What to expect when you're calibrating: measuring the effect of calibration on the estimation of macroeconomic models

Nikolay Iskrev

Abstract

I propose two measures of the impact of calibration on the estimation of macroeconomic models. The first quantifies the amount of information introduced with respect to each estimated parameter as a result of fixing the value of one or more calibrated parameters. The second is a measure of the sensitivity of parameter estimates to perturbations in the calibration values. The purpose of the measures is to show researchers how much and in what way calibration affects their estimation results – by shifting the location and reducing the spread of the marginal posterior distributions of the estimated parameters. Such analysis is often appropriate since macroeconomists do not always agree on whether and how to calibrate structural parameters in macroeconomic models. The methodology is illustrated using the models estimated in Smets and Wouters (2007) and Schmitt-Grohé and Uribe (2012).

Keywords: DSGE models, information content, calibration, estimation, identification

JEL classification: C32, C51, C52, E32

Bank of Portugal, Research in Economics and Mathematics (REM), and Research Unit on Complexity and Economics (UECE). UECE is supported by Fundação para a Ciência e a Tecnologia.
Av. Almirante Reis 71, 1150-012, Lisboa, Portugal, +351213130006, nikolay.iskrev@bportugal.pt, https://github.com/niskrev
Acknowledgments: I would like to thank the editor, Thomas Lubik, and two anonymous referees for their valuable comments and suggestions. I would also like to thank the participants at the 71st European Meeting of the Econometric Society, the 2018 North American Summer Meeting of the Econometric Society, the 2018 Computing in Economics and Finance conference, and the 14th annual DYNARE Conference.

1 Introduction

It is a common practice in the empirical macroeconomic literature to mix estimation of some model parameters with calibration of others. The rationale behind this approach is either that some parameters are difficult to identify from available data, or that their values have been well-established elsewhere in the literature. While these may be reasonable arguments in some cases, the list of calibrated parameters often includes some for which the empirical evidence is far from settled, and whose values are simply taken from previous studies, often based on very different models and data patterns. Convenience and ease of estimation may be a more plausible explanation of the common practice of fixing some parameters than the possession of true knowledge of their values. It is therefore important to understand the impact, if any, parameter calibration has on model estimation.

The practice of mixing calibration and estimation can have two potentially important consequences. First, the values of the calibrated parameters may affect the point estimates of the free parameters.[1] Thus, mis-calibration could result in biased estimates of some estimated parameters. Second, from the point of view of estimation, calibration of some parameters is equivalent to assuming that their values are known. This may introduce information about parameters that are estimated. Put differently, by eliminating all uncertainty with respect to calibrated parameters, one may also remove some of the uncertainty about freely estimated parameters.

Clearly, not all free parameters are affected equally by calibration. In general, the size of the impact will depend on the interactions between free and calibrated parameters in the context of a given model. Except in very simple cases with a small number of parameters, it is generally difficult to identify, by intuition or heuristic reasoning alone, which estimated parameters will be affected, in what way and by how much, as a result of calibrating one or more model parameters.

One possible way of quantifying the amount of information introduced by calibration is to re-estimate the model in the absence of calibration, and compare the resulting uncertainty with that of the restricted model. Similarly, the effect of changing the calibration values can be assessed be re-estimating the model multiple times conditional on different values of the fixed parameters. Whether or not these are reasonable ways to proceed depends on how feasible it is to estimate the larger unrestricted model, or to

[1] Or, in Bayesian context, the location of the posterior distribution of the estimated parameters.

estimate multiple times the restricted model, and also how strongly one feels about the reasons for calibration in the first place. Note that estimating the unrestricted model is almost certain to result in point estimates of the previously fixed parameters that are different from the calibration values. This might be undesirable if one has strong views about what those values should be. Furthermore, the point estimates of at least some freely estimated parameters are likely to be different in the unrestricted model. This will complicate the comparison of the estimation uncertainty in the restricted and unrestricted cases.[2]

The purpose of this paper is to present an alternative approach, which does not require estimating models more than once, and only uses the estimation results under the original calibration. The method is based on the asymptotic posterior distribution of the parameters in the unrestricted case, which is used to construct two different measures. The first is a measure of the amount of information gained with respect to each free parameter as a result of knowing the value of one or more calibrated parameters. It shows the reduction of asymptotic uncertainty as a percent of the uncertainty in the unrestricted case. The second is a measure of the sensitivity of parameter estimates to perturbations in the values of different calibrated parameters. In particular, it shows the sign and the magnitude of the response of different estimated parameters to changes in the values of the calibrated ones.

The intuition behind the proposed approach is simple: the effect of calibration will depend on how different parameters interact in a given model. From the point of view of estimation, these interactions are captured by the parameters' impact on the model log-likelihood function. Closely-related parameters are difficult to distinguish on the basis of their effect on the log-likelihood. Fixing one or more of them provides a lot of information about the other related parameters, which are also very responsive to changes in the calibration values. The opposite holds true for unrelated parameters whose effects on the likelihood function are orthogonal to each other. For instance, consider a standard business cycle model. In such models there are typically a few parameters that determine the steady state of the economy. Calibrating some of them will naturally have a stronger impact on the other steady state-related parameters, both in terms of location

[2]It is straightforward to think of examples where, because of the choice of calibration values of the fixed parameters, the estimation uncertainty is much larger than it would be if those parameters were estimated instead. For instance, if two parameters are nearly unidentifiable when a third one is in a particular region of the parameter space, but very well identified elsewhere, estimation uncertainty will be much smaller if the unrestricted model is in a well-identified part of the parameter space, compared to a restricted model with calibrated value from the poorly identified region.

and spread of their posterior distribution. On the other hand, more weakly-related parameters, such as variance coefficients of shocks, are likely to be unaffected.

The measures I propose formalize this intuition. Specifically, I use the asymptotic Gaussianity of the posterior distribution of the model parameters, and study the effect of calibration by comparing the mean and variance of the distribution in the unrestricted case to the same moments in the restricted case, i.e. conditional on some parameters being known and fixed. Simple closed-form expressions show that the impact of calibration depends on the model-implied interdependence between free and calibrated parameters, which is captured by the correlation structure of the asymptotic posterior distribution.

From a Bayesian perspective, calibration of some model parameters could be interpreted as having very strong prior beliefs about the values of those parameters. In this sense, my paper is similar to Müller (2012), who proposed measures of prior sensitivity and prior informativeness in Bayesian models. As Müller (2012) observes, "likelihood information about different parameters can be far from independent, so that the marginal posterior distributions crucially depend on the interaction of the likelihood with the whole prior." The same argument implies that calibrating some parameters can have a significant impact on the posterior distributions of freely-estimated parameters. Unlike the sensitivity and informativeness measures in this paper, the measures of Müller (2012) cannot be applied to parameters that are held fixed during estimation since computing them requires sampling from the posterior distribution of the full parameter vector. As noted earlier, combining estimation, both frequentist and Bayesian, with calibration is a rather common practice in the DSGE literature, which makes my contribution complementary to that of Müller (2012).[3]

In terms of methodology, my paper is most closely related to Andrews et al. (2017), who introduced a measure of sensitivity of parameter estimates to the empirical moments they are based on. The purpose of their analysis is to identify the most influential moments, which, if misspecified, could result in a large estimation bias. Even though my measure of sensitivity is with respect to calibrated parameters and not moments, its derivation is based on the same idea: I use the joint asymptotic distribution of free and calibrated parameters, whereas Andrews et al. (2017) use the joint asymptotic

[3]My measures also have somewhat different interpretations from those of Müller (2012). In particular, I measure the amount of information due to calibration by comparing posterior uncertainty with and without calibration, while Müller (2012) compares the posterior to the prior uncertainty. Also, my sensitivity measure shows not only the magnitude of the effect of perturbations in the calibration values, but also the sign of the effect. Müller's (2012) sensitivity only shows the magnitude.

distribution of free parameters and empirical moments. In both cases sensitivity is measured locally and can be used as an indicator of how robust the estimation results are to small perturbations in either the calibration values or the moment conditions. My paper also shares Andrews et al. (2017) larger goal, namely, to help increase the transparency of estimated structural models by providing easy-to-use tools for assessing the importance of different estimation assumptions. In the context of DSGE models, I believe it is important for researchers to discuss not only the reasons for and methods of calibration, but also the likely impact of calibration on the estimation results. The measures derived in this paper serve precisely that purpose and can be easily incorporated into the standard estimation output usually reported in empirical DSGE research.

The remainder of the paper is organized as follows. Section 2 defines and motivates my measures of information gains and sensitivity. Section 3 illustrates the use of the proposed measures using two different DSGE models. The models are a new Keynesian model estimated in Smets and Wouters (2007), and a real business cycle model with news shocks estimated in Schmitt-Grohé and Uribe (2012). In each case I show how calibration used by the authors affects their estimation results. Section 4 offers some concluding remarks.

2 Methodology

This section describes the methodology I use to measure the impact of calibration of some parameters on the estimation of the free parameters of a model. I assume the following setup: a researcher has a model that fully characterizes the density function $p_T(\boldsymbol{y}_T|\boldsymbol{\theta})$ of a data vector $\boldsymbol{Y}_T = (Y_1, \ldots, Y_T)$, as a function of a parameter vector $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^{n_\theta}$. The true value of $\boldsymbol{\theta}$ is unknown, and is estimated using maximum likelihood or Bayesian methods subject to the restriction that some elements of $\boldsymbol{\theta}$ are known, and are therefore held fixed in the estimation. Further, I assume that estimation of the full set of parameters is either not feasible or too costly. Hence, the objective is to characterize the consequences of calibration using only the estimates of the constrained model.

2.1 Asymptotic normality of the posterior distribution

A well-known property of Bayesian estimation procedures is that, asymptotically, they inherit the properties of the classical maximum likelihood estimator. This is because the variation in the prior distribution is dominated by the variation in the likelihood function, resulting in a posterior distribution whose shape moves arbitrarily close to the shape of the likelihood function. Hence, asymptotically, the posterior distribution is Gaussian centered at the maximum likelihood estimate with covariance matrix equal to the inverse of the expected Fisher's information matrix. This result is commonly known as the Bernstein-Von Mises theorem, first established for independent data by Walker (1969), and extended to stationary time series by Heyde and Johnstone (1979) and Chen (1985), and to non-stationary time series by Phillips and Ploberger (1996) and Kim (1998).

More formally, suppose that $\hat{\boldsymbol{\theta}}$ is the maximum likelihood estimate of $\boldsymbol{\theta}$ and that $\widehat{\mathcal{I}}$ is the expected Fisher's information matrix evaluated at $\hat{\boldsymbol{\theta}}$, i.e.

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\operatorname{argmax}}\, p_T\left(\boldsymbol{y}_T | \boldsymbol{\theta}\right) \tag{2.1}$$

$$\widehat{\mathcal{I}} = -\lim_{T \to \infty} \frac{1}{T} \operatorname{E}\left[\frac{\partial^2 \log p_T(\boldsymbol{y}_T | \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right] \tag{2.2}$$

Let $\pi(\boldsymbol{\theta})$ be the prior density of $\boldsymbol{\theta}$. Then, the posterior density is defined as

$$\pi_T\left(\boldsymbol{\theta} | \boldsymbol{Y}_T\right) = \frac{p_T(\boldsymbol{Y}_T | \boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\int_{\boldsymbol{\Theta}} p_T(\boldsymbol{Y}_T | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{2.3}$$

Under suitable regularity conditions and for large T, the posterior distribution of $\boldsymbol{\theta}$ is approximately equal to the normal density with mean $\hat{\boldsymbol{\theta}}$ and covariance matrix $\widehat{\boldsymbol{\Sigma}}$ given by the inverse of the Fisher's information matrix

$$\pi_T\left(\boldsymbol{\theta} | \boldsymbol{Y}_T\right) \approx \mathcal{N}\left(\hat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Sigma}}\right), \text{ where } \widehat{\boldsymbol{\Sigma}} = \widehat{\mathcal{I}}^{-1}/T \tag{2.4}$$

Note that a natural implication of the asymptotic normality of the posterior distribution is that the posterior mean and mode are asymptotically the same, and, as the sample size grows, both converge to the maximum likelihood estimator. Therefore, instead of MLE we could equivalently use the mean or the mode of the posterior distribution. Which one should be used in practice will depend on the point estimates one wishes to focus on.

2.2 Uncertainty reduction due to calibration

I use the asymptotic distribution to determine the impact of parameter calibration on the posterior uncertainty of the free parameters. For this, I assume that the calibrated values are not "wrong", in the sense of being different from the MLE (or posterior mean or mode) of the unrestricted model parameter values. Admittedly, this is a strong assumption, but I make it here in order to determine the pure effect calibration has on parameter uncertainty, i.e. in the absence of mis-calibration of the fixed parameters. I will consider the case of erroneous calibration later.

My approach consists of comparing two covariance matrices – that of the asymptotic posterior distribution when all elements of $\boldsymbol{\theta}$ are treated as free, and the one of the asymptotic posterior distribution of a subset of $\boldsymbol{\theta}$, conditional of the remaining parameters being fixed. For concreteness, let $\boldsymbol{\theta} = [\boldsymbol{\theta}_1', \boldsymbol{\theta}_2']'$ and partition $\boldsymbol{\Sigma}$ and $\mathcal{I}$ as follows:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\theta}_1} & \boldsymbol{\Sigma}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2} \\ \boldsymbol{\Sigma}_{\boldsymbol{\theta}_2\boldsymbol{\theta}_1} & \boldsymbol{\Sigma}_{\boldsymbol{\theta}_2} \end{bmatrix}, \qquad \mathcal{I} = \begin{bmatrix} \mathcal{I}_{\boldsymbol{\theta}_1} & \mathcal{I}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2} \\ \mathcal{I}_{\boldsymbol{\theta}_2\boldsymbol{\theta}_1} & \mathcal{I}_{\boldsymbol{\theta}_2} \end{bmatrix} \tag{2.5}$$

From (2.4), the asymptotic marginal posterior distribution of $\boldsymbol{\theta}_1$ is

$$\pi_T\left(\boldsymbol{\theta}_1|\boldsymbol{Y}_T\right) \approx \mathcal{N}\left(\hat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1}\right) \tag{2.6}$$

Now, suppose that $\boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2$ is known. The derivatives of the log-likelihood function with respect to $\boldsymbol{\theta}_2$ are zero, hence the Fisher's information matrix is given by $\widehat{\mathcal{I}}_{\boldsymbol{\theta}_1}$. Therefore, the asymptotic posterior distribution of $\boldsymbol{\theta}_1$ conditional on $\boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2$ is

$$\pi_T\left(\boldsymbol{\theta}_1|\boldsymbol{Y}_T, \hat{\boldsymbol{\theta}}_2\right) \approx \mathcal{N}\left(\hat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1|\boldsymbol{\theta}_2}\right), \text{ where } \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1|\boldsymbol{\theta}_2} = \widehat{\mathcal{I}}_{\boldsymbol{\theta}_1}^{-1}/T \tag{2.7}$$

An alternative expression for the covariance matrix in (2.7) is obtained by noting that $\pi_T\left(\boldsymbol{\theta}_1|\boldsymbol{Y}_T, \hat{\boldsymbol{\theta}}_2\right)$ is simply the conditional distribution of $\boldsymbol{\theta}_1$ given $\boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2$. From (2.6) we know that the joint distribution of these two vectors (given $\boldsymbol{Y}_T$) is asymptotically Gaussian. Therefore, when $\boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2$ is known, the variance of the conditional distribution of $\boldsymbol{\theta}_1$ is:

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1|\boldsymbol{\theta}_2} = \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1} - \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_2}^{-1}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_2\boldsymbol{\theta}_1} \tag{2.8}$$

Unless $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2} = \mathbf{0}$, i.e. $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are asymptotically independent, the marginal covariance matrix $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1}$ is larger than the conditional covariance matrix $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1|\boldsymbol{\theta}_2}$. In other words, knowing $\boldsymbol{\theta}_2$ reduces the uncertainty about the vector $\boldsymbol{\theta}_1$ as a whole. To quantify the effect of fixing $\boldsymbol{\theta}_2$ on the uncertainty about individual elements of $\boldsymbol{\theta}_1$, I define a measure of information gain (IG) with respect to a parameter θ_i as the percent reduction in the asymptotic standard deviation of that parameter, i.e.:

$$\mathrm{IG}_{\theta_i}(\boldsymbol{\theta}_2) = \left(\frac{\mathrm{std}_{\theta_i} - \mathrm{std}_{\theta_i|\boldsymbol{\theta}_2}}{\mathrm{std}_{\theta_i}}\right) \times 100, \tag{2.9}$$

where std_{θ_i} and $\mathrm{std}_{\theta_i|\boldsymbol{\theta}_2}$ are the square roots of the diagonal elements of $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1}$ and $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1|\boldsymbol{\theta}_2}$, respectively. Since $\mathrm{std}_{\theta_i} \geq \mathrm{std}_{\theta_i|\boldsymbol{\theta}_2} > 0$, the value of $\mathrm{IG}_{\theta_i}(\boldsymbol{\theta}_2)$ lies in the range between 0 and 100, with $\mathrm{IG}_{\theta_i}(\boldsymbol{\theta}_2) \approx 0$ implying that knowledge of $\boldsymbol{\theta}_2$ provides little or no information about θ_i, while $\mathrm{IG}_{\theta_i}(\boldsymbol{\theta}_2) \approx 100$ indicates that knowing $\boldsymbol{\theta}_2$ removes most of the uncertainty about θ_i.[4] It can be seen from (2.8) that the size of the information gain depends on how correlated θ_i and $\boldsymbol{\theta}_2$ are. In particular, the information gain will be small if the elements of $\widehat{\boldsymbol{\Sigma}}_{\theta_i\boldsymbol{\theta}_2}$ are close to zero, i.e. θ_i and the parameters in $\boldsymbol{\theta}_2$ are asymptotically close to being orthogonal. On the other hand, if one or more parameters in $\boldsymbol{\theta}_2$ are strongly correlated with θ_i, knowing $\boldsymbol{\theta}_2$ will provide a lot of information with respect to θ_i.

2.3 Sensitivity to errors in calibration

So far I have maintained the assumption that the calibrated parameter values are correct, i.e. they coincide with the values one would obtain if all model parameters were estimated freely. This, of course, is an unrealistic assumption and it is generally difficult to predict exactly how errors in the value of the fixed parameters will affect the ones that are estimated. Here I present a simple method for gauging the sign and the relative magnitude of the bias in a given estimated parameter as a result of errors in calibration. As before, I use the Gaussian approximation of the posterior distribution of $\boldsymbol{\theta}$. Suppose that the value of $\boldsymbol{\theta}_2$ is fixed at $\hat{\boldsymbol{\theta}}_2 + \triangle\hat{\boldsymbol{\theta}}_2$. The conditional mean of $\boldsymbol{\theta}_1$ given $\boldsymbol{\theta}_2$ is:

$$\mathrm{E}\left[\boldsymbol{\theta}_1|\boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2 + \triangle\hat{\boldsymbol{\theta}}_2\right] = \hat{\boldsymbol{\theta}}_1 + \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_2}^{-1}\triangle\hat{\boldsymbol{\theta}}_2 \tag{2.10}$$

Note that the first term on the right-hand side is the conditional mean of $\boldsymbol{\theta}_1$ given $\boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2$. Therefore, small deviations of $\boldsymbol{\theta}_2$ in the neighborhood of $\hat{\boldsymbol{\theta}}_2$ will shift the

[4] We can have information gain of 100% if a parameter θ_i is only identifiable when one or more other parameters are fixed, i.e. $\mathrm{std}_{\theta_i|\boldsymbol{\theta}_2} < \mathrm{std}_{\theta_i} = \infty$. In that case $\frac{\mathrm{std}_{\theta_i} - \mathrm{std}_{\theta_i|\boldsymbol{\theta}_2}}{\mathrm{std}_{\theta_i}} = \frac{\infty}{\infty}$ which I take to equal 1.

conditional mean of $\boldsymbol{\theta}_1$ by approximately $\mathbf{S}_{\boldsymbol{\theta}_1,\boldsymbol{\theta}_2}\triangle\hat{\boldsymbol{\theta}}_2$, where the sensitivity matrix $\mathbf{S}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2}$ is defined as

$$\mathbf{S}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2} = \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}_2}^{-1} = -\widehat{\mathcal{I}}_{\boldsymbol{\theta}_1}^{-1}\widehat{\mathcal{I}}_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2}, \tag{2.11}$$

where the second equality follows trivially from the properties of the inverse of partitioned matrices (see Exercise 5.16 in Magnus and Abadir (2005)). For an arbitrary pair of parameters $\theta_i \in \boldsymbol{\theta}_1$ and $\theta_j \in \boldsymbol{\theta}_2$, the corresponding element $\mathbf{S}_{\theta_i,\theta_j}$ of the sensitivity matrix shows the effect of perturbing the value of calibrated parameter θ_j on the asymptotic posterior mean value of free parameter θ_i.

The sensitivity measure in (2.11) is similar to the one proposed by Andrews et al. (2017) to measure the sensitivity of parameter estimates to reduced-form statistics. Instead of studying the effect of calibration, Andrews et al. (2017) are interested in the estimation bias one can expect as a result of violations in certain identifying assumptions. These violations are expressed as perturbations in the moment conditions on which a given estimation procedure, such as the generalized method of moments, is based. Similar to my approach, Andrews et al. (2017) derive their local sensitivity measure using the asymptotic Gaussian approximation of the joint distribution of structural parameters and moment conditions.

It is important to stress that the analysis described above is only applicable when the Fisher's information matrix is invertible (see equation (2.4)), i.e. when $\boldsymbol{\theta}$ is locally identifiable. At first glance this may appear to be a drawback of the methodology since calibration is often motivated by the difficulty to identify certain parameters. In particular, it may be that some parameters are not identifiable unless others are fixed. This means that the information matrix of $\boldsymbol{\theta}$ is singular and the information gain and sensitivity measures are not defined. Remember, however, that the purpose of the proposed methodology is to help researchers assess the impact of calibration on their estimation results. If some parameters have to be fixed in order to make others identifiable, the effect of calibration is clear – it makes the identification of some estimated parameters possible. The existing literature on identification in DSGE models has already discussