and capital utilization adjustment cost function (ψ) parameters result in more pronounced changes in the dynamics of the observables variables. In variance decomposition, the price and wage shocks gain much more influence.

References

Ackerberg, D., Geweke, J., and Hahn, J. (2009). Comments on "convergence properties of the likelihood of computed dynamic models". *Econometrica*, 77(6):2009–2017.

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

Adolfson, M., Laséen, S., Lindé, J., and Villani, M. (2007). Evaluating an estimated new keynesian small open economy model. Working Paper Series 203, Sveriges Riksbank (Central Bank of Sweden).

Adolfson, M. and Lindé, J. (2011). Parameter identification in a estimated new keynesian open economy model. Working Paper Series 251, Sveriges Riksbank (Central Bank of Sweden).

An, S. and Schorfheide, F. (2007). Bayesian analysis of dsge models. *Econometric Reviews*, 26(2-4):211–219.

Anderson, G. (2008). Solving linear rational expectations models: A horse race. *Computational Economics*, 31(2):95–113.

Aruoba, S. B., Fernandez-Villaverde, J., and Rubio-Ramirez, J. F. (2006). Comparing solution methods for dynamic equilibrium economies. *Journal of Economic Dynamics and Control*, 30(12):2477–2508.

Caglar, E., Chadha, J. S., and Shibayama, K. (2012). Bayesian estimation of dsge models: Is the workhorse model identified? Koç University-TUSIAD Economic Research Forum Working Papers 1205, Koc University-TUSIAD Economic Research Forum.

Canova, F. and Sala, L. (2009). Back to square one: Identification issues in dsge models. *Journal of Monetary Economics*, 56(4):431–449.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45.

Christiano, L. J., Trabandt, M., and Walentin, K. (2010). Dsge models for monetary policy analysis. In Friedman, B. M. and Woodford, M., editors, *Handbook of Monetary Economics*, volume 3 of *Handbook of Monetary Economics*, chapter 7, pages 285–367. Elsevier.

Del Negro, M. and Schorfheide, F. (2012). Dsge model-based forecasting. Technical report.

Duffie, D. and Singleton, K. J. (1993). Simulated moments estimation of markov models of asset prices. *Econometrica*, 61(4):929–52.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., and Santos, M. S. (2006). Convergence properties of the likelihood of computed dynamic models. *Econometrica*, 74(1):93–119.

Iskrev, N. (2008). Evaluating the information matrix in linearized dsge models. *Economics Letters*, 99(3):607–610.

Iskrev, N. (2010a). Evaluating the strength of identification in dsge models. an a priori approach. Technical report.

Iskrev, N. (2010b). Local identification in dsge models. *Journal of Monetary Economics*, 57(2):189–202.

Kim, J. and Ruge-Murcia, F. J. (2009). How much inflation is necessary to grease the wheels? *Journal of Monetary Economics*, 56(3):365–377.

Kim, J. and Ruge-Murcia, F. J. (2011). Monetary policy when wages are downwardly rigid: Friedman meets tobin. *Journal of Economic Dynamics and Control*, 35(12):2064–2077.

Komunjer, I. and Ng, S. (2011). Dynamic identification of dynamic stochastic general equilibrium models. *Econometrica*, 79(6):1995–2032.

Koop, G., Pesaran, M. H., and Smith, R. P. (2011). On identification of bayesian dsge models. IZA Discussion Papers 5638, Institute for the Study of Labor (IZA).

Newey, W. K. and McFadden, D. (1994). Chapter 36 large sample estimation and hypothesis testing. volume 4 of *Handbook of Econometrics*, pages 2111 – 2245. Elsevier.

Ríos-Rull, J.-V., Schorfheide, F., Fuentes-Albero, C., Kryshko, M., and Santaeulàlia-Llopis, R. (2012). Methods versus substance: Measuring the effects of technology shocks. *Journal of Monetary Economics*, 59(8):826 – 846.

Rotemberg, J. and Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy. In *NBER Macroeconomics Annual 1997, Volume 12*, NBER Chapters, pages 297–361. National Bureau of Economic Research, Inc.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica*, 39(3):577–91.

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.

Ruge-Murcia, F. J. (2007). Methods to estimate dynamic stochastic general equilibrium models. *Journal of Economic Dynamics and Control*, 31(8):2599–2636.

Santos, M. S. and Peralta-Alva, A. (2005). Accuracy of simulations for stochastic dynamic models. *Econometrica*, 73(6):1939–1976.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606.

A The Smets and Wouters (2007) model

A.1 Model Equation

The log-linearized model consists of:

$$
\begin{aligned}
y_t &= c_y c_t + i_y i_t + \bar{r}^k k_y z_t + \varepsilon_t^a & \text{(A.1)}\\
c_t &= \frac{\lambda/\gamma}{1+\lambda/\gamma} c_{t-1} + \frac{1}{1+\lambda/\gamma} E_t c_{t+1} + \frac{\bar{W}\bar{l}(\sigma_c - 1)}{\bar{C}\sigma_c(1+\lambda/\gamma)}(l_t - E_t l_{t+1}) - \frac{1-\lambda/\gamma}{(1+\lambda/\gamma)\sigma_c}(r_t - E_t\pi_{t+1}) \\
&\quad - \frac{1-\lambda/\gamma}{(1+\lambda/\gamma)\sigma_c}\varepsilon_t^b & \text{(A.2)}\\
i_t &= \frac{1}{1+\beta\gamma^{(1-\sigma_c)}} i_{t-1} + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}} E_t i_{t+1} + \frac{1}{\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})} q_t + \varepsilon_t^i & \text{(A.3)}\\
q_t &= \beta(1-\delta)\gamma^{-\sigma_c} E_t q_{t+1} - (1-\beta(1-\delta)\gamma^{-\sigma_c}) E_t r_{t+1}^k - (r_t - E_t\pi_{t+1} + \varepsilon_t^b) & \text{(A.4)}\\
y_t &= \phi_p(\alpha k_t^s + (1-\alpha) l_t + \varepsilon_t^a) & \text{(A.5)}\\
k_t^s &= k_{t-1} + z_t & \text{(A.6)}\\
z_t &= \frac{1-\psi}{\psi} r_t^k & \text{(A.7)}\\
k_t &= \frac{(1-\delta)}{\gamma} k_{t-1} + \frac{\gamma-(1-\delta)}{\gamma} i_t + \frac{\gamma-(1-\delta)}{\gamma}\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})\varepsilon_t^p & \text{(A.8)}\\
\mu_t^p &= \alpha(k_t^s - l_t) - w_t + \varepsilon_t^a & \text{(A.9)}\\
\pi_t &= \frac{\iota_p}{1+\iota_p\beta\gamma^{(1-\sigma_c)}}\pi_{t-1} + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\iota_p\beta\gamma^{(1-\sigma_c)}} E_t\pi_{t+1} \\
&\quad + \frac{(1-\beta\gamma^{(1-\sigma_c)}\zeta_p)(1-\zeta_p)}{(1+\iota_p\beta\gamma^{(1-\sigma_c)})(1+(\phi_p-1)\epsilon_p)\zeta_p}\mu_t^p + \varepsilon_t^p & \text{(A.10)}\\
r_t^k &= l_t + w_t - k_t & \text{(A.11)}\\
\mu_t^w &= w_t - (\sigma_l l_t \frac{1}{1-\lambda/\gamma}(c_t - \lambda/\gamma c_{t-1})) & \text{(A.12)}\\
w_t &= \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}(w_{t-1} + \iota_w\pi_{t-1}) + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}}(E_t w_{t+1} + E_t\pi_{t+1}) \\
&\quad - \frac{(1-\beta\gamma^{(1-\sigma_c)}\zeta_w)(1-\zeta_w)}{(1+\beta\gamma^{(1-\sigma_c)})(1+(\phi_w-1)\epsilon_w)\zeta_w}\mu_t^w + \varepsilon_t^w & \text{(A.13)}\\
r_t &= \rho_{t-1} r_{t-1} + (1-\rho)[r_\pi\pi_t + r_Y(y_t - y_t^f)] + r_{\Delta y}[(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)] + \varepsilon_t^r & \text{(A.14)}\\
\varepsilon_t^a &= \rho_a\varepsilon_{t-1}^a + \eta_t^a & \text{(A.15)}\\
\varepsilon_t^b &= \rho_b\varepsilon_{t-1}^b + \eta_t^b & \text{(A.16)}\\
\varepsilon_t^g &= \rho_g\varepsilon_{t-1}^g + \rho_{ga}\eta_t^a + \eta_t^g & \text{(A.17)}\\
\varepsilon_t^i &= \rho_i\varepsilon_{t-1}^i + \eta_t^i & \text{(A.18)}\\
\varepsilon_t^r &= \rho_r\varepsilon_{t-1}^r + \eta_t^r & \text{(A.19)}\\
\varepsilon_t^p &= \rho_p\varepsilon_{t-1}^p + \eta_t^p - \mu_p\eta_{t-1}^p & \text{(A.20)}
\end{aligned}
$$

$$\varepsilon_t^w = \rho_w \varepsilon_{t-1}^w + \eta_t^w - \mu_w \eta_{t-1}^w \tag{A.21}$$

$$\tag{A.22}$$

where, y_t^f, is the potential output, the output of the *flexible* economy, defined as being the analogue of the original without price or wage rigidities and in the absence of price and wage mark-up shocks. The model equations of the flexible economy are:

$$y_t^f = c_y c_t^f + i_y i_t^f + \bar{r}^k k_y z_t^f + \varepsilon_t^g \tag{A.23}$$

$$c_t^f = \frac{\lambda/\gamma}{1+\lambda/\gamma} c_{t-1}^f + \frac{1}{1+\lambda/\gamma} E_t c_{t+1}^f + \frac{\bar{W}\bar{l}(\sigma_c - 1)}{\bar{C}\sigma_c(1+\lambda/\gamma)} - \frac{1-\lambda/\gamma}{(1+\lambda/\gamma)\sigma_c}(r_t - E_t \pi_{t+1}) - \frac{1-\lambda/\gamma}{(1+\lambda/\gamma)\sigma_c}\varepsilon_t^b \tag{A.24}$$

$$i_t^f = \frac{1}{1+\beta\gamma^{(1-\sigma_c)}} i_{t-1}^f + \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}} E_t i_{t+1}^f + \frac{1}{\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})} q_t^f + \varepsilon_t^i \tag{A.25}$$

$$q_t^f = \beta(1-\delta)\gamma^{-\sigma_c} E_t q_{t+1}^f - (1\beta(1-\delta)\gamma^{-\sigma_c}) E_t r_{t+1}^{kf} - r_t^f + \varepsilon_t^b) \tag{A.26}$$

$$y_t^f = \phi_p(\alpha k_t^{sf} + (1-\alpha) l_t^f + \varepsilon_t^a) \tag{A.27}$$

$$k_t^{sf} = k_{t-1}^f + z_t^f \tag{A.28}$$

$$z_t^f = \frac{1-\psi}{\psi} r_t^{kf} \tag{A.29}$$

$$k_t^f = \frac{(1-\delta)}{\gamma} k_{t-1}^f + \frac{\gamma-(1-\delta)}{\gamma} i_t^* + \frac{\gamma-(1-\delta)}{\gamma}\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})\varepsilon_t^p \tag{A.30}$$

$$\mu_t^{pf} = \alpha(k_t^s - l_t) - w_t + \varepsilon_t^a \tag{A.31}$$

$$\mu_t^{pf} = 0 \tag{A.32}$$

$$r_t^{kf} = l_t^f + w_t^f - k_t^f \tag{A.33}$$

$$w_t^f = \sigma_l l_t^f + \frac{1}{1-\lambda/\gamma}(c_t^f - \lambda/\gamma c_{t-1}^f) \tag{A.34}$$

$$\tag{A.35}$$

A.2 SW07 Bounds and Posteriors

Table A.1: SW07 Bounds and Posteriors

Parameter	**LB**	**UB**	**Posterior**
σ_a	0.01	3	0.45
σ_b	0.025	5	0.23
σ_g	0.01	3	0.53
σ_I	0.01	3	0.45
σ_r	0.01	3	0.24
σ_p	0.01	3	0.14
σ_w	0.01	3	0.24
ρ_a	0.01	0.9999	0.95
ρ_b	0.01	0.9999	0.22
ρ_g	0.01	0.9999	0.97
ρ_I	0.01	0.9999	0.71
ρ_r	0.01	0.9999	0.15
ρ_p	0.01	0.9999	0.89
ρ_w	0.001	0.9999	0.96
μ_p	0.01	0.9999	0.69
μ_w	0.01	0.9999	0.84
φ	2	15	5.74
σ_c	0.25	3	1.38
h	0.001	0.99	0.71
ζ_w	0.3	0.95	0.7
σ_l	0.25	10	1.83
ζ_p	0.3	0.95	0.66
ι_w	0.01	0.99	0.58
ι_p	0.01	0.99	0.24
ψ	0.01	1	0.54
ϕ	1	3	1.6
r_π	1	3	2.04
ρ	0.5	0.975	0.81
r_y	0.001	0.5	0.08
$r_{\Delta y}$	0.001	0.5	0.22
$\bar{\pi}$	0.1	2	0.78
$100(\beta^{-1} - 1)$	0.01	2	0.16
$\bar{l}$	-10	10	0.53
$\bar{\gamma}$	0.01	0.8	0.43
ρ_{ga}	0.01	2	0.52
α	0.01	1	0.19

A.3 Estimated Parameters Description

Table A.2: Estimated Parameters Description

Symbol	Description
σ_a	Sd. of the TFP innovation
σ_b	Sd. of the innovation in the wedge between the interest rate controlled by the central bank and the return on assets held by the household
σ_g	Sd. of the innovation of exogenous spending
σ_I	Sd. of the innovation to the investment-specific technology
σ_r	Sd. of the innovation of the monetary policy shock
σ_p	Sd. of the innovation of the price inflation shock
σ_w	Sd. of the innovation of the wage inflation shock
ρ_a	AR parameter of the tecnological shock
ρ_b	AR parameter of the wedge between the interest rate controlled by the central bank and the return on assets held by the household
ρ_g	AR parameter of the exogenous spending
ρ_I	AR parameter of the investment-specific technology
ρ_r	AR parameter of the monetary policy shock
ρ_p	AR parameter of the price inflation
ρ_w	AR parameter of the wage inflation
μ_p	MA parameter of the price inflation
μ_w	MA parameter of the wage inflation
φ	Steady-state elasticity of the capital adjustment cost function
σ_c	Elasticity of intertemporal substitution of consumption
h	External habit formation
ζ_w	Wage stickiness
σ_l	Elasticity of labor supply with respect to the real wage
ζ_p	Price stickiness
ι_w	Wage indexation
ι_p	Price indexation
ψ	Parameter of the capital utilization adjustment cost function
ϕ_p	1+ Share of fixed costs in steady-state production
r_π	Parameter of price inflation in the Taylor rule
ρ	AR parameter of the Taylor rule
r_y	Parameter of the output gap in the Taylor rule
$r_{\Delta y}$	Parameter of the variation of the output gap in the Taylor rule
$\bar{\pi}$	Steady-state price inflation
$100 * (\beta^{-1} - 1)$	Function of the intertemporal discont factor
$\bar{l}$	Labor in steady-state
$\bar{\gamma}$	Common quarterly trend growth rate
ρ_{ga}	Corr. Between the productivity and expenditure shocks
α	Share of capital in production

Table A.3: Variables Description

Variable	Type	Observable	Description[34]
y_t	End.	Yes	Output
c_t	End.	Yes	Consumption
i_t	End.	Yes	Investment
q_t	End.	No	Value of Capital
k_t^s	End.	No	Utilized Capital
k_t	End.	No	Installed Capital
z_t	End.	No	Capacity Utilization
r_t^k	End.	No	Rental rate of capital
μ_t^p	End.	No	Price Markup
μ_t^w	End.	No	Wage Markup
π_t^p	End.	Yes	Price Inflation
w_t	End.	Yes	Real wage
l_t	End.	Yes	Hours Worked
r_t	End.	Yes	Nominal interest rate
ε_t^a	Exo. Shock	No	Total factor productivity
ε_t^i	Exo. Shock	No	Investment specific tecnology
ε_t^b	Exo. Shock	No	Risk premium
ε_t^g	Exo. Shock	No	Exogenous spending
ε_t^p	Exo. Shock	No	Monetary policy
ε_t^w	Exo. Shock	No	Wage markup
ε_t^r	Exo. Shock	No	Price markup

[34] All variables are log-deviations from steady-state

B Results from Monte-Carlo Experiments

B.1 Correct Specification

Table B.1: Empirical Distribution of ML Estimates

Parameter	**DGP**	**Samples**								
		100			**200**			**400**		
		Mean	*Median*	*MSE*	*Mean*	*Median*	*MSE*	*Mean*	*Median*	*MSE*
σ_a	0.45	0.437	0.433	0.002	0.447	0.444	0.001	0.447	0.446	0.000
σ_b	0.23	0.232	0.235	0.001	0.232	0.232	0.000	0.231	0.231	0.000
σ_g	0.53	0.519	0.518	0.002	0.526	0.525	0.001	0.528	0.528	0.000
σ_I	0.45	0.465	0.463	0.004	0.465	0.463	0.002	0.456	0.459	0.001
σ_r	0.24	0.234	0.234	0.001	0.239	0.239	0.000	0.240	0.240	0.000
σ_p	0.14	0.147	0.147	0.001	0.143	0.144	0.000	0.140	0.140	0.000
σ_w	0.24	0.239	0.241	0.001	0.240	0.241	0.001	0.241	0.240	0.000
ρ_a	0.95	0.911	0.932	0.006	0.939	0.943	0.001	0.945	0.947	0.000
ρ_b	0.22	0.195	0.189	0.011	0.200	0.197	0.007	0.212	0.211	0.003
ρ_g	0.97	0.944	0.959	0.003	0.962	0.966	0.000	0.967	0.968	0.000
ρ_I	0.71	0.694	0.697	0.007	0.695	0.700	0.004	0.706	0.709	0.002
ρ_r	0.15	0.166	0.150	0.009	0.148	0.143	0.005	0.143	0.143	0.003
ρ_p	0.89	0.873	0.901	0.012	0.885	0.896	0.004	0.887	0.891	0.001
ρ_w	0.96	0.929	0.948	0.006	0.947	0.955	0.001	0.954	0.956	0.000
μ_p	0.69	0.661	0.680	0.037	0.668	0.696	0.023	0.667	0.680	0.013
μ_w	0.84	0.776	0.795	0.017	0.815	0.826	0.005	0.831	0.834	0.002
φ	5.74	5.574	5.063	5.262	5.850	5.461	3.737	5.657	5.470	1.598
σ_c	1.38	1.327	1.305	0.038	1.372	1.359	0.019	1.383	1.381	0.010
h	0.71	0.700	0.703	0.005	0.700	0.702	0.003	0.702	0.704	0.001
ζ_w	0.7	0.690	0.696	0.010	0.698	0.697	0.005	0.700	0.700	0.003
σ_l	1.83	2.294	1.848	2.627	2.159	1.893	1.715	1.963	1.818	0.631
ζ_p	0.66	0.631	0.631	0.008	0.644	0.645	0.003	0.651	0.651	0.001
ι_w	0.58	0.542	0.546	0.046	0.560	0.569	0.032	0.579	0.582	0.015
ι_p	0.24	0.299	0.288	0.026	0.256	0.256	0.014	0.237	0.239	0.008
ψ	0.54	0.544	0.542	0.037	0.560	0.559	0.015	0.556	0.558	0.006
ϕ_p	1.6	1.632	1.623	0.031	1.615	1.614	0.014	1.608	1.602	0.006
r_π	2.04	1.993	1.958	0.173	2.072	2.038	0.102	2.109	2.092	0.079
ρ	0.81	0.786	0.790	0.004	0.803	0.804	0.001	0.811	0.811	0.001
r_y	0.08	0.074	0.061	0.003	0.080	0.075	0.001	0.083	0.080	0.001
$r_{\Delta y}$	0.22	0.221	0.212	0.004	0.227	0.224	0.002	0.225	0.223	0.001
$\bar{\pi}$	0.78	0.832	0.826	0.070	0.784	0.788	0.031	0.780	0.784	0.014
$100(\beta^{-1}-1)$	0.16	0.229	0.222	0.018	0.180	0.178	0.008	0.165	0.163	0.005
$\bar{l}$	0.53	0.384	0.389	3.250	0.649	0.619	1.485	0.514	0.471	0.621
$\bar{\gamma}$	0.43	0.427	0.427	0.001	0.430	0.430	0.000	0.430	0.430	0.000
ρ_{ga}	0.52	0.508	0.516	0.024	0.518	0.523	0.009	0.523	0.524	0.005
α	0.19	0.196	0.197	0.001	0.193	0.193	0.000	0.192	0.192	0.000

Table B.2: Empirical Distribution of SMM Estimates

Parameter	DGP	Samples								
		100			**200**			**400**		
		Mean	*Median*	*MSE*	*Mean*	*Median*	*MSE*	*Mean*	*Median*	*MSE*
σ_a	0.45	0.399	0.344	0.099	0.347	0.312	0.083	0.343	0.312	0.078
σ_b	0.23	0.277	0.230	0.050	0.228	0.195	0.041	0.214	0.210	0.019
σ_g	0.53	0.460	0.398	0.129	0.437	0.422	0.094	0.429	0.427	0.076
σ_I	0.45	0.479	0.450	0.110	0.455	0.450	0.106	0.442	0.450	0.070
σ_r	0.24	0.186	0.170	0.022	0.166	0.158	0.020	0.160	0.144	0.021
σ_p	0.14	0.271	0.210	0.073	0.217	0.169	0.041	0.201	0.178	0.026
σ_w	0.24	0.273	0.238	0.044	0.232	0.226	0.025	0.241	0.240	0.019
ρ_a	0.95	0.623	0.662	0.189	0.595	0.661	0.219	0.644	0.709	0.176
ρ_b	0.22	0.194	0.147	0.033	0.164	0.109	0.033	0.158	0.105	0.035
ρ_g	0.97	0.709	0.788	0.138	0.735	0.827	0.126	0.766	0.858	0.106
ρ_I	0.71	0.360	0.325	0.175	0.304	0.236	0.221	0.320	0.280	0.208
ρ_r	0.15	0.362	0.320	0.112	0.397	0.369	0.136	0.419	0.397	0.146
ρ_p	0.89	0.510	0.520	0.216	0.479	0.490	0.245	0.466	0.491	0.266
ρ_w	0.96	0.520	0.556	0.274	0.515	0.520	0.288	0.526	0.540	0.274
μ_p	0.69	0.611	0.651	0.055	0.649	0.690	0.071	0.693	0.743	0.059
μ_w	0.84	0.686	0.758	0.079	0.691	0.801	0.095	0.711	0.819	0.090
φ	5.74	6.885	6.001	12.460	6.137	5.781	6.996	5.931	5.740	5.281
σ_c	1.38	1.091	1.032	0.366	1.110	1.026	0.416	1.068	1.016	0.425
h	0.71	0.524	0.549	0.092	0.525	0.581	0.101	0.544	0.579	0.097
ζ_w	0.7	0.663	0.699	0.027	0.667	0.700	0.036	0.673	0.700	0.036
σ_l	1.83	2.975	1.958	7.151	2.513	1.654	5.872	1.967	1.438	3.189
ζ_p	0.66	0.593	0.608	0.036	0.611	0.629	0.041	0.618	0.627	0.041
ι_w	0.58	0.380	0.357	0.101	0.369	0.316	0.119	0.348	0.300	0.122
ι_p	0.24	0.240	0.175	0.049	0.227	0.143	0.056	0.241	0.164	0.056
ψ	0.54	0.524	0.540	0.069	0.606	0.623	0.081	0.571	0.564	0.082
ϕ_p	1.6	2.024	1.969	0.457	2.079	2.007	0.557	2.119	2.127	0.589
r_π	2.04	2.070	2.095	0.259	2.079	2.055	0.300	2.152	2.173	0.314
ρ	0.81	0.724	0.725	0.020	0.729	0.739	0.021	0.727	0.743	0.023
r_y	0.08	0.298	0.316	0.067	0.322	0.352	0.079	0.334	0.373	0.087
$r_{\Delta y}$	0.22	0.264	0.255	0.019	0.285	0.287	0.022	0.297	0.302	0.025
$\bar{\pi}$	0.78	0.861	0.780	0.157	0.834	0.816	0.126	0.827	0.788	0.123
$100(\beta^{-1}-1)$	0.16	0.621	0.506	0.420	0.526	0.419	0.301	0.551	0.448	0.344
$\bar{l}$	0.53	0.543	0.530	12.100	0.429	0.482	7.321	0.500	0.472	1.695
$\bar{\gamma}$	0.43	0.352	0.359	0.035	0.364	0.385	0.032	0.378	0.393	0.028
ρ_{ga}	0.52	0.732	0.602	0.260	0.633	0.520	0.245	0.602	0.520	0.192
α	0.19	0.234	0.190	0.032	0.235	0.190	0.037	0.215	0.184	0.028

Table B.3: Inference from ML Estimates

Parameter	DGP	Samples					
		100		**200**		**400**	
		ASE	SD	ASE	SD	ASE	SD
σ_a	0.45	0.039	0.044	0.028	0.029	0.019	0.020
σ_b	0.23	0.034	0.029	0.023	0.022	0.016	0.016
σ_g	0.53	0.042	0.044	0.029	0.030	0.020	0.020
σ_I	0.45	0.068	0.065	0.046	0.046	0.031	0.030
σ_r	0.24	0.021	0.022	0.015	0.015	0.011	0.012
σ_p	0.14	0.025	0.026	0.019	0.019	0.014	0.015
σ_w	0.24	0.033	0.032	0.023	0.023	0.016	0.016
ρ_a	0.95	0.031	0.064	0.016	0.025	0.010	0.012
ρ_b	0.22	0.114	0.101	0.079	0.078	0.055	0.055
ρ_g	0.97	0.024	0.046	0.012	0.019	0.007	0.010
ρ_I	0.71	0.082	0.085	0.057	0.059	0.040	0.041
ρ_r	0.15	0.116	0.095	0.081	0.071	0.057	0.057
ρ_p	0.89	0.068	0.106	0.050	0.060	0.035	0.037
ρ_w	0.96	0.035	0.070	0.020	0.033	0.013	0.015
μ_p	0.69	0.184	0.190	0.129	0.149	0.100	0.109
μ_w	0.84	0.094	0.116	0.057	0.064	0.038	0.038
φ	5.74	2.024	2.288	1.572	1.930	1.134	1.262
σ_c	1.38	0.204	0.188	0.147	0.137	0.101	0.100
h	0.71	0.076	0.073	0.052	0.052	0.036	0.036
ζ_w	0.7	0.097	0.100	0.069	0.073	0.049	0.051
σ_l	1.83	1.362	1.553	0.953	1.268	0.644	0.783
ζ_p	0.66	0.073	0.085	0.051	0.055	0.036	0.036
ι_w	0.58	0.250	0.210	0.176	0.177	0.125	0.123
ι_p	0.24	0.167	0.149	0.121	0.116	0.089	0.087
ψ	0.54	0.188	0.192	0.117	0.122	0.075	0.076
ϕ_p	1.6	0.154	0.174	0.108	0.116	0.074	0.079
r_π	2.04	0.495	0.414	0.377	0.317	0.273	0.273
ρ	0.81	0.058	0.055	0.039	0.035	0.026	0.026
r_y	0.08	0.046	0.055	0.034	0.034	0.025	0.024
$r_{\Delta y}$	0.22	0.060	0.062	0.043	0.044	0.030	0.030
$\bar{\pi}$	0.78	0.215	0.260	0.154	0.175	0.111	0.119
$100(\beta^{-1}-1)$	0.16	0.145	0.115	0.105	0.086	0.074	0.069
$\bar{l}$	0.53	1.570	1.797	1.052	1.213	0.771	0.788
$\bar{\gamma}$	0.43	0.023	0.027	0.009	0.010	0.003	0.003
ρ_{ga}	0.52	0.135	0.154	0.093	0.095	0.065	0.068
α	0.19	0.024	0.024	0.017	0.017	0.012	0.012

Table B.4: Inference from SMM Estimates

Parameter	DGP	Samples 100		200		400	
		ASE	*SD*	*ASE*	*SD*	*ASE*	*SD*
σ_a	0.45	0.115	0.311	0.135	0.269	0.130	0.258
σ_b	0.23	0.115	0.311	0.135	0.269	0.130	0.258
σ_g	0.53	0.076	0.352	0.076	0.291	0.068	0.256
σ_I	0.45	0.094	0.331	0.098	0.326	0.077	0.265
σ_r	0.24	0.093	0.137	0.112	0.121	0.113	0.120
σ_p	0.14	0.057	0.236	0.075	0.188	0.068	0.150
σ_w	0.24	0.055	0.207	0.073	0.159	0.066	0.139
ρ_a	0.95	0.242	0.287	0.360	0.305	0.298	0.287
ρ_b	0.22	0.251	0.180	0.376	0.172	0.272	0.175
ρ_g	0.97	0.137	0.264	0.184	0.265	0.133	0.254
ρ_I	0.71	0.158	0.228	0.179	0.237	0.141	0.235
ρ_r	0.15	0.746	0.259	0.950	0.273	0.895	0.271
ρ_p	0.89	0.361	0.267	0.495	0.277	0.449	0.294
ρ_w	0.96	0.476	0.284	0.655	0.300	0.558	0.291
μ_p	0.69	0.681	0.221	0.902	0.264	0.829	0.244
μ_w	0.84	0.523	0.236	0.694	0.270	0.505	0.270
φ	5.74	2.965	3.339	3.399	2.615	3.215	2.290
σ_c	1.38	0.186	0.532	0.236	0.586	0.237	0.573
h	0.71	0.106	0.240	0.139	0.258	0.135	0.264
ζ_w	0.7	0.133	0.162	0.160	0.188	0.153	0.187
σ_l	1.83	2.096	2.417	2.149	2.325	1.679	1.781
ζ_p	0.66	0.152	0.178	0.198	0.196	0.211	0.198
ι_w	0.58	0.476	0.247	0.647	0.272	0.686	0.261
ι_p	0.24	0.374	0.222	0.444	0.237	0.454	0.238
ψ	0.54	0.238	0.262	0.352	0.277	0.315	0.285
ϕ_p	1.6	0.386	0.527	0.569	0.572	0.733	0.565
r_π	2.04	1.642	0.508	2.402	0.546	2.626	0.549
ρ	0.81	0.201	0.114	0.272	0.121	0.269	0.127
r_y	0.08	0.424	0.138	0.618	0.144	0.662	0.150
$r_{\Delta y}$	0.22	0.162	0.131	0.222	0.134	0.215	0.139
$\bar{\pi}$	0.78	0.054	0.388	0.060	0.350	0.051	0.348
$100(\beta^{-1}-1)$	0.16	0.086	0.455	0.106	0.409	0.095	0.437
$\bar{l}$	0.53	0.257	3.479	0.308	2.704	0.304	1.302
$\bar{\gamma}$	0.43	0.032	0.168	0.034	0.166	0.029	0.160
ρ_{ga}	0.52	0.435	0.463	0.636	0.481	0.575	0.430
α	0.19	0.038	0.174	0.045	0.186	0.038	0.165

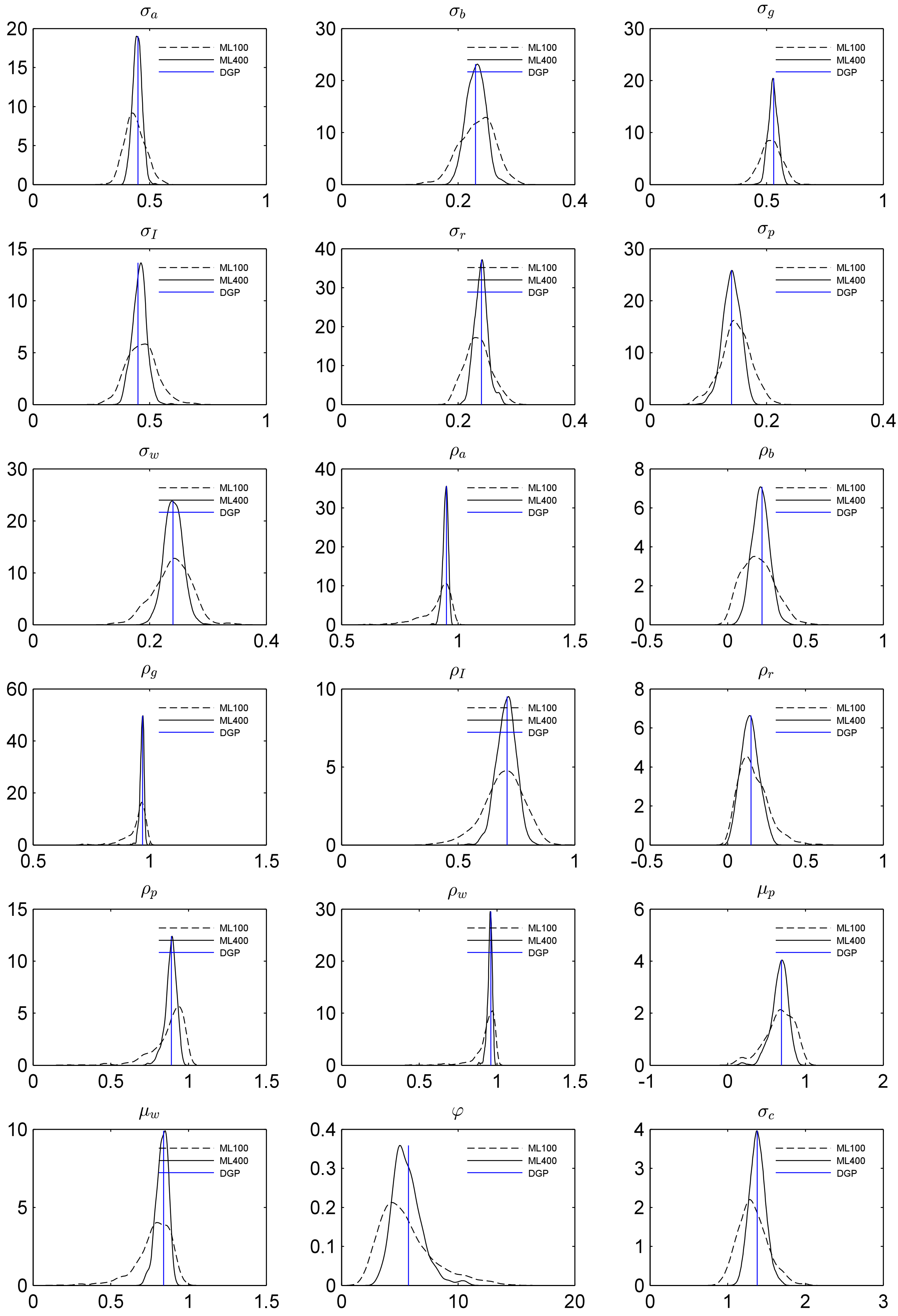

Figure B.1: Kernel density estimates of the empirical distributions of the estimates for ML

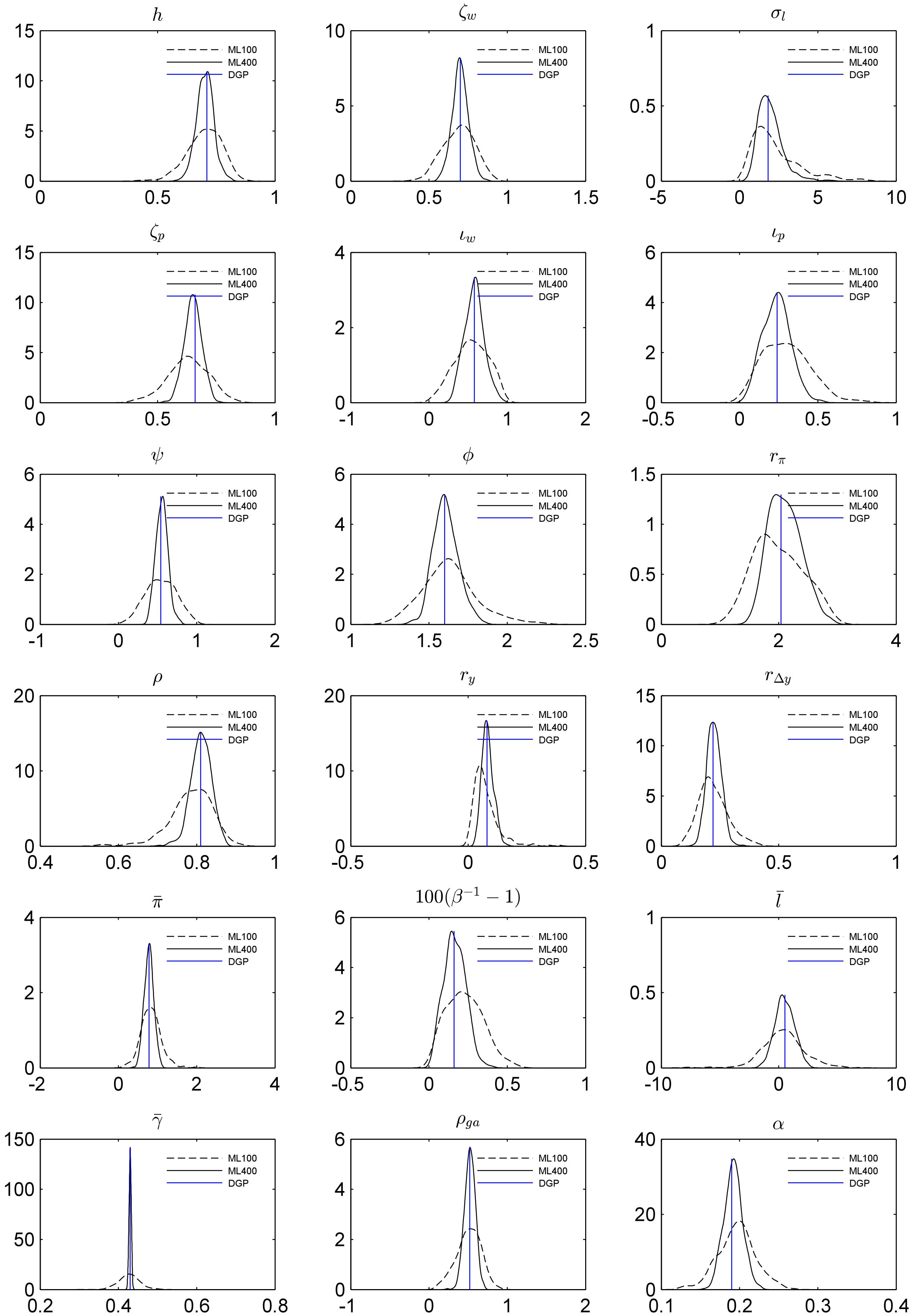

Figure B.1: Kernel density estimates of the empirical distributions of the estimates for ML (continued)

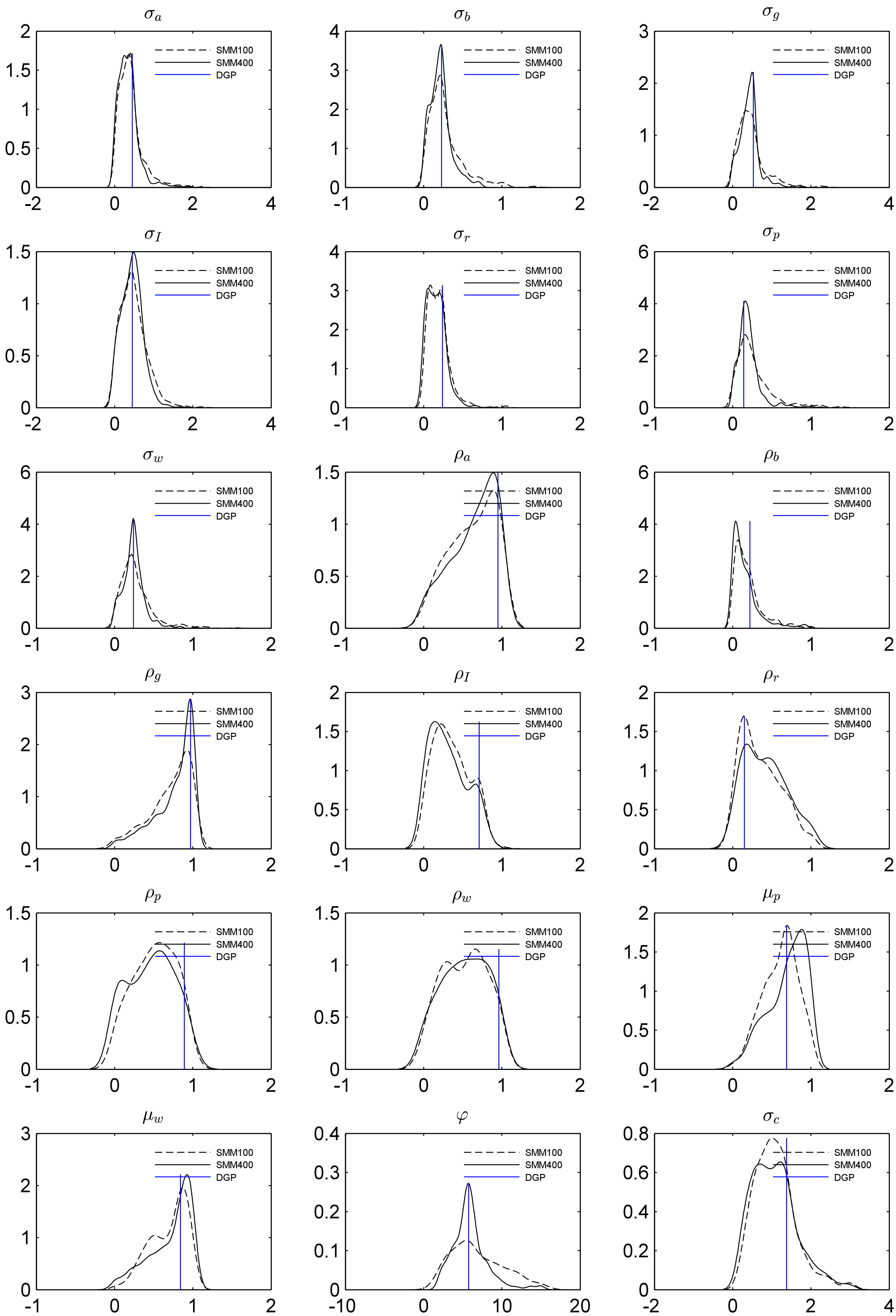

Figure B.2: Kernel density estimates of the empirical distributions of the estimates for SMM

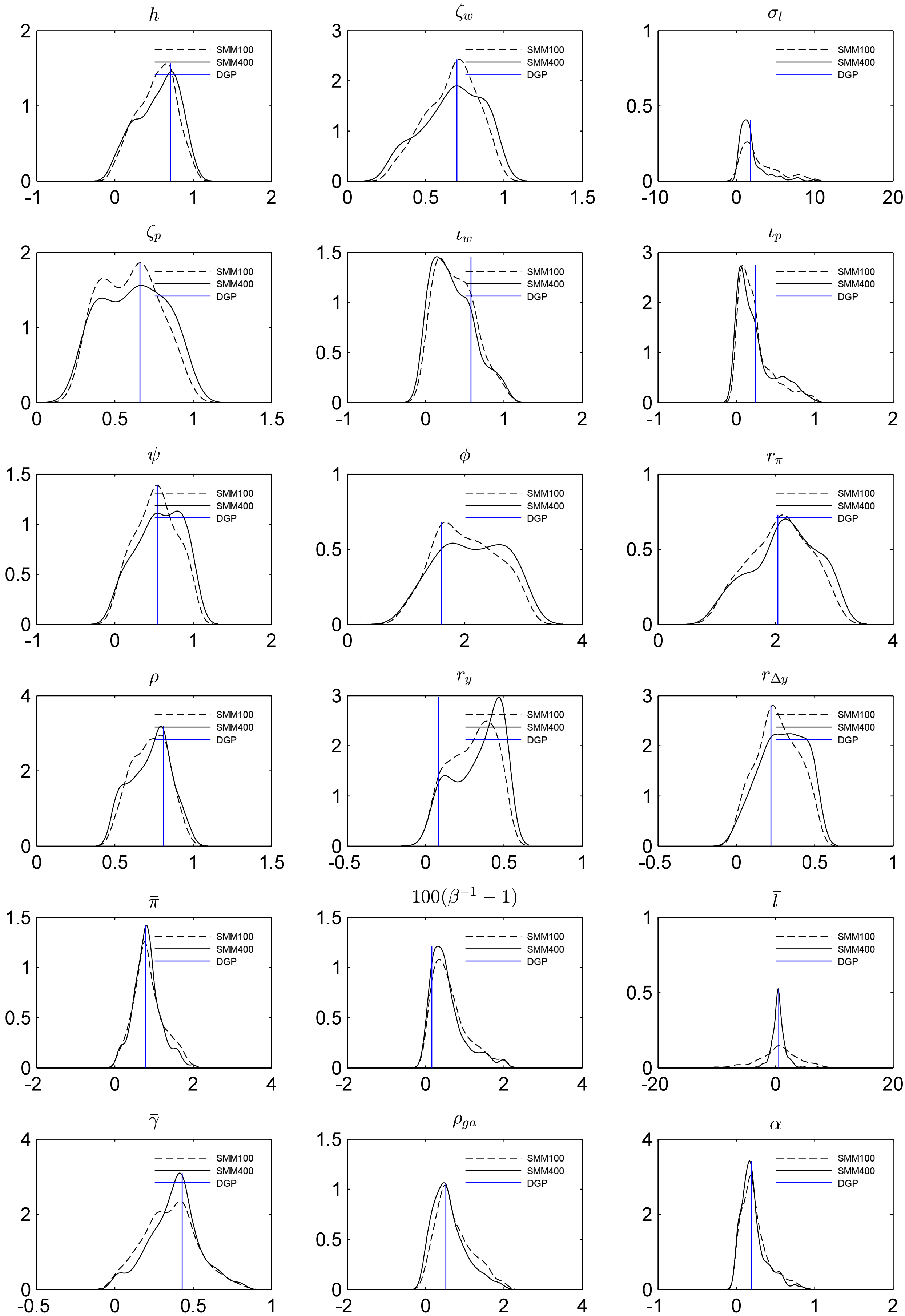

Figure B.2: Kernel density estimates of the empirical distributions of the estimates for SMM (continued)

B.2 Misspecification: Second-order DGP

B.2.1 Starting from the Posteriors (DGP)

Table B.5: Empirical Distribution of Estimates under Misspecification: DGP=Second Order, x0=Posterior Means

Parameter	DGP	*ML*						*SMM*					
		Mean	*Median*	*STD*	*Bias*	*Bias*(%)	*MSE*	*Mean*	*Median*	*STD*	*Bias*	*Bias*(%)	*MSE*
σ_a	0.45	0.50	0.47	0.08	0.05	11.92	0.01	0.43	0.45	0.15	−0.02	−3.44	0.02
σ_b	0.23	0.50	0.48	0.13	0.27	118.37	0.09	0.31	0.23	0.35	0.08	34.44	0.13
σ_g	0.53	0.92	0.73	0.45	0.39	73.00	0.36	0.56	0.53	0.25	0.03	6.26	0.06
σ_I	0.45	0.96	0.82	0.38	0.51	112.33	0.40	0.57	0.45	0.42	0.12	26.40	0.19
σ_r	0.24	0.33	0.30	0.08	0.09	37.56	0.01	0.23	0.24	0.10	−0.01	−3.55	0.01
σ_p	0.14	0.12	0.13	0.03	−0.02	−14.51	0.00	0.16	0.14	0.12	0.02	12.62	0.01
σ_w	0.24	0.20	0.22	0.05	−0.04	−15.56	0.00	0.23	0.24	0.10	−0.01	−2.54	0.01
ρ_a	0.95	0.89	0.91	0.08	−0.06	−5.93	0.01	0.88	0.95	0.20	−0.07	−7.40	0.05
ρ_b	0.22	0.44	0.40	0.13	0.22	101.55	0.07	0.22	0.22	0.09	0.00	0.30	0.01
ρ_g	0.97	0.92	0.95	0.09	−0.05	−5.24	0.01	0.90	0.97	0.19	−0.07	−7.08	0.04
ρ_I	0.71	0.94	0.94	0.03	0.23	31.71	0.05	0.67	0.71	0.14	−0.04	−5.49	0.02
ρ_r	0.15	0.18	0.15	0.08	0.03	21.16	0.01	0.20	0.15	0.15	0.05	30.55	0.02
ρ_p	0.89	0.93	0.93	0.03	0.04	4.81	0.00	0.79	0.89	0.23	−0.10	−11.48	0.06
ρ_w	0.96	0.93	0.93	0.04	−0.03	−3.17	0.00	0.88	0.96	0.20	−0.08	−8.22	0.05
μ_p	0.69	0.66	0.67	0.04	−0.03	−4.50	0.00	0.66	0.69	0.13	−0.03	−3.95	0.02
μ_w	0.84	0.84	0.85	0.06	0.00	−0.36	0.00	0.81	0.84	0.13	−0.03	−3.55	0.02
φ	5.74	5.71	5.71	0.02	−0.03	−0.60	0.00	6.38	5.74	1.91	0.64	11.17	4.04
σ_c	1.38	1.28	1.31	0.09	−0.10	−7.19	0.02	1.39	1.38	0.33	0.01	0.73	0.11
h	0.71	0.61	0.61	0.09	−0.10	−14.05	0.02	0.69	0.71	0.10	−0.02	−2.46	0.01
ζ_w	0.7	0.72	0.70	0.03	0.02	2.26	0.00	0.70	0.70	0.07	0.00	0.29	0.00
σ_l	1.83	1.82	1.83	0.02	−0.01	−0.41	0.00	2.02	1.83	1.24	0.19	10.51	1.58
ζ_p	0.66	0.70	0.69	0.03	0.04	5.41	0.00	0.66	0.66	0.07	0.00	0.33	0.01
ι_w	0.58	0.59	0.58	0.04	0.01	1.02	0.00	0.53	0.58	0.14	−0.05	−8.16	0.02
ι_p	0.24	0.25	0.25	0.03	0.01	2.15	0.00	0.25	0.24	0.12	0.01	4.65	0.01
ψ	0.54	0.38	0.46	0.16	−0.16	−28.80	0.05	0.51	0.54	0.13	−0.03	−5.86	0.02
ϕ_p	1.6	1.68	1.62	0.16	0.08	4.70	0.03	1.74	1.60	0.35	0.14	8.83	0.14
r_π	2.04	2.05	2.04	0.04	0.01	0.65	0.00	2.11	2.04	0.24	0.07	3.53	0.06
ρ	0.81	0.72	0.73	0.06	−0.09	−11.20	0.01	0.78	0.81	0.07	−0.03	−3.33	0.01
r_y	0.08	0.14	0.11	0.08	0.06	73.48	0.01	0.13	0.08	0.11	0.05	59.19	0.01
$r_{\Delta y}$	0.22	0.21	0.22	0.05	−0.01	−2.89	0.00	0.22	0.22	0.06	0.00	−0.66	0.00
$\bar{\pi}$	0.78	0.78	0.78	0.00	0.00	0.02	0.00	0.80	0.78	0.23	0.02	3.05	0.05
$100(\beta^{-1}-1)$	0.16	0.16	0.16	0.01	0.00	−2.18	0.00	0.31	0.16	0.40	0.15	93.82	0.18
$\bar{l}$	−0.06	−0.06	−0.07	0.01	0.00	7.82	0.00	0.03	−0.06	1.54	0.09	−152.21	2.38
$\bar{\gamma}$	0.43	0.42	0.43	0.05	−0.01	−1.28	0.00	0.43	0.43	0.03	0.00	−0.24	0.00
ρ_{ga}	0.52	0.55	0.53	0.07	0.03	6.30	0.01	0.60	0.52	0.29	0.08	16.04	0.09
α	0.19	0.02	0.01	0.04	−0.17	−88.38	0.03	0.21	0.19	0.12	0.02	9.26	0.01

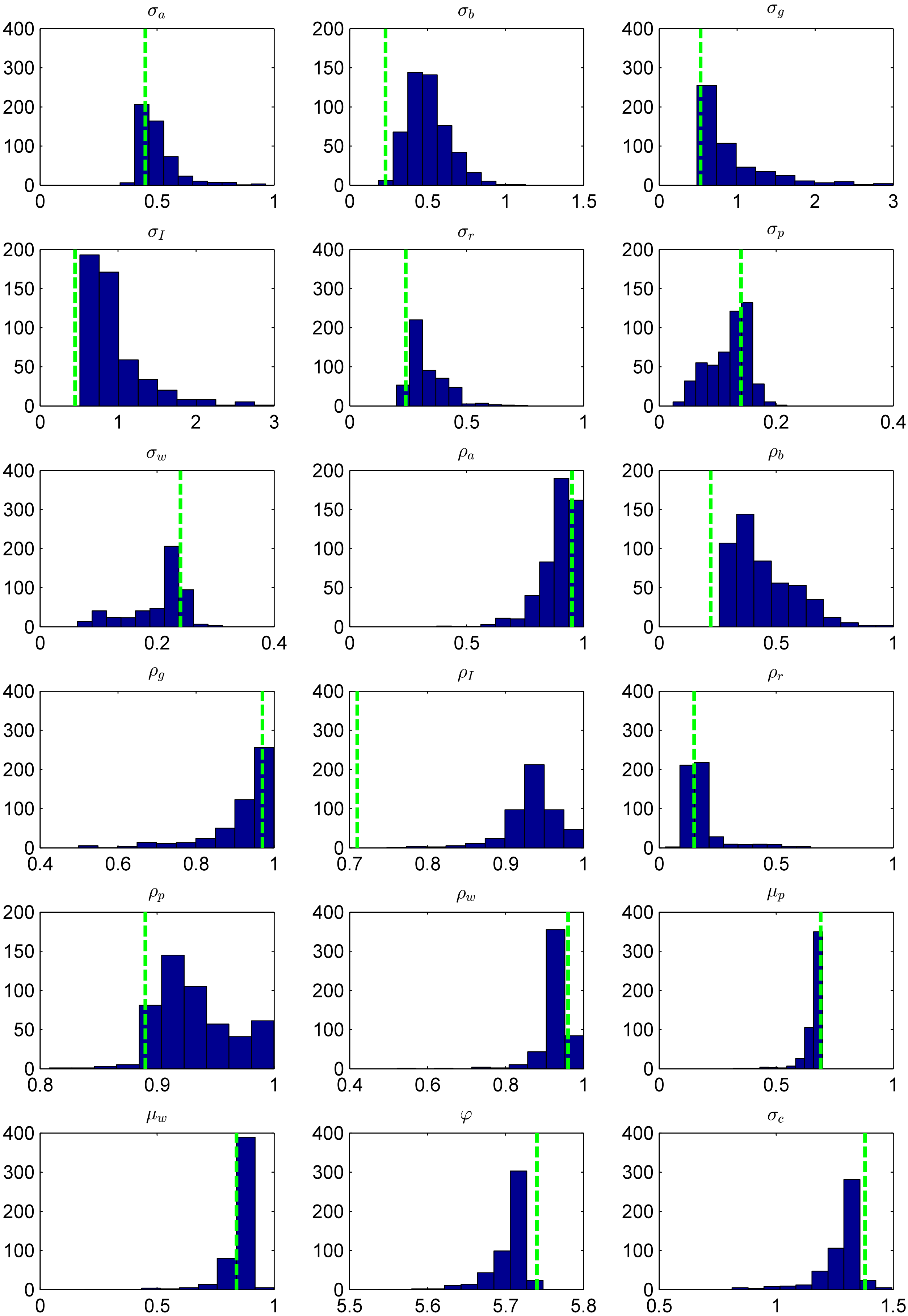

Figure B.3: Histogram of the empirical distributions of the ML estimates: x0=DGP

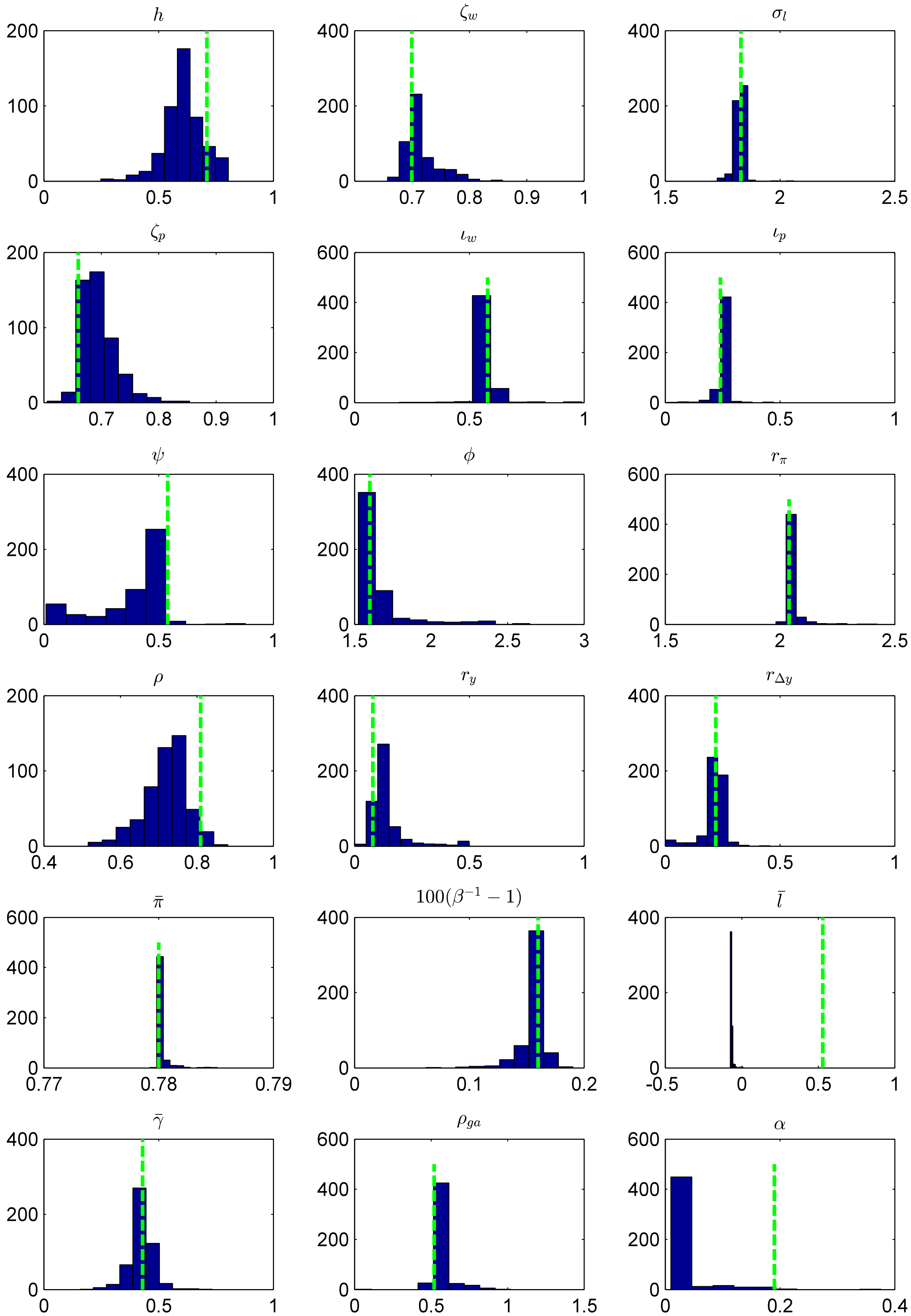

Figure B.3: Histogram of the empirical distributions of the ML estimates: x0=DGP (continued)

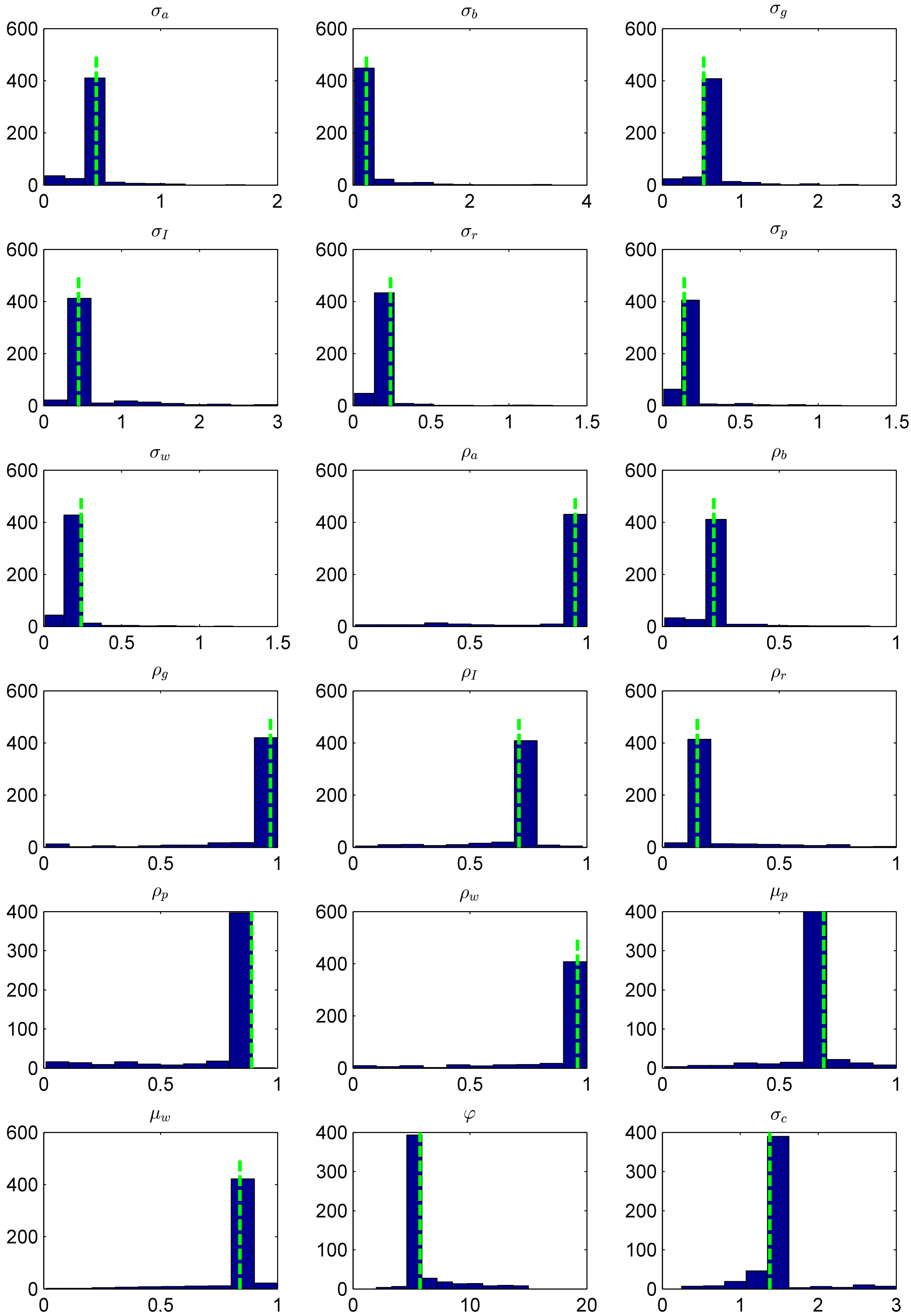

Figure B.4: Histogram of the empirical distributions of the SMM estimates: x0=DGP

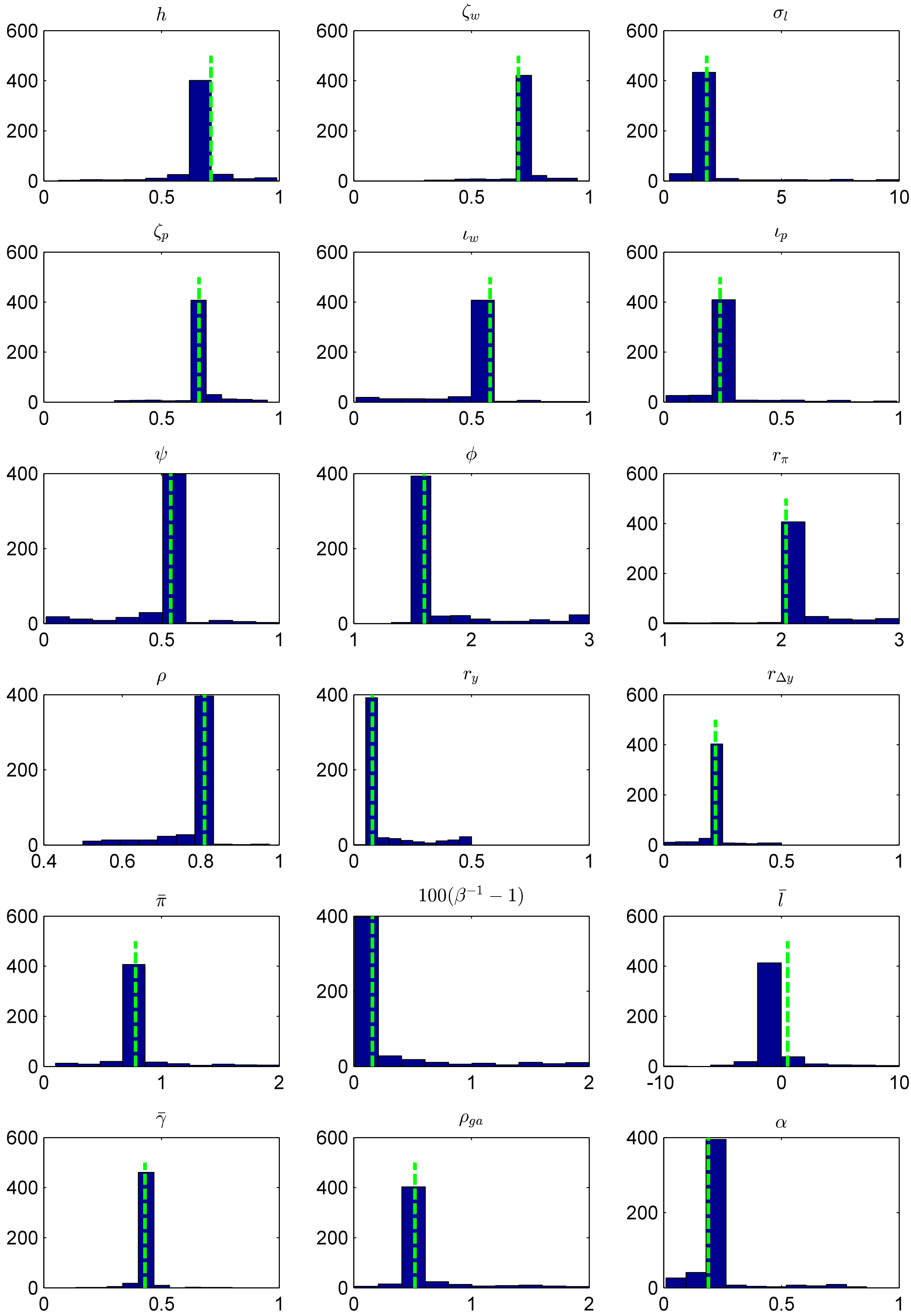

Figure B.4: Histogram of the empirical distributions of the SMM estimates: x0=DGP (continued)

B.2.2 Starting from the Priors

Table B.6: Empirical Distribution of Estimates under Misspecification: DGP=Second Order, x0=Prior Means

Parameter	DGP	x0	*ML*						*SMM*					
			Mean	*Median*	*STD*	*Bias*	*Bias*(%)	*MSE*	*Mean*	*Median*	*STD*	*Bias*	*Bias*(%)	*MSE*
σ_a	0.45	0.1	0.64	0.59	0.21	0.19	42.00	0.08	0.24	0.18	0.21	−0.21	−45.91	0.09
σ_b	0.23	0.1	0.47	0.46	0.12	0.24	105.11	0.07	0.32	0.21	0.33	0.09	39.98	0.12
σ_g	0.53	0.1	1.25	1.16	0.49	0.72	136.79	0.76	0.37	0.22	0.38	−0.16	−29.63	0.17
σ_I	0.45	0.1	1.49	1.40	0.59	1.04	231.42	1.43	0.85	0.68	0.77	0.40	88.81	0.75
σ_r	0.24	0.1	0.35	0.34	0.07	0.11	47.20	0.02	0.16	0.13	0.11	−0.08	−34.42	0.02
σ_p	0.14	0.1	0.14	0.12	0.10	0.00	3.36	0.01	0.11	0.10	0.07	−0.03	−19.49	0.01
σ_w	0.24	0.1	0.19	0.14	0.13	−0.05	−22.53	0.02	0.12	0.10	0.07	−0.12	−51.26	0.02
ρ_a	0.95	0.5	0.82	0.84	0.13	−0.13	−13.89	0.03	0.44	0.43	0.21	−0.51	−53.92	0.30
ρ_b	0.22	0.5	0.71	0.71	0.08	0.49	223.63	0.25	0.34	0.32	0.18	0.12	54.16	0.05
ρ_g	0.97	0.5	0.93	0.96	0.07	−0.04	−4.39	0.01	0.54	0.52	0.22	−0.43	−44.66	0.24
ρ_I	0.71	0.5	0.90	0.92	0.08	0.19	26.18	0.04	0.47	0.48	0.20	−0.24	−33.52	0.10
ρ_r	0.15	0.5	0.44	0.43	0.05	0.29	191.00	0.08	0.46	0.47	0.21	0.31	208.72	0.14
ρ_p	0.89	0.5	0.69	0.65	0.11	−0.20	−22.79	0.05	0.38	0.37	0.19	−0.51	−57.84	0.30
ρ_w	0.96	0.5	0.61	0.59	0.06	−0.35	−36.60	0.13	0.49	0.50	0.22	−0.47	−48.72	0.27
μ_p	0.69	0.5	0.39	0.39	0.04	−0.30	−43.56	0.09	0.54	0.53	0.20	−0.15	−21.84	0.06
μ_w	0.84	0.5	0.40	0.41	0.06	−0.44	−52.14	0.19	0.51	0.51	0.19	−0.33	−39.43	0.15
φ	5.74	4	3.94	3.95	0.04	−1.80	−31.34	3.24	5.27	4.57	2.17	−0.47	−8.13	4.91
σ_c	1.38	1.5	1.29	1.32	0.11	−0.09	−6.60	0.02	1.36	1.31	0.50	−0.02	−1.45	0.25
h	0.71	0.7	0.64	0.65	0.09	−0.07	−10.27	0.01	0.64	0.67	0.17	−0.07	−9.75	0.03
ζ_w	0.7	0.5	0.73	0.73	0.06	0.03	3.97	0.00	0.64	0.64	0.13	−0.06	−8.20	0.02
σ_l	1.83	2	1.95	1.96	0.05	0.12	6.56	0.02	1.96	1.71	1.20	0.13	7.07	1.45
ζ_p	0.66	0.5	0.70	0.70	0.07	0.04	6.26	0.01	0.57	0.56	0.12	−0.09	−13.66	0.02
ι_w	0.58	0.5	0.53	0.51	0.06	−0.05	−8.14	0.01	0.49	0.50	0.22	−0.09	−15.33	0.06
ι_p	0.24	0.5	0.49	0.53	0.10	0.25	104.78	0.07	0.43	0.43	0.22	0.19	80.57	0.08
ψ	0.54	0.5	0.34	0.38	0.13	−0.20	−37.10	0.06	0.43	0.43	0.21	−0.11	−19.89	0.06
ϕ_p	1.6	1.25	1.62	1.49	0.31	0.02	1.55	0.10	2.00	1.96	0.48	0.40	25.10	0.39
r_π	2.04	1.5	1.58	1.54	0.13	−0.46	−22.56	0.23	2.14	2.14	0.44	0.10	4.70	0.20
ρ	0.81	0.75	0.65	0.66	0.06	−0.16	−19.38	0.03	0.71	0.71	0.08	−0.10	−12.54	0.02
r_y	0.08	0.125	0.22	0.17	0.12	0.14	173.52	0.03	0.26	0.26	0.11	0.18	222.29	0.04
$r_{\Delta y}$	0.22	0.125	0.05	0.04	0.06	−0.17	−77.75	0.03	0.17	0.14	0.11	−0.05	−22.59	0.01
$\bar{\pi}$	0.78	0.625	0.63	0.63	0.00	−0.15	−19.83	0.02	0.80	0.62	0.57	0.02	2.34	0.32
$100(\beta^{-1}-1)$	0.16	0.25	0.27	0.27	0.02	0.11	71.03	0.01	0.40	0.25	0.38	0.24	150.09	0.20
$\bar{l}$	−0.06	0	0.01	0.01	0.02	0.07	−109.64	0.00	0.32	0.24	1.17	0.38	−638.41	1.52
$\bar{\gamma}$	0.43	0.4	0.43	0.43	0.03	0.00	−0.53	0.00	0.37	0.37	0.12	−0.06	−14.95	0.02
ρ_{ga}	0.52	0.5	0.68	0.62	0.20	0.16	31.73	0.07	1.08	1.08	0.44	0.56	108.21	0.51
α	0.19	0.3	0.06	0.01	0.07	−0.13	−70.13	0.02	0.30	0.27	0.18	0.11	58.44	0.04

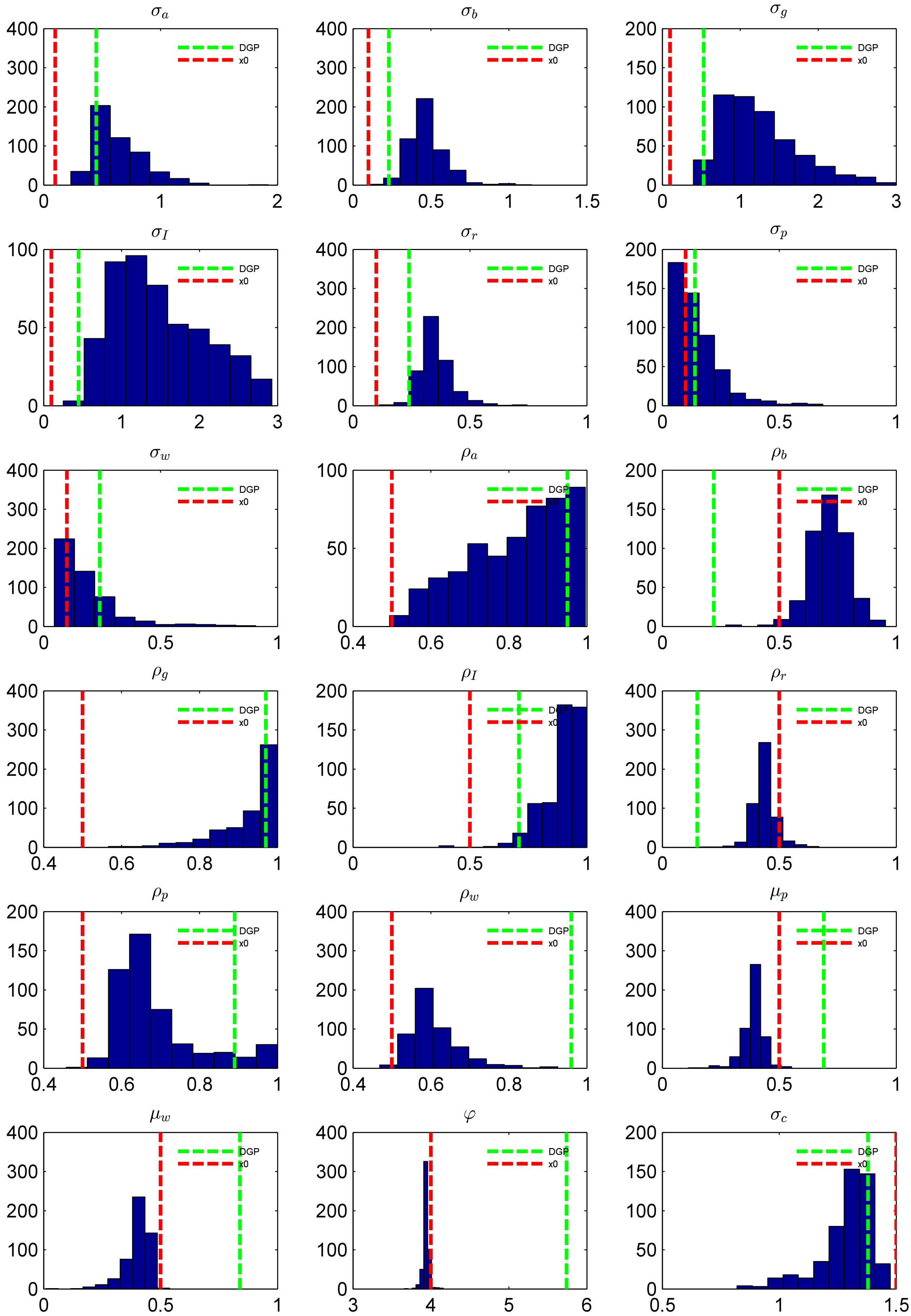

Figure B.5: Histogram of the empirical distributions of the ML estimates: x0=Priors

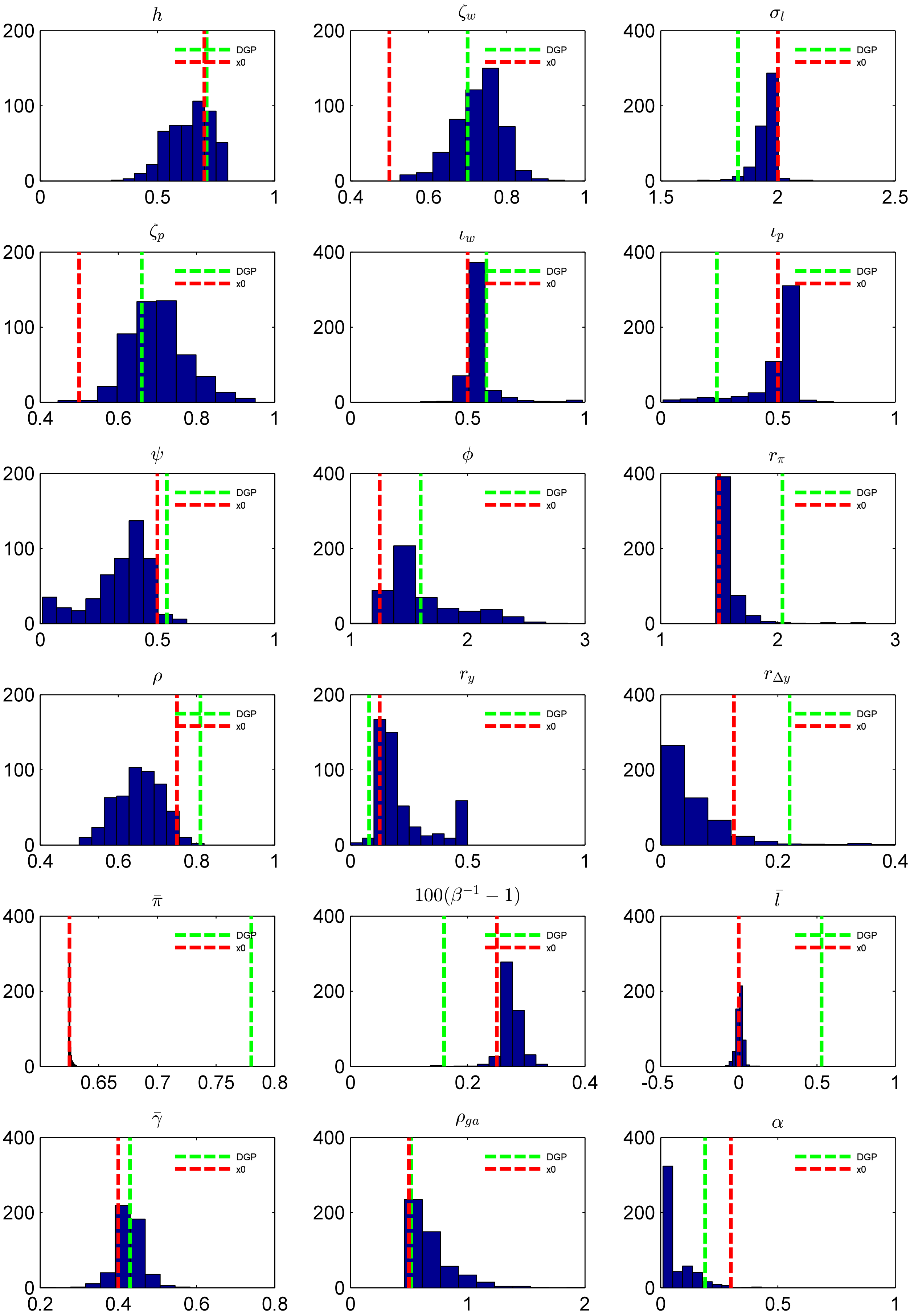

Figure B.5: Histogram of the empirical distributions of the ML estimates: x0=Priors (continued)

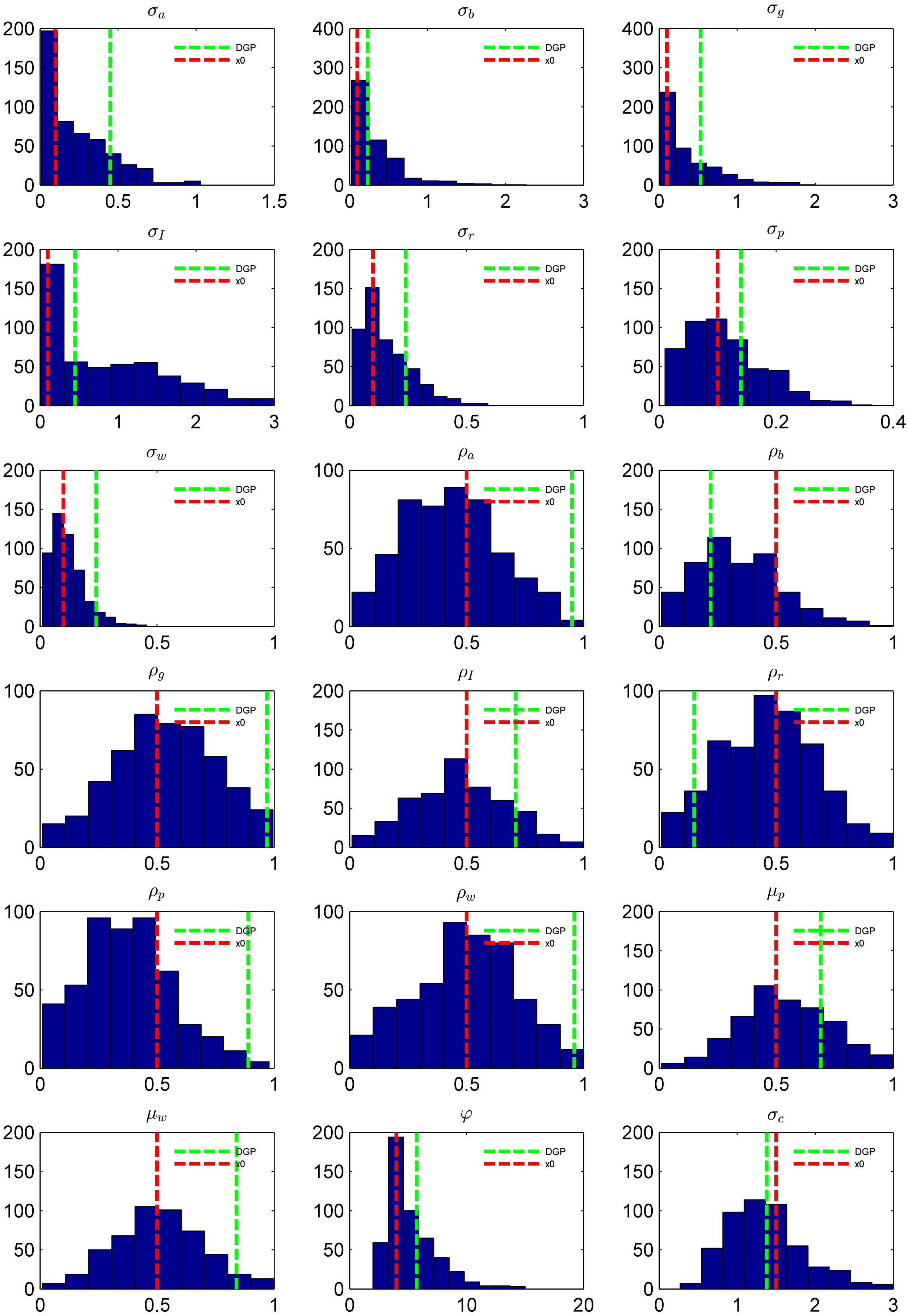

Figure B.6: Histogram of the empirical distributions of the SMM estimates: x0=Priors

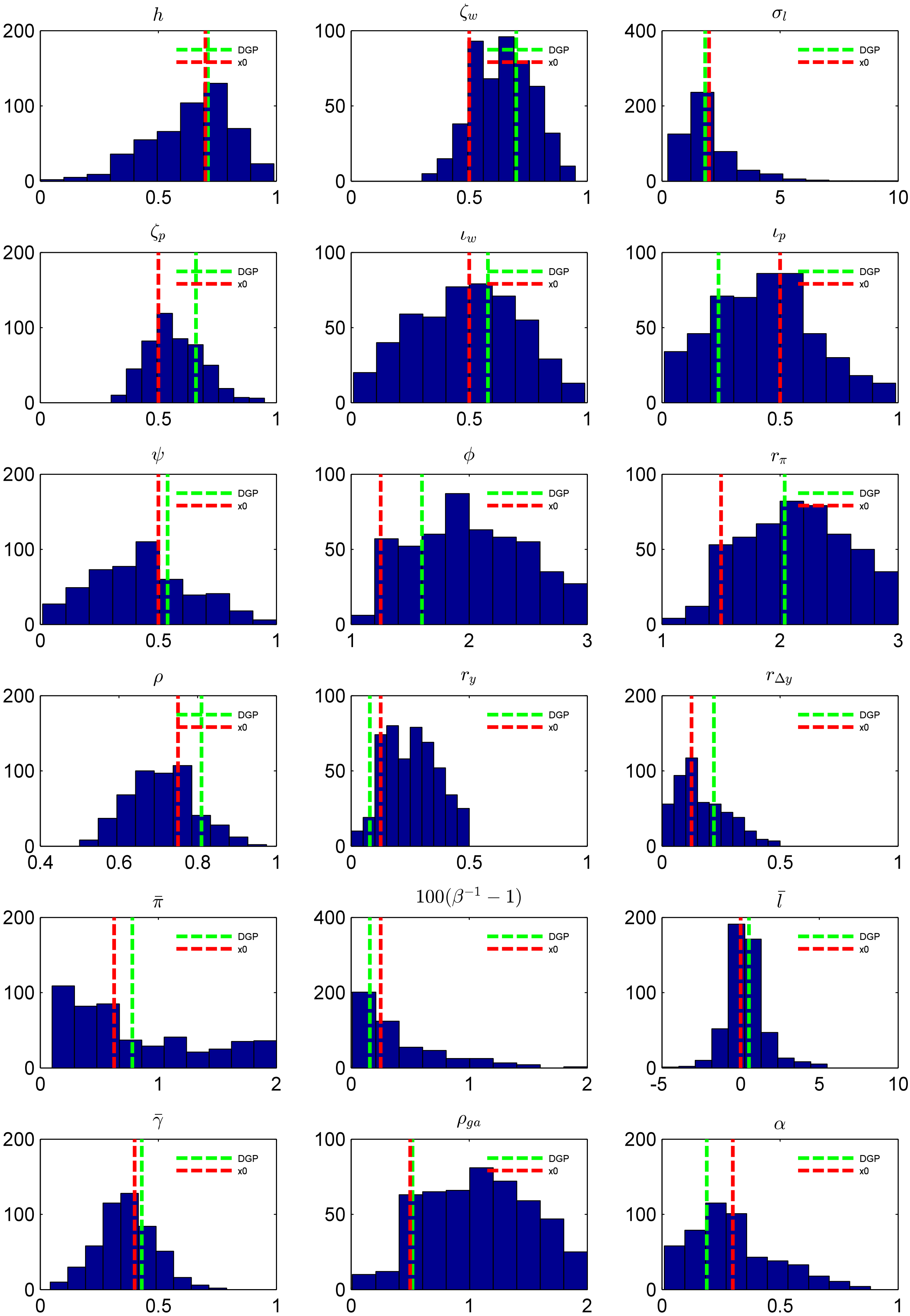

Figure B.6: Histogram of the empirical distributions of the SMM estimates: x0=Priors (continued)

B.3 Misspecification: Estimated Model has i.i.d. Shocks

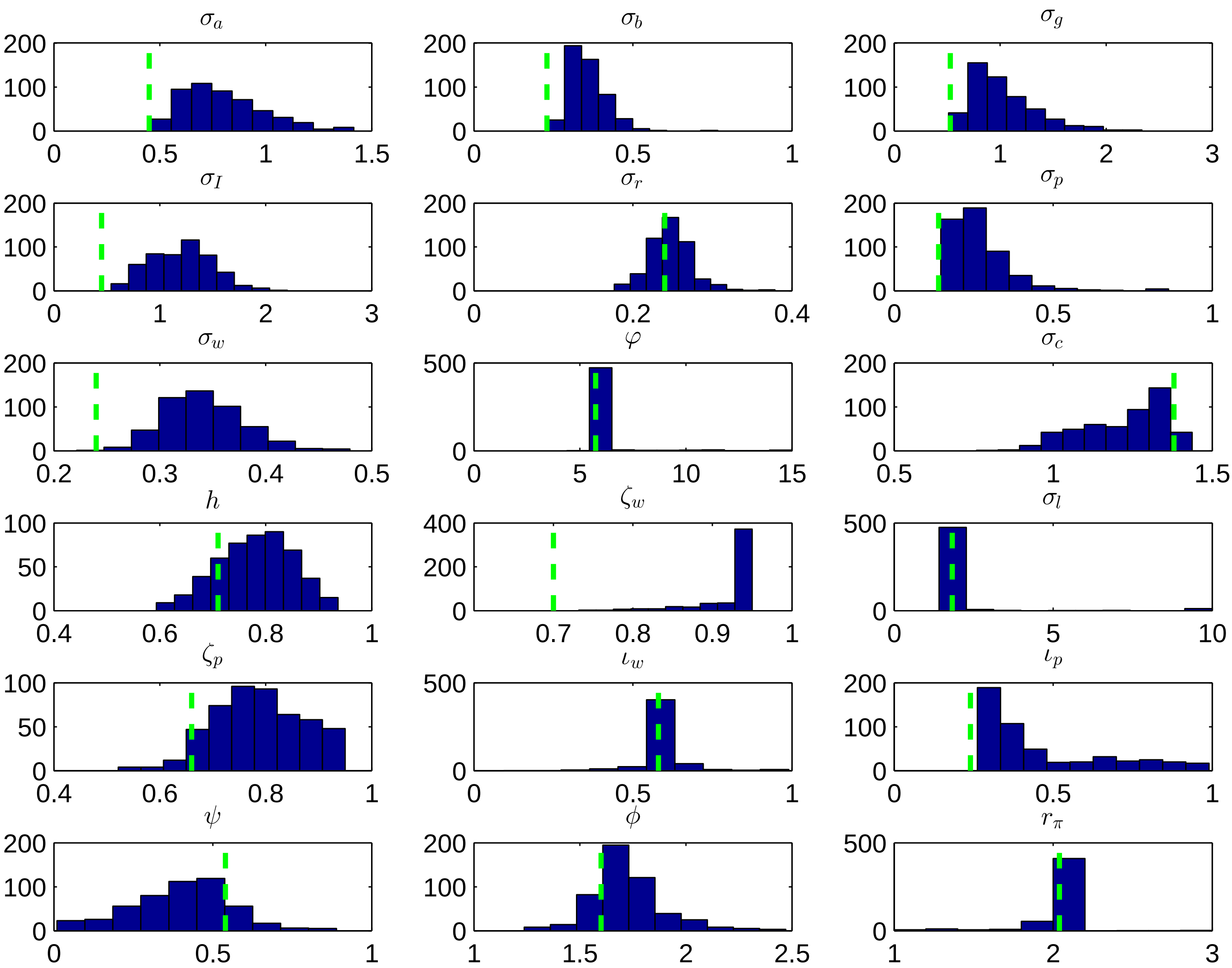

Figure B.7: Histogram of the empirical distributions of the ML estimates: x0=DGP

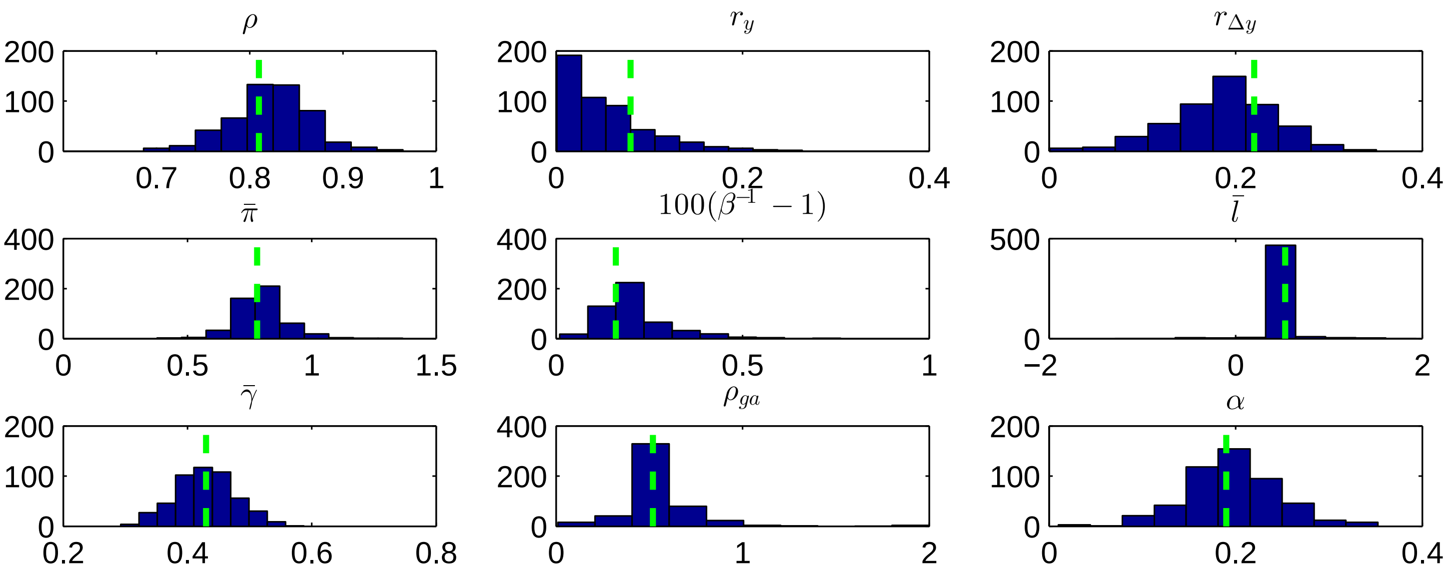

Figure B.7: Histogram of the empirical distributions of the ML estimates: x0=DGP (continued)

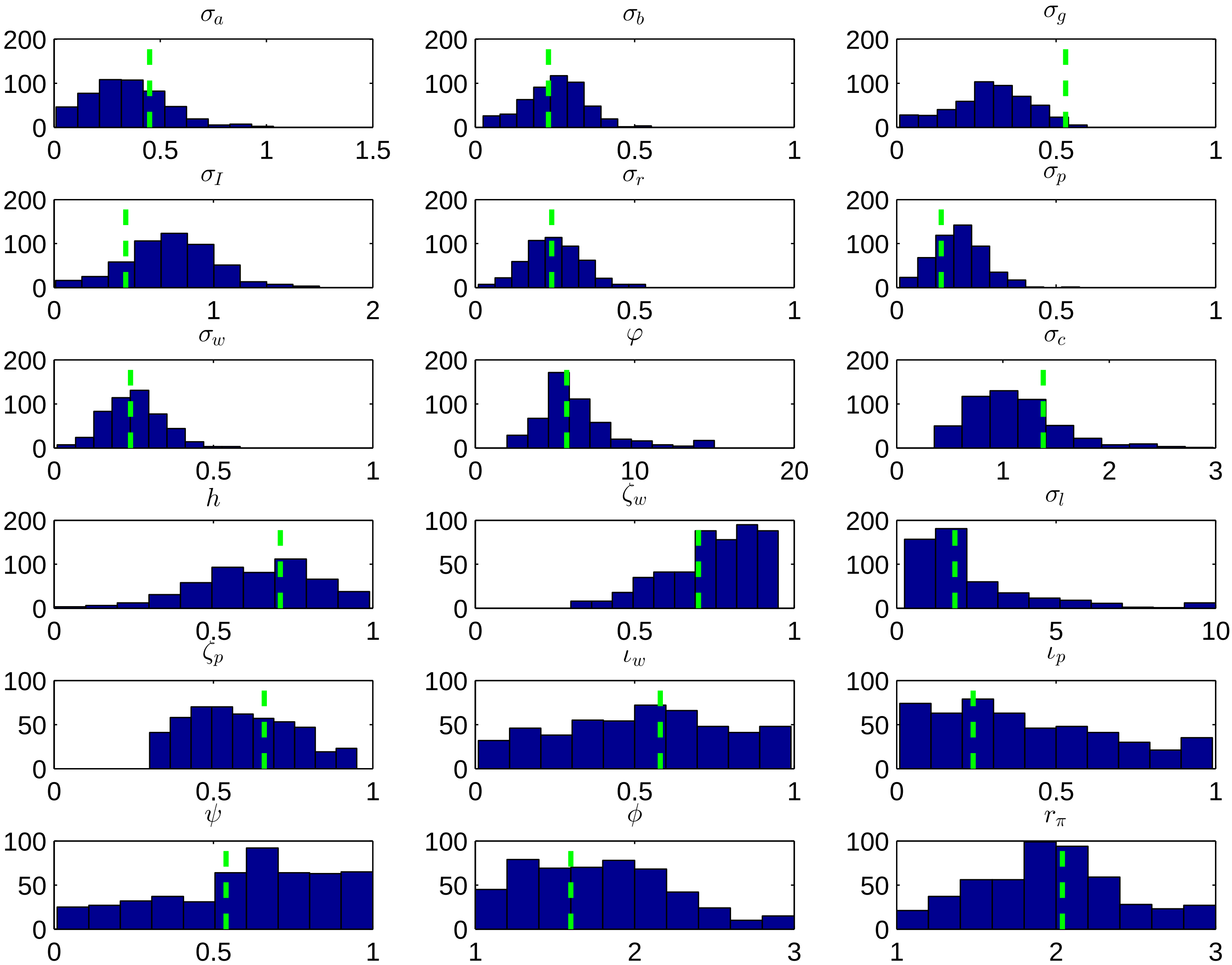

Figure B.8: Histogram of the empirical distributions of the SMM estimates: x0=DGP

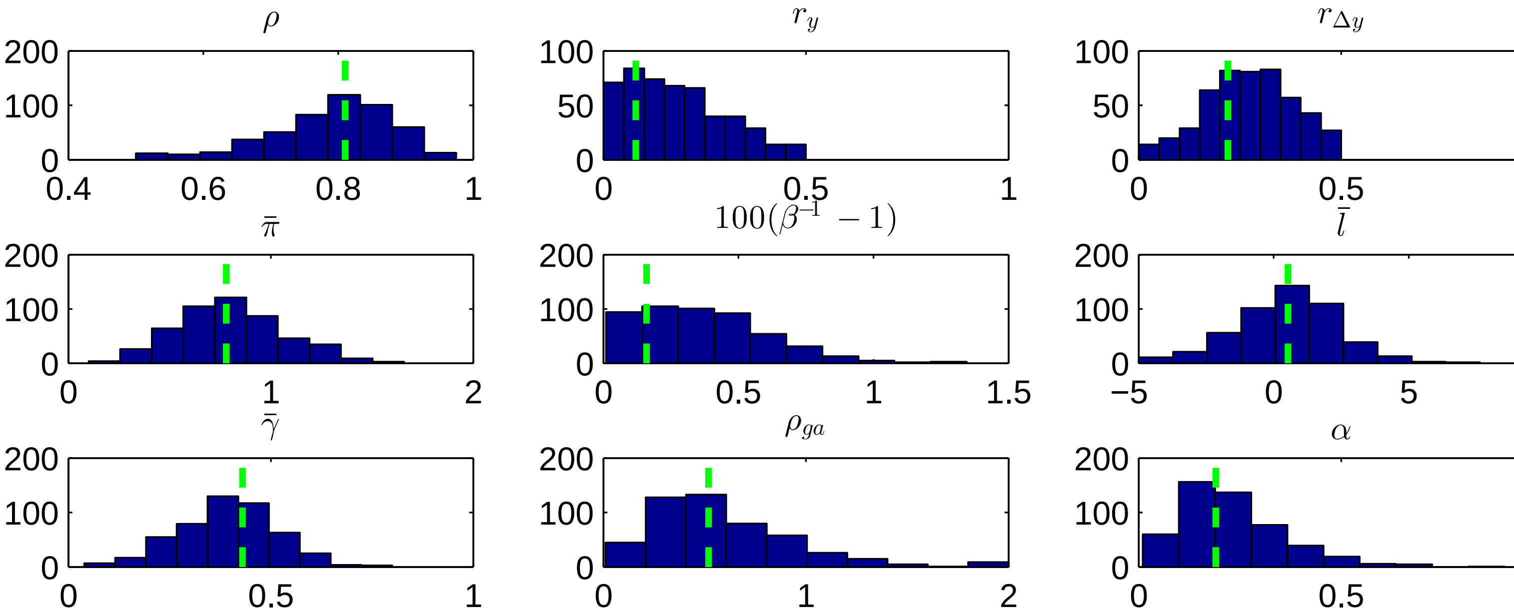

Figure B.8: Histogram of the empirical distributions of the SMM estimates: x0=DGP (continued)

Table B.7: Empirical Distribution of Estimates under Misspecification: DGP=Linear, Model= i.i.d., x0=Posterior Means

Parameter	DGP	*ML*						*SMM*					
		Mean	*Median*	*STD*	*Bias*	*Bias*(%)	*MSE*	*Mean*	*Median*	*STD*	*Bias*	*Bias*(%)	*MSE*
σ_a	0.45	0.80	0.77	0.18	0.35	77.36	0.16	0.35	0.34	0.18	−0.10	−23.30	0.04
σ_b	0.23	0.36	0.35	0.06	0.13	55.40	0.02	0.25	0.25	0.09	0.02	7.46	0.01
σ_g	0.53	1.03	0.96	0.30	0.50	93.70	0.34	0.29	0.30	0.12	−0.24	−44.71	0.07
σ_I	0.45	1.20	1.22	0.29	0.75	166.11	0.64	0.73	0.75	0.28	0.28	62.11	0.15
σ_r	0.24	0.25	0.25	0.03	0.01	2.72	0.00	0.25	0.24	0.09	0.01	3.74	0.01
σ_p	0.14	0.27	0.25	0.10	0.13	90.69	0.03	0.19	0.19	0.08	0.05	39.26	0.01
σ_w	0.24	0.34	0.34	0.04	0.10	42.25	0.01	0.25	0.25	0.09	0.01	5.25	0.01
φ	5.74	6.00	5.75	1.18	0.26	4.58	1.47	6.33	5.80	2.52	0.59	10.30	6.72
σ_c	1.38	1.22	1.26	0.13	−0.16	−11.47	0.04	1.10	1.04	0.42	−0.28	−20.51	0.25
h	0.71	0.78	0.78	0.07	0.07	9.88	0.01	0.64	0.65	0.18	−0.07	−10.33	0.04
ζ_w	0.7	0.93	0.95	0.04	0.23	32.89	0.05	0.74	0.77	0.15	0.04	6.13	0.02
σ_l	1.83	2.08	1.82	1.30	0.25	13.46	1.74	2.25	1.63	1.95	0.42	23.20	3.97
ζ_p	0.66	0.79	0.78	0.09	0.13	19.40	0.02	0.58	0.57	0.16	−0.08	−11.71	0.03
ι_w	0.58	0.60	0.60	0.08	0.02	2.82	0.01	0.52	0.54	0.27	−0.06	−9.71	0.07
ι_p	0.24	0.46	0.37	0.20	0.22	92.76	0.09	0.41	0.36	0.27	0.17	69.39	0.10
ψ	0.54	0.40	0.42	0.16	−0.14	−26.01	0.05	0.60	0.63	0.26	0.06	11.64	0.07
ϕ_p	1.6	1.72	1.70	0.17	0.12	7.56	0.04	1.79	1.77	0.46	0.19	11.92	0.25
r_π	2.04	1.99	2.03	0.18	−0.05	−2.45	0.04	1.97	1.96	0.46	−0.07	−3.65	0.22
ρ	0.81	0.82	0.82	0.04	0.01	1.35	0.00	0.79	0.80	0.09	−0.02	−2.69	0.01
r_y	0.08	0.05	0.04	0.05	−0.03	−35.86	0.00	0.18	0.17	0.12	0.10	128.97	0.03
$r_{\Delta y}$	0.22	0.19	0.19	0.06	−0.03	−15.81	0.00	0.27	0.28	0.11	0.05	24.71	0.02
$\bar{\pi}$	0.78	0.80	0.79	0.11	0.02	2.07	0.01	0.80	0.78	0.27	0.02	2.79	0.07
$100(\beta^{-1}-1)$	0.16	0.21	0.18	0.10	0.05	30.63	0.01	0.37	0.34	0.24	0.21	131.76	0.10
$\bar{l}$	−0.06	0.52	0.53	0.19	−0.01	−1.04	0.04	0.53	0.57	1.92	0.00	−0.01	3.68
$\bar{\gamma}$	0.43	0.43	0.43	0.05	0.00	−0.32	0.00	0.39	0.40	0.12	−0.04	−8.22	0.02
ρ_{ga}	0.52	0.56	0.54	0.21	0.04	7.01	0.04	0.59	0.52	0.36	0.07	12.85	0.13
α	0.19	0.20	0.20	0.05	0.01	3.41	0.00	0.23	0.21	0.13	0.04	21.92	0.02

C Does it make a difference?

C.1 Misspecification: Second-order DGP

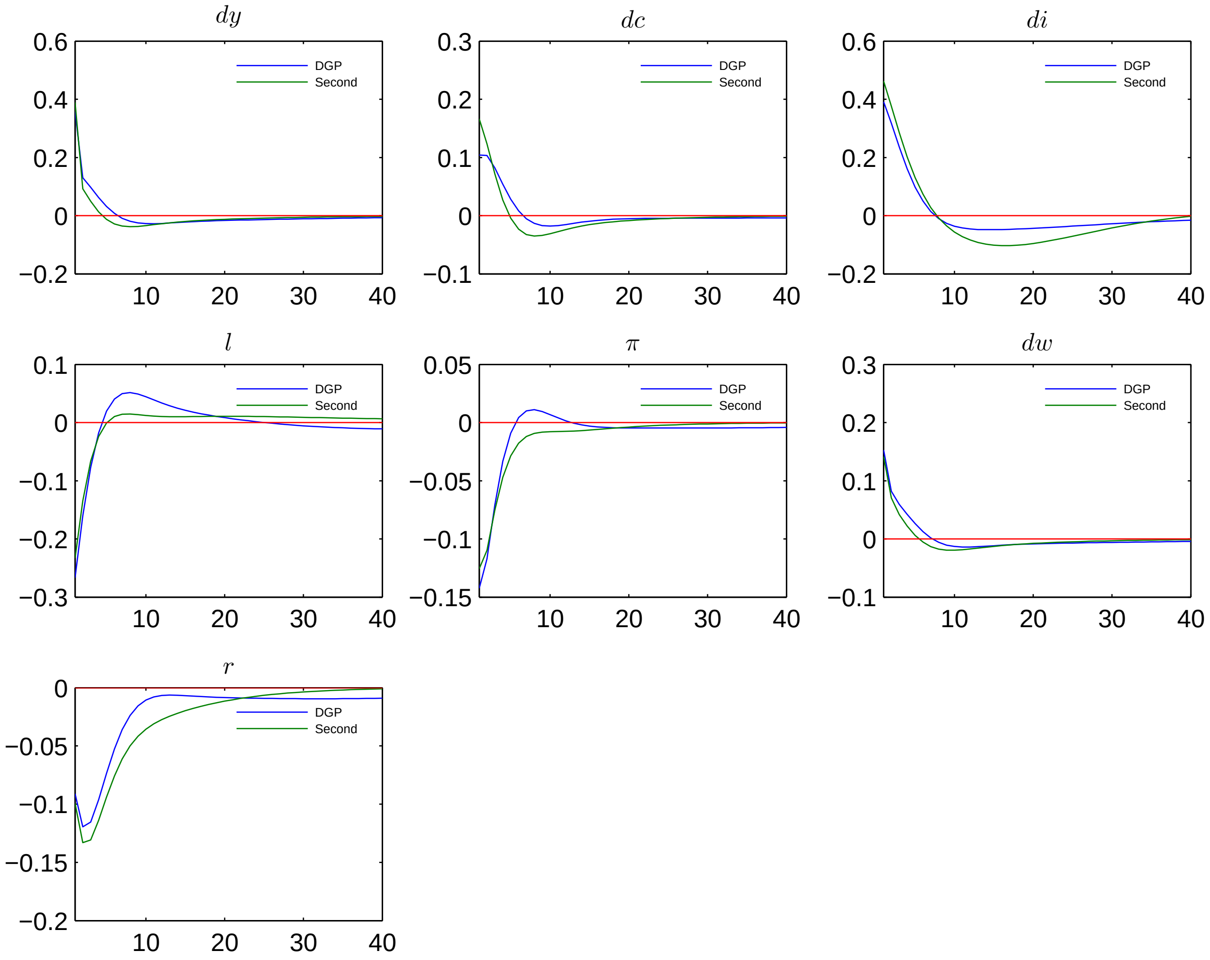

Figure C.1: Responses of the linearized model to a ε_a shock: DGP versus parameter estimated from second order data

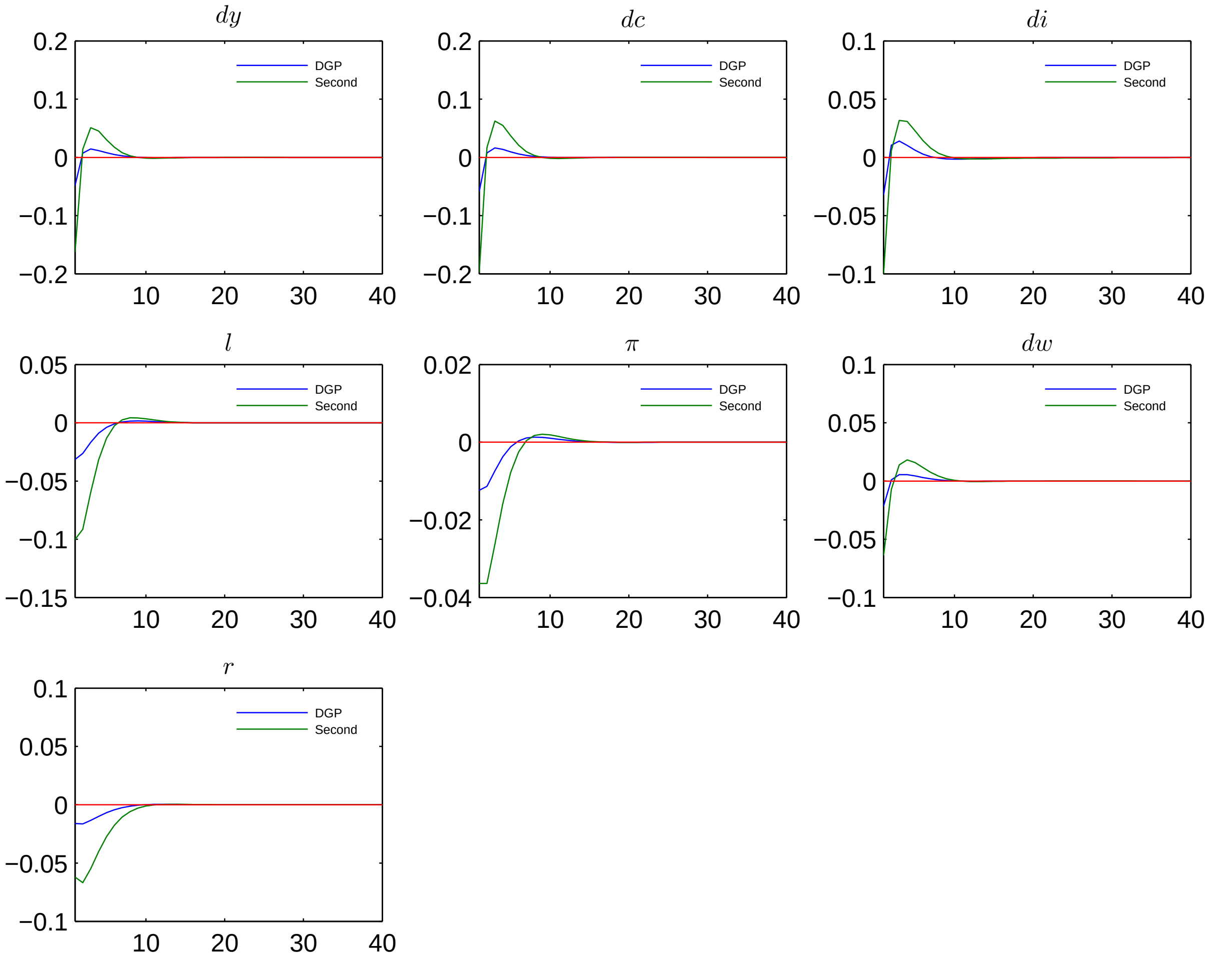

Figure C.2: Responses of the linearized model to a ε_b shock: DGP versus parameter estimated from second order data

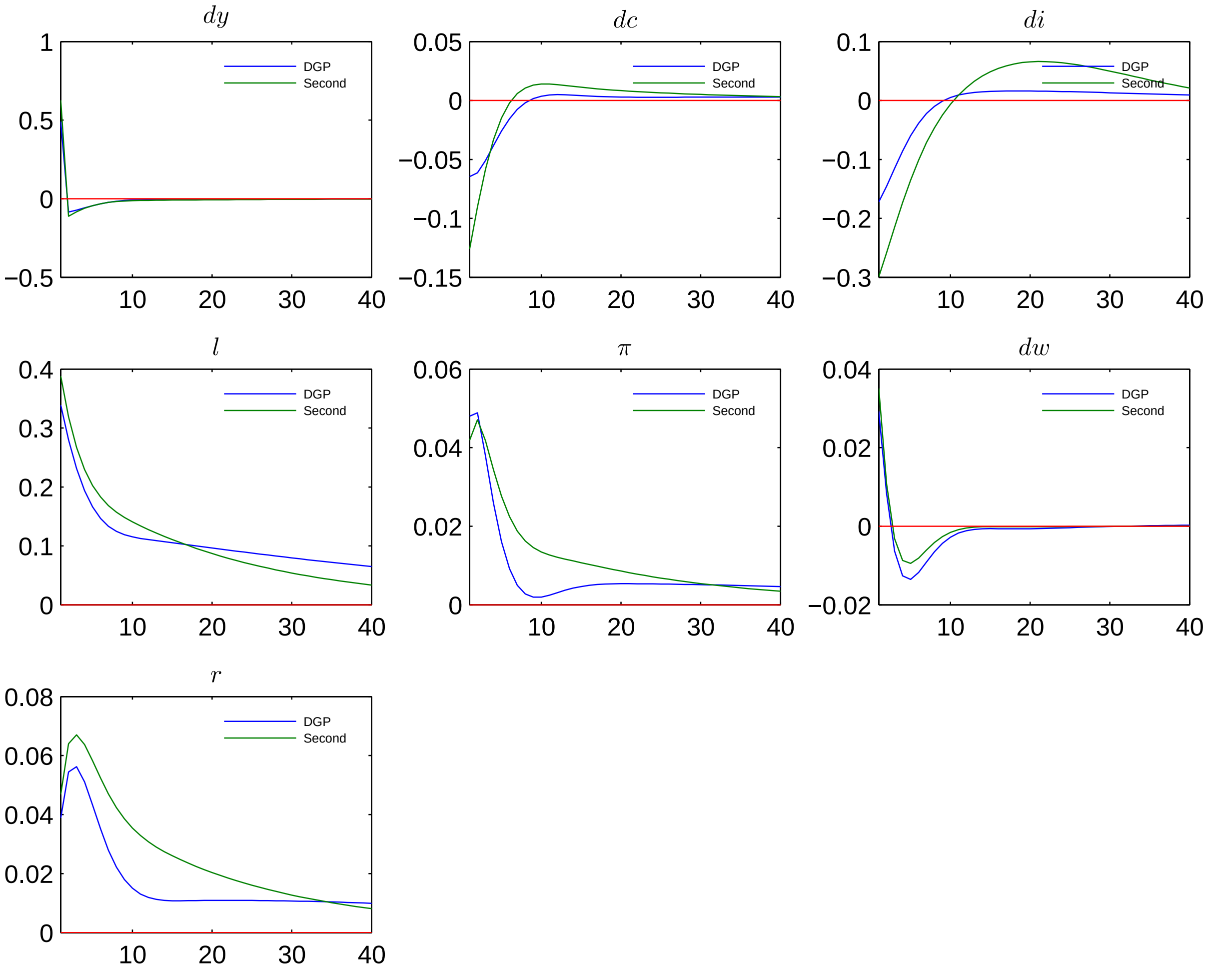

Figure C.3: Responses of the linearized model to a ε_g shock: DGP versus parameter estimated from second order data

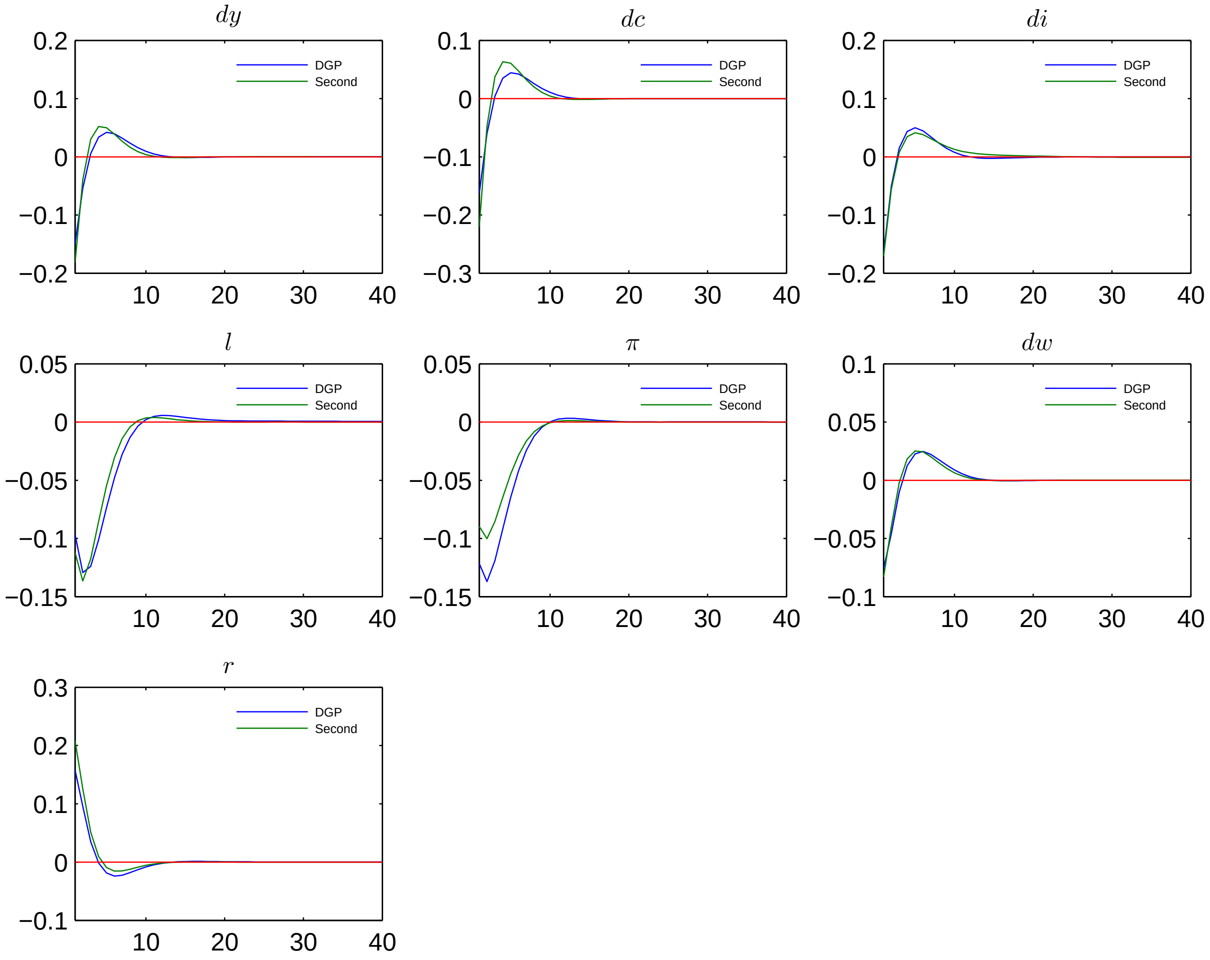

Figure C.4: Responses of the linearized model to a ε_m shock: DGP versus parameter estimated from second order data

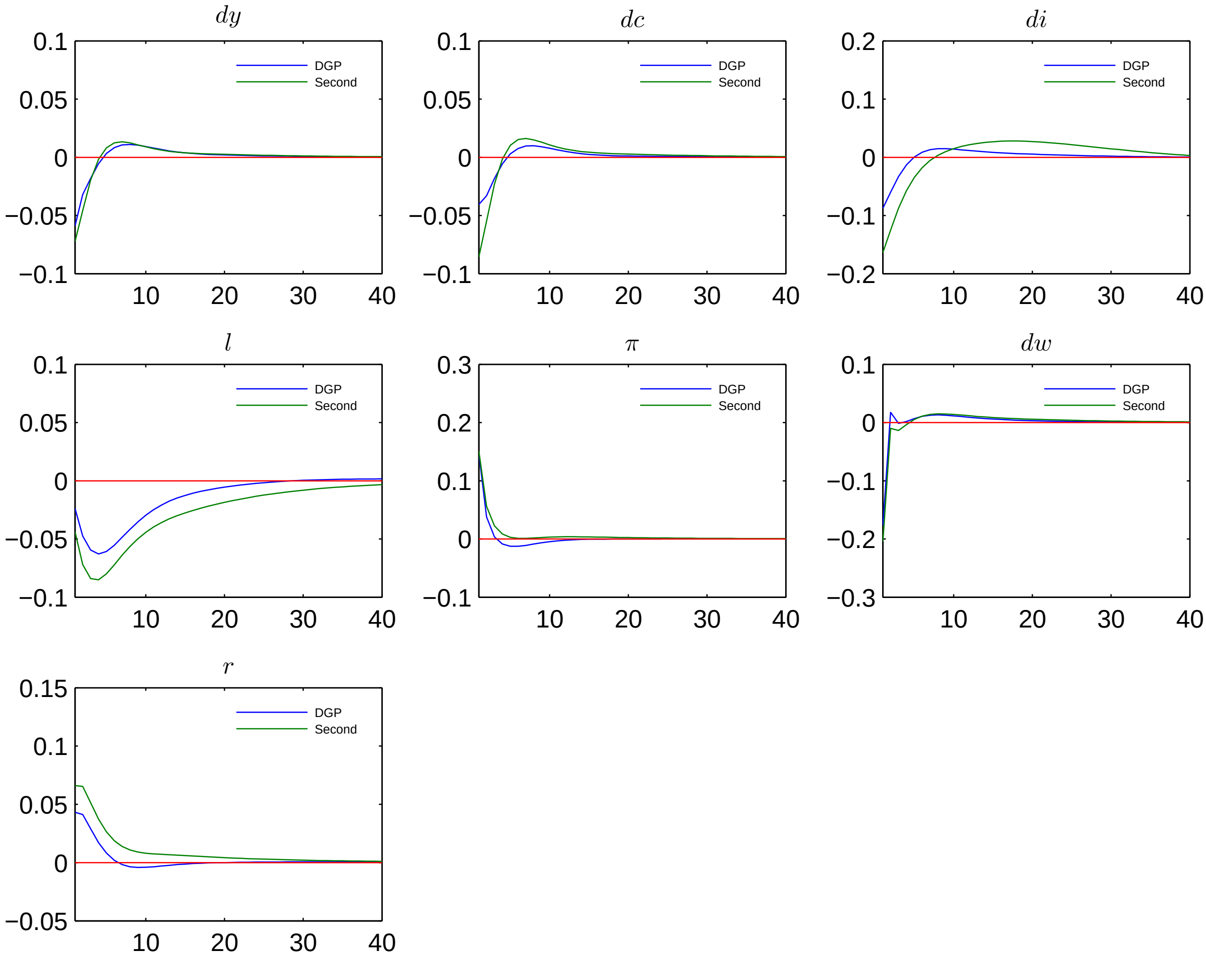

Figure C.5: Responses of the linearized model to a ε_π shock: DGP versus parameter estimated from second order data

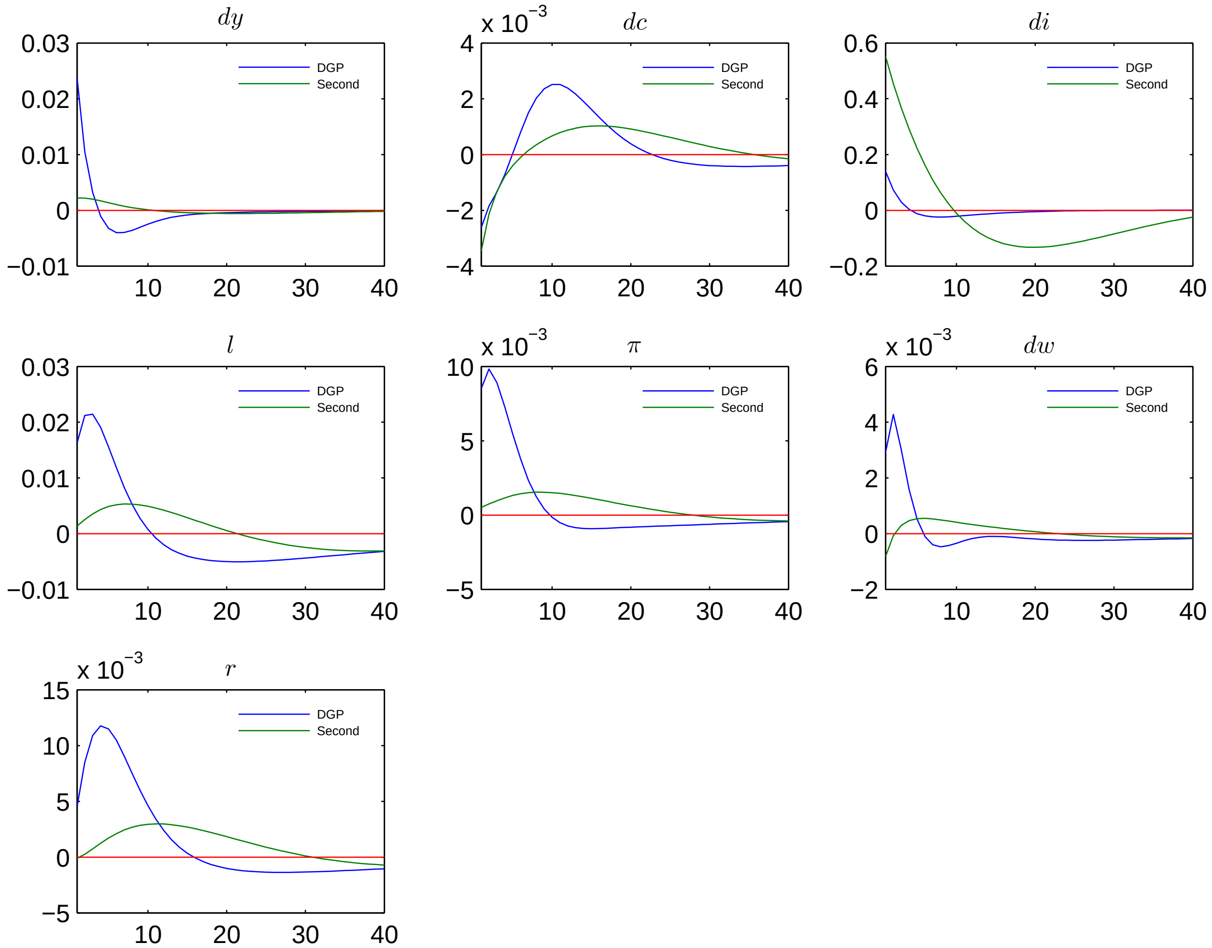

Figure C.6: Responses of the linearized model to a ε_i shock: DGP versus parameter estimated from second order data

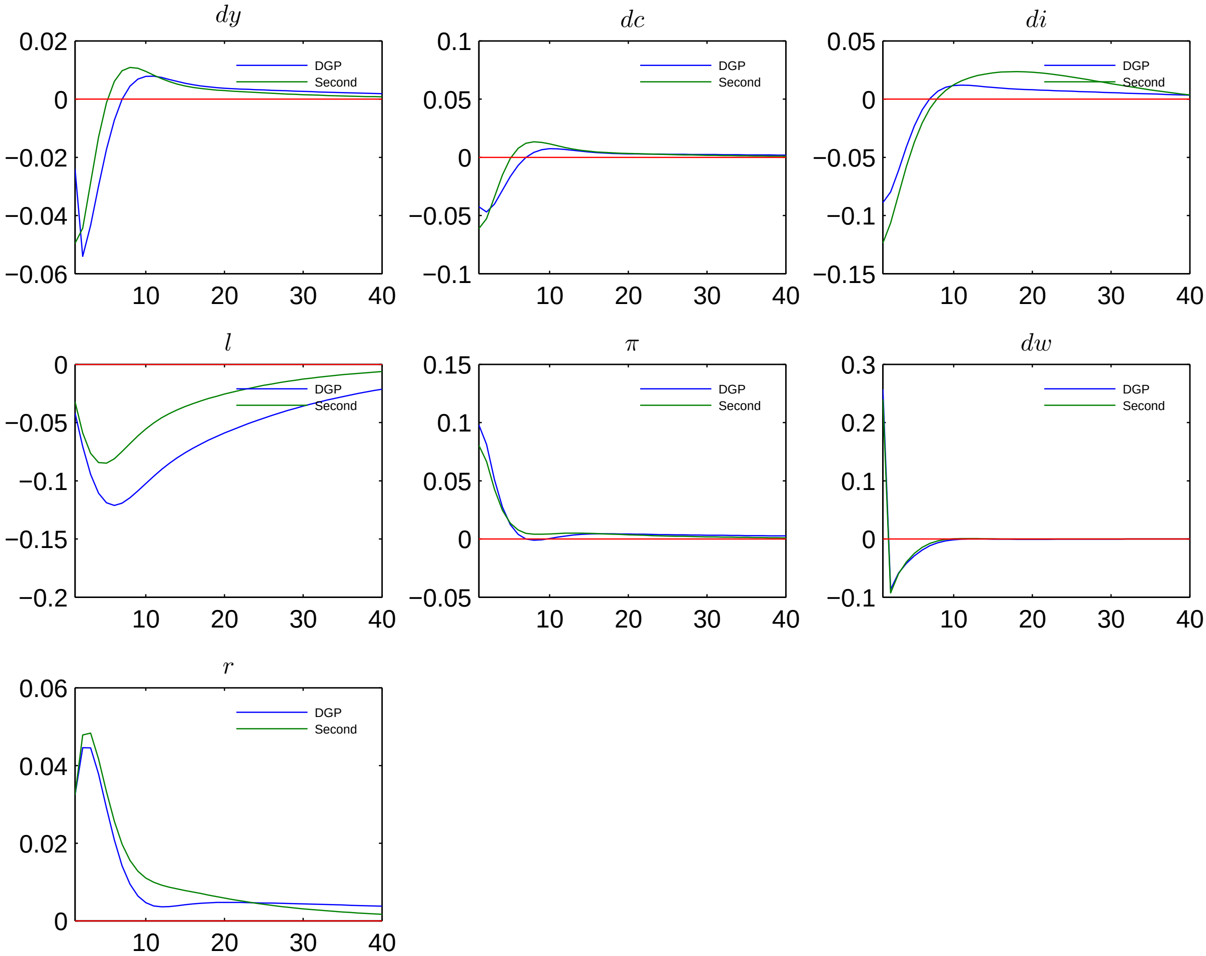

Figure C.7: Responses of the linearized model to a ε_w shock: DGP versus parameter estimated from second order data

Table C.1: Variance Decomposition: Parameter estimated from second order data

	ε_a	ε_b	ε_g	ε_i	ε_m	ε_p	ε_w
dy	13.09	5.24	74.61	0.01	6.45	0.44	0.16
dc	7.34	30.69	20.85	0.04	37.68	2.5	0.91
di	10.41	1.68	20.65	64.04	2.52	0.4	0.29
l	7.3	4.04	81.76	0.07	5.73	0.66	0.44
π	23.71	16.07	21.08	0.14	24.53	11.06	3.43
dw	13.72	12.85	5.45	0	18.19	23.61	26.18
r	22.75	21.36	27.99	0.17	21.88	4.09	1.77

Table C.2: Variance Decomposition: DGP

	ε_a	ε_b	ε_g	ε_i	ε_m	ε_p	ε_w
dy	35.46	0.59	54.49	0.16	6.62	1.17	1.52
dc	34.34	4.04	13.43	0.06	37.31	3.58	7.23
di	67.26	0.26	14.44	5.12	6.6	2.46	3.86
l	11.01	0.18	63.57	0.24	5.34	2.25	17.41
π	26.33	0.23	5.39	0.23	39.89	15.07	12.85
dw	23.58	0.33	0.98	0.02	6.29	20.09	48.7
r	44.44	0.66	16.07	0.64	27.87	3.64	6.68

C.2 Misspecification: Estimated Model has i.i.d. shocks

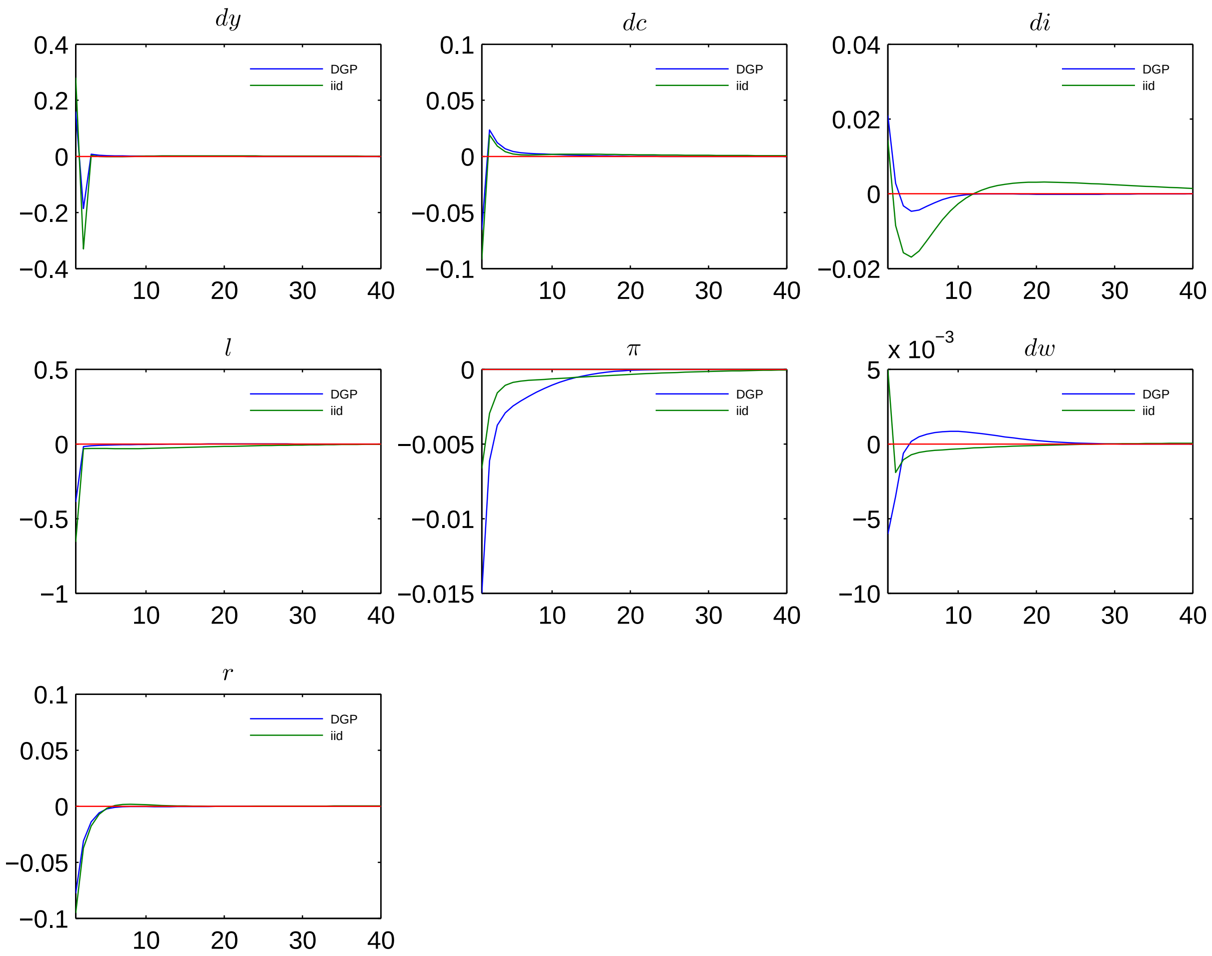

Figure C.8: Responses of the linearized model to a ε_a shock: DGP versus parameters estimated from i.i.d model

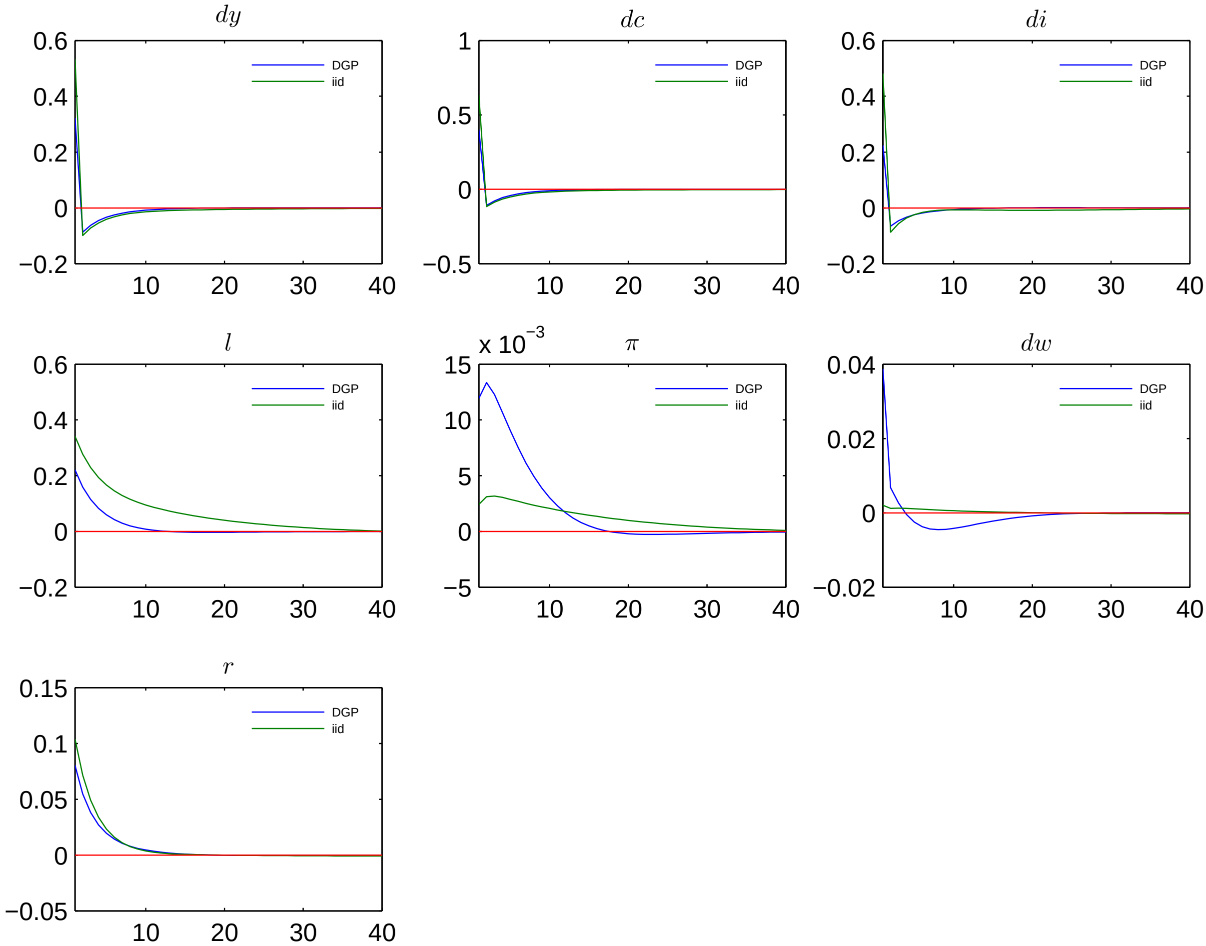

Figure C.9: Responses of the linearized model to a ε_b shock: DGP versus parameters estimated from i.i.d model

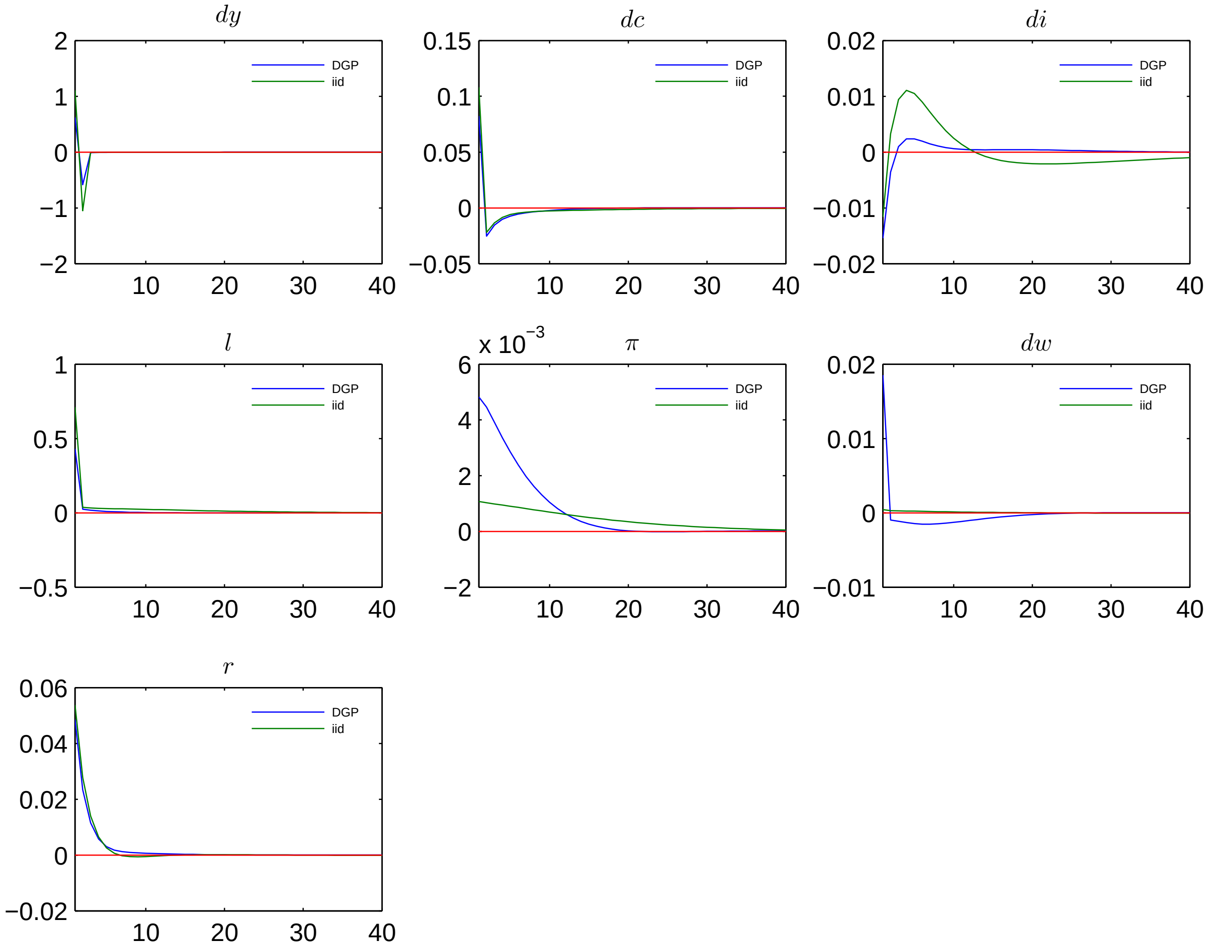

Figure C.10: Responses of the linearized model to a ε_g shock: DGP versus parameters estimated from i.i.d model

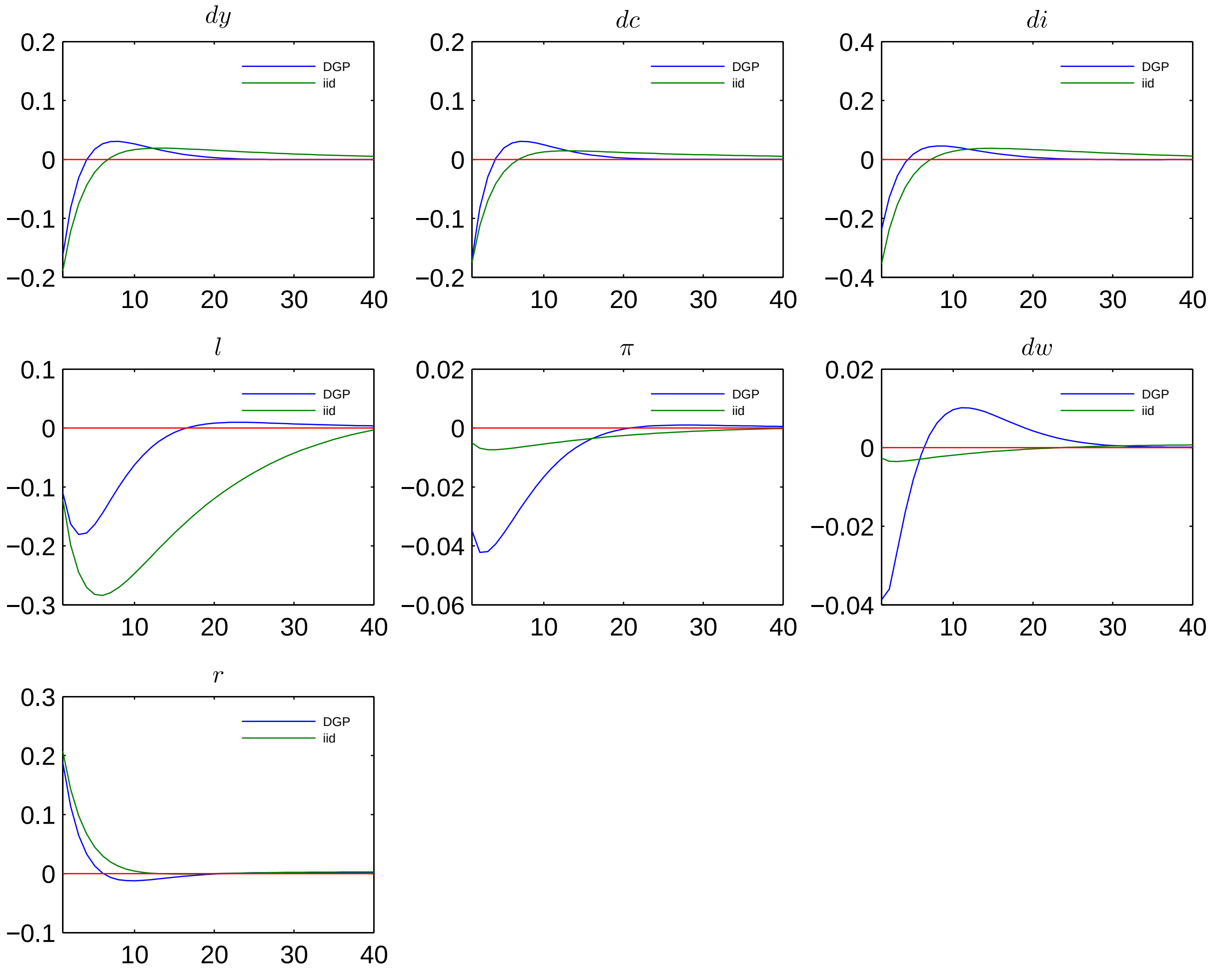

Figure C.11: Responses of the linearized model to a ε_m shock: DGP versus parameters estimated from i.i.d model

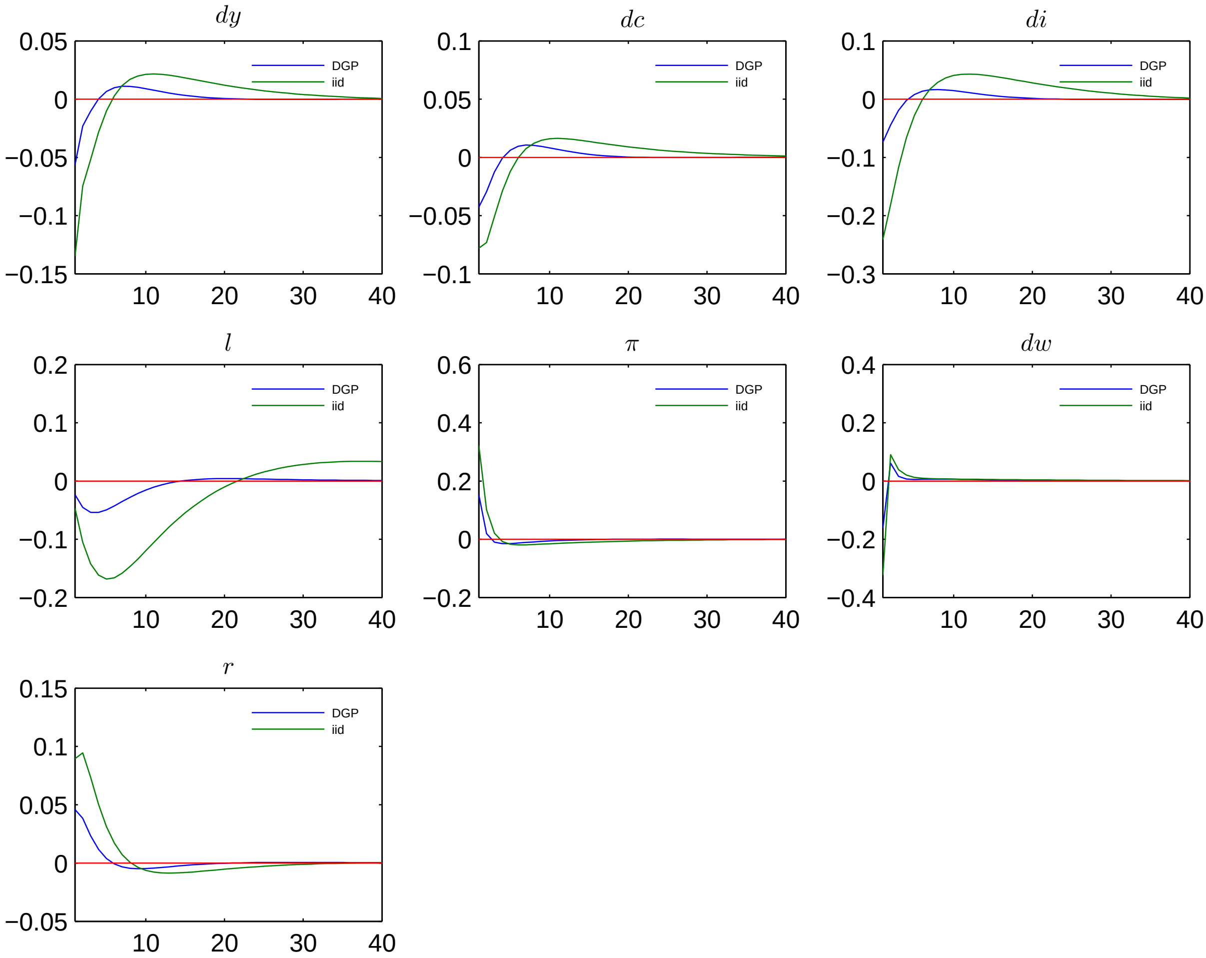

Figure C.12: Responses of the linearized model to a ε_π shock: DGP versus parameters estimated from i.i.d model

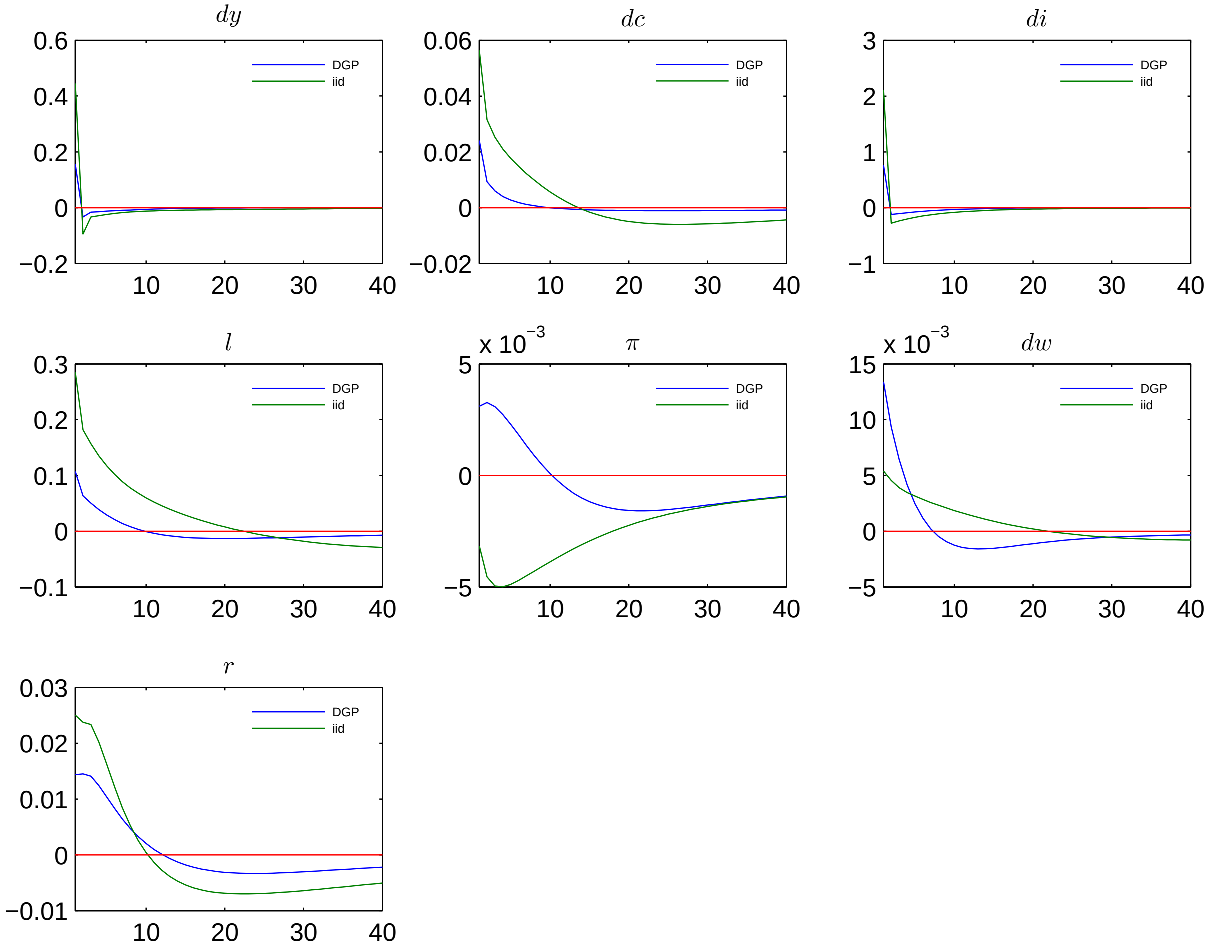

Figure C.13: Responses of the linearized model to a ε_i shock: DGP versus parameters estimated from i.i.d model

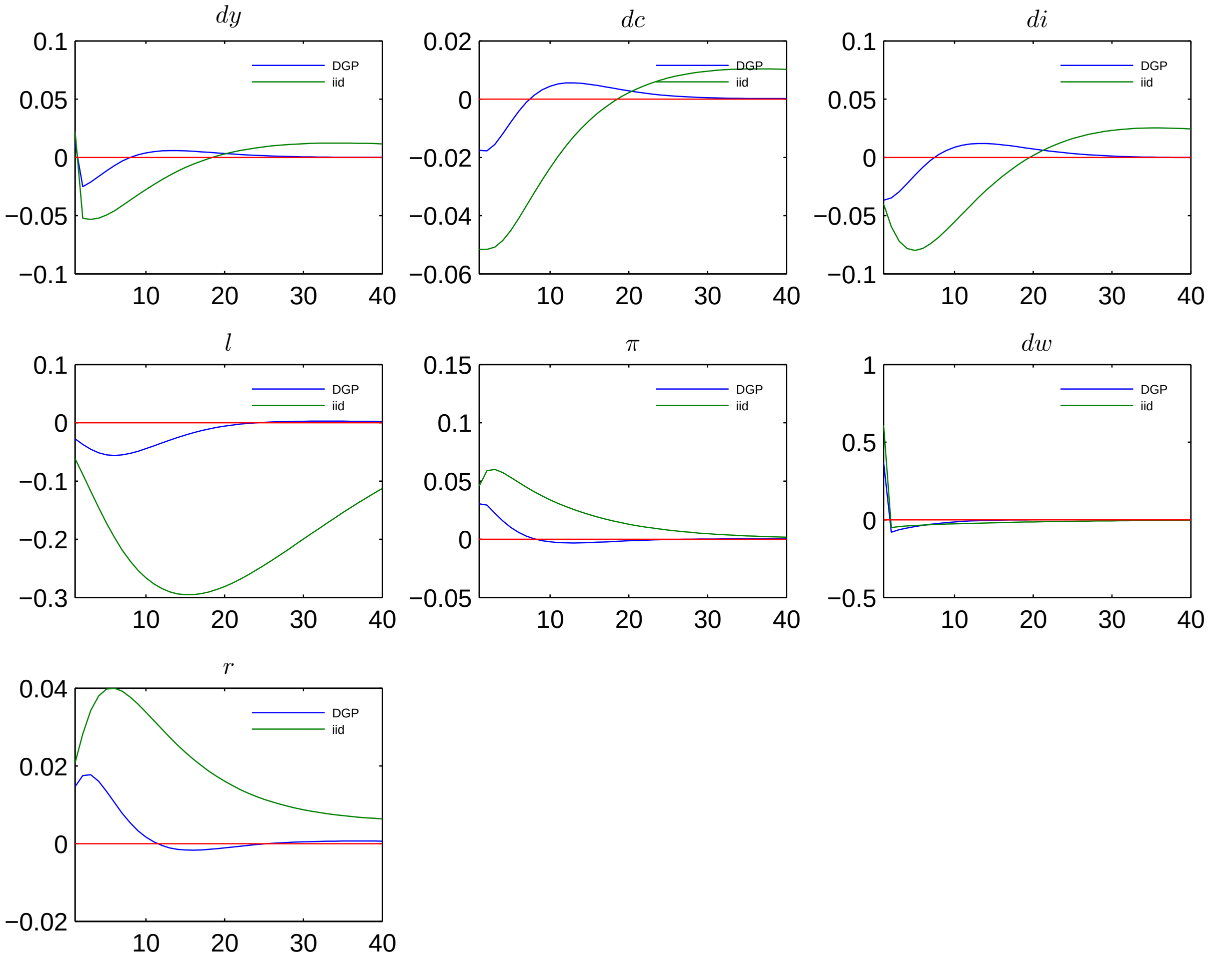

Figure C.14: Responses of the linearized model to a ε_w shock: DGP versus parameters estimated from i.i.d model

Table C.3: Variance Decomposition: Parameter estimated from i.i.d model

	ε_a	ε_b	ε_g	ε_i	ε_m	ε_p	ε_w
dy	5.96	9.71	73.78	6.72	2.05	1.03	0.75
dc	1.61	78.07	2.26	1.26	9.63	3.19	3.98
di	0.03	4.53	0.01	87.28	4.44	2.42	1.29
l	8.89	8.58	10.31	5.03	20.36	5.26	41.58
π	0.04	0.07	0.01	0.23	0.41	79.46	19.78
dw	0.01	0	0	0.03	0.02	22.77	77.17
r	6.42	12.15	2.32	2.43	48.36	15.93	12.39

Table C.4: Variance Decomposition: DGP

	ε_a	ε_b	ε_g	ε_i	ε_m	ε_p	ε_w
dy	6.18	12.25	74.06	2.71	4.13	0.47	0.2
dc	2.08	75.59	3.2	0.31	16.86	1.41	0.55
di	0.07	7.44	0.03	79.08	11.48	1.19	0.7
l	21.71	14.4	26.58	3.59	27.34	2.25	4.12
π	0.77	2.06	0.24	0.27	27.76	62	6.9
dw	0.03	0.88	0.19	0.19	2.43	15.69	80.59
r	8.52	14.68	3.64	1.53	64.64	5.2	1.79

D A nonlinear version of The Smets and Wouters (2007) model

D.1 Model Equations

The nonlinear version of SW07 model consist of the equations below. The notation is basically the same of the linear (and original) version (see tables A.2 and A.3), but some variables enter only the nonlinear model. This is the case for ξ (lagrange multiplier of the budget constraint) , y^{obs} (sum of the production by the intermediate firms [35]), l^{obs} (sum of intermediate labor), $\mathcal{A}_t$ (function of the capacity utilization cost), $D\mathcal{A}_t$ (derivative of the function with respect to the capacity utilization cost), S_t (investment adjustment cost), DS_t (derivative of the function with respect to the change in investment).

The other variables in capital letters represent integral and infinite summations in recursive form. Equations (D.4), (D.5) and (D.6) describe the Calvo price setting of the intermediate firms. Equations, (D.20), (D.21) and (D.22) describe the same for the intermediate labor unions. PDISP and WDISP represent, respectively, the price and wage dispersion of the economy.

$$y_t^{obs} = \varepsilon_t^a \, k_t^{s\,\alpha} \, l_t^{1-\alpha} - (\phi - 1)\, \bar{y}^{obs} \tag{D.1}$$

$$rk_t = l_t \, \frac{\frac{\alpha}{1-\alpha}\, w_t}{k_t^s} \tag{D.2}$$

$$mc_t = \frac{w_t^{1-\alpha}\, rk_t^{\alpha}}{\varepsilon_t^a \, \alpha^{\alpha}\, (1-\alpha)^{1-\alpha}} \tag{D.3}$$

$$SP_t = \frac{\xi_t \, y_t}{\lambda_p - 1}\, \tilde{p}_t^{\frac{(-1)}{\lambda_p - 1}} + \zeta_p \, \bar{\beta}\, \gamma \left(\frac{\pi_t^{\iota_p}\, \bar{\pi}^{1-\iota_p}}{\pi_{t+1}} \, \frac{\tilde{p}_t}{\tilde{p}_{t+1}} \right)^{\frac{(-1)}{\lambda_p - 1}} SP_{t+1} \tag{D.4}$$

$$SMC_t = y_t \, mc_t \, \xi_t \, \frac{\lambda_p}{\lambda_p - 1}\, \tilde{p}_t^{\frac{(-\lambda_p)}{\lambda_p - 1}} + \zeta_p \, \bar{\beta}\, \gamma \left(\frac{\pi_t^{\iota_p}\, \bar{\pi}^{1-\iota_p}}{\pi_{t+1}} \, \frac{\tilde{p}_t}{\tilde{p}_{t+1}} \right)^{\frac{(-\lambda_p)}{\lambda_p - 1}} SMC_{t+1} \tag{D.5}$$

$$SP_t = SMC_t \, \varepsilon_t^p \tag{D.6}$$

$$\tilde{p}_t = \left(\frac{1 - \zeta_p \left(\frac{\bar{\pi}^{1-\iota_p}\, \pi_{t-1}^{\iota_p}}{\pi_t} \right)^{\frac{1}{\lambda_p - 1}}}{1 - \zeta_p} \right)^{\lambda_p - 1} \tag{D.7}$$

$$\xi_t = \frac{\sigma_c - 1}{1 + \sigma_l}\, l_t^{obs\,1+\sigma_l} \left(c_t - \frac{h}{\gamma}\, c_{t-1} \right)^{(-\sigma_c)} (1 - h\,\beta) \tag{D.8}$$

$$w_t^h = \frac{\left(c_t - \frac{h}{\gamma}\, c_{t-1} \right) l_t^{obs\,\sigma_l}}{1 - h\,\beta} \tag{D.9}$$

$$\xi_t = \frac{\bar{\beta}\, r_t \, \xi_{t+1}}{\pi_{t+1}}\, \varepsilon_t^b \tag{D.10}$$

[35]See SW07 online appendix (p.8) for more details. In the linear case, $y = y^{obs}$

$$k^s_t = \frac{z_t\, k_{t-1}}{\gamma} \tag{D.11}$$

$$S_t = 0.5\, \left(exp\left(\sqrt{\varphi}\, \left(\gamma\, \frac{i_t}{i_{t-1}} - \gamma\right)\right) + exp\left(\left(\gamma\, \frac{i_t}{i_{t-1}} - \gamma\right)\, (-\sqrt{\varphi})\right) - 2\right) \tag{D.12}$$

$$DS_t = 0.5\, \left(\sqrt{\varphi}\, exp\left(\sqrt{\varphi}\, \left(\gamma\, \frac{i_t}{i_{t-1}} - \gamma\right)\right) - \sqrt{\varphi}\, exp\left(-\sqrt{\varphi}\left(\gamma\, \frac{i_t}{i_{t-1}} - \gamma\right)\right)\right) \tag{D.13}$$

$$k_t = k_{t-1}\, \frac{1-\delta}{\gamma} + i_t\, \varepsilon^i_t\, (1 - S_t) \tag{D.14}$$

$$1 = \varepsilon^i_t\, q_t\, \left(1 - S_t - \frac{i_t}{i_{t-1}}\, \gamma\, DS_t\right) + \bar{\beta}\, \left(\frac{\xi_{t+1}}{\xi_t}\, q_{t+1}\, \varepsilon^i_{t+1}\, DS_{t+1}\, \left(\gamma\, \frac{i_{t+1}}{i_t}\right)^2\right) \tag{D.15}$$

$$\mathcal{A}_t = .5\, \bar{rk}_t\, \sigma_{\mathcal{A}}\, z^2_t + z_t\, \bar{rk}_t\, (1 - \sigma_{\mathcal{A}}) + \bar{rk}_t\, \left(\frac{\sigma_{\mathcal{A}}}{2} - 1\right) \tag{D.16}$$

$$\mathcal{DA}_t = \bar{rk}_t\, (1 - \sigma_{\mathcal{A}}) + z_t\, \bar{rk}_t\, \sigma_{\mathcal{A}} \tag{D.17}$$

$$q_t = \bar{\beta}\left(\frac{\xi_{t+1}}{\xi_t}\, (rk_{t+1}\, z_{t+1} - \mathcal{A}_{t+1} + (1-\delta)\, q_{t+1})\right) \tag{D.18}$$

$$rk_t = \mathcal{DA}_t \tag{D.19}$$

$$SWH_t = l_t\, w^h_t\, \frac{\xi_t\, \lambda_w}{\lambda_w - 1}\, \left(\frac{\tilde{w}_t}{w_t}\right)^{\frac{(-\lambda_w)}{\lambda_w - 1}} + \gamma\, \bar{\beta}\, \zeta_w\, \left(\frac{\tilde{w}_t\, \frac{\pi_t^{\iota_w}\, \bar{\pi}^{1-\iota_w}}{\pi_{t+1}}}{\tilde{w}_{t+1}}\right)^{\frac{(-\lambda_w)}{\lambda_w - 1}} SWH_{t+1} \tag{D.20}$$

$$SW_t = l_t\, \frac{\xi_t}{\lambda_w - 1}\, \tilde{w}_t^{\frac{(-1)}{\lambda_w - 1}}\, \left(\frac{1}{w_t}\right)^{\frac{(-\lambda_w)}{\lambda_w - 1}} + \gamma\, \bar{\beta}\, \zeta_w\, \left(\frac{\tilde{w}_t\, \frac{\pi_t^{\iota_w}\, \bar{\pi}^{1-\iota_w}}{\pi_{t+1}}}{\tilde{w}_{t+1}}\right)^{\frac{(-1)}{\lambda_w - 1}} SW_{t+1} \tag{D.21}$$

$$SW_t = SWH_t\, \varepsilon^w_t \tag{D.22}$$

$$\tilde{w}_t = \left(\frac{w_t^{\frac{(-1)}{\lambda_w - 1}} - \zeta_w\, \left(\frac{\bar{\pi}^{1-\iota_w}\, \pi_{t-1}^{\iota_w}}{\pi_t}\, w_{t-1}\right)^{\frac{(-1)}{\lambda_w - 1}}}{1 - \zeta_w}\right)^{(-(\lambda_w - 1))} \tag{D.23}$$

$$y_t = c_t + i_t + \bar{y}\, (c_g + \varepsilon^g_t) + \frac{k_{t-1}\, \mathcal{A}_t}{\gamma} \tag{D.24}$$

$$PDISP_t = \tilde{p}_t^{\frac{(-\lambda_p)}{\lambda_p - 1}}\, (1 - \zeta_p) + \zeta_p\, \left(\frac{\bar{\pi}^{1-\iota_p}\, \pi_{t-1}^{\iota_p}}{\pi_t}\right)^{\frac{(-\lambda_p)}{\lambda_p - 1}} PDISP_{t-1} \tag{D.25}$$

$$y^{obs}_t = y_t\, PDISP_t \tag{D.26}$$

$$WDISP_t = \left(\frac{\tilde{w}_t}{w_t}\right)^{\frac{(-\lambda_w)}{\lambda_w - 1}}\, (1 - \zeta_w) + \zeta_w\, \left(\frac{\bar{\pi}^{1-\iota_w}\, \pi_{t-1}^{\iota_w}\, \frac{w_{t-1}}{w_t}}{\pi_t}\right)^{\frac{(-\lambda_w)}{\lambda_w - 1}} WDISP_{t-1} \tag{D.27}$$

$$l^{obs}_t = l_t\, WDISP_t \tag{D.28}$$

$$\frac{r_t}{\bar{r}} = \left(\frac{r_{t-1}}{\bar{r}}\right)^{\rho}\, \left(\left(\frac{\pi_t}{\bar{\pi}}\right)^{r_\pi}\, \left(\frac{y^{obs}_t}{{y^f}_t}\right)^{r_y}\right)^{1-\rho}\, \left(\frac{\frac{y^{obs}_t}{{y^f}_t}\, y^{obs}_{t-1}}{{y^f}_{t-1}}\right)^{r_{\Delta y}} \varepsilon^r_t \tag{D.29}$$

$$y^f{}_t = \varepsilon_t^a\, k^{sf}{}_t^{\alpha}\, l^f{}_t^{1-\alpha} - (\phi - 1)\ \bar{y}^f \tag{D.30}$$

$$rk^f{}_t = l^f{}_t\ \frac{\frac{\alpha}{1-\alpha}\, w^f{}_t}{k_t^{sf}} \tag{D.31}$$

$$\frac{1}{\lambda_p} = \frac{w^f{}_t^{1-\alpha}\, rk^f{}_t^{\alpha}}{\varepsilon_t^a\, \alpha^\alpha\ (1-\alpha)^{1-\alpha}} \tag{D.32}$$

$$\xi^f{}_t = (1 - h\,\beta)\ \frac{\sigma_c - 1}{1+\sigma_l}\, l^f{}_t^{1+\sigma_l}\ \left(c^f{}_t - \frac{h}{\gamma}\, c^f{}_{t-1}\right)^{(-\sigma_c)} \tag{D.33}$$

$$\xi^f{}_t = \varepsilon_t^b\, \bar{\beta}\, r^f{}_t\, \xi^f{}_{t+1} \tag{D.34}$$

$$k^{sf}{}_t = \frac{z^f{}_t\, k^f{}_{t-1}}{\gamma} \tag{D.35}$$

$$S_t^f = 0.5\ \left(exp\left(\sqrt{\varphi}\ \left(\gamma\, \frac{i_t^f}{i_{t-1}^f} - \gamma\right)\right) + exp\left(\left(\gamma\, \frac{i_t^f}{i_{t-1}^f} - \gamma\right)\ (-\sqrt{\varphi})\right) - 2\right) \tag{D.36}$$

$$DS_t^f = 0.5\ \left(\sqrt{\varphi}\, exp\left(\sqrt{\varphi}\ \left(\gamma\, \frac{i_t^f}{i_{t-1}^f} - \gamma\right)\right) - \sqrt{\varphi}\, exp\left(-\sqrt{\varphi}\left(\gamma\, \frac{i_t^f}{i_{t-1}^f} - \gamma\right)\right)\right) \tag{D.37}$$

$$k^f{}_t = \frac{1-\delta}{\gamma}\, k^f{}_{t-1} + i^f{}_t\, \varepsilon_t^i\ \left(1 - S^f{}_t\right) \tag{D.38}$$

$$1 = \varepsilon_t^i\, q^f{}_t\ \left(1 - S^f{}_t - \frac{i^f{}_t}{i^f{}_{t-1}}\, \gamma\, DS^f{}_t\right) + \bar{\beta}\ \left(\frac{\xi^f{}_{t+1}}{\xi^f{}_t}\, \varepsilon_{t+1}^i\, q^f{}_{t+1}\, DS^f{}_{t+1}\ \left(\gamma\, \frac{i^f{}_{t+1}}{i^f{}_t}\right)^2\right) \tag{D.39}$$

$$\mathcal{A}_t^f = \sigma_{\mathcal{A}}\, .5\, r\bar{k}^f{}_t\, z^f{}_t^2 + z^f{}_t\ (1-\sigma_{\mathcal{A}})\ r\bar{k}^f{}_t + \left(\frac{\sigma_{\mathcal{A}}}{2} - 1\right)\ r\bar{k}^f{}_t \tag{D.40}$$

$$\mathcal{DA}_t^f = (1-\sigma_{\mathcal{A}})\ r\bar{k}^f{}_t + z^f{}_t\, \sigma_{\mathcal{A}}\, r\bar{k}^f{}_t \tag{D.41}$$

$$q^f{}_t = \bar{\beta}\ \left(\frac{\xi_{t+1}^f}{\xi^f{}_t}\ \left(rk^f{}_{t+1}\, z^f{}_{t+1} - \mathcal{A}_{t+1}^f\right) + (1-\delta)\ q^f{}_{t+1}\right) \tag{D.42}$$

$$rk^f{}_t = \mathcal{DA}_t^f \tag{D.43}$$

$$w^f{}_t = \frac{\lambda_w\ \left(c^f{}_t - \frac{h}{\gamma}\, c^f{}_{t-1}\right)\ l^f{}_t^{\sigma_l}}{1 - h\,\beta} \tag{D.44}$$

$$y^f{}_t = c^f{}_t + i^f{}_t + (c_g + \varepsilon_t^g\,)\ \bar{y}^f + \frac{k^f{}_{t-1}\, \mathcal{A}_t^f}{\gamma} \tag{D.45}$$

$$\varepsilon_t^a = \rho_a\, \varepsilon_{t-1}^a + \eta_t^a \tag{D.46}$$

$$\varepsilon_t^b = \rho_b\, \varepsilon_{t-1}^b + \eta_t^b \tag{D.47}$$

$$\varepsilon_t^g = \rho_g\, \varepsilon_{t-1}^g + \eta_t^g + \eta_t^a\, \rho_{ga} \tag{D.48}$$

$$\varepsilon_t^i = \rho_I\, \varepsilon_{t-1}^i + \eta_t^I \tag{D.49}$$

$$\varepsilon_t^r = \rho_r\, \varepsilon_{t-1}^r + \eta_t^m \tag{D.50}$$

$$\varepsilon_t^p = \rho_p\, \varepsilon_{t-1}^p + \eta_t^p - \mu_p\, \eta_{t-1}^p \tag{D.51}$$

$$\varepsilon_t^w = \rho_w\, \varepsilon_{t-1}^w + \eta_t^w - \mu_w\, \eta_{t-1}^w \tag{D.52}$$

$$dy_t^{obs} = y_t^{obs} - y_{t-1}^{obs} + \bar{\gamma} \tag{D.53}$$

$$dc_t = \bar{\gamma} + c_t - c_{t-1} \tag{D.54}$$

$$di_t = \bar{\gamma} + i_t - i_{t-1} \tag{D.55}$$

$$dw_t = \bar{\gamma} + w_t - w_{t-1} \tag{D.56}$$

$$\pi_t^{obs} = \pi_t\, 100 \tag{D.57}$$

$$r_t^{obs} = r_t\, 100 \tag{D.58}$$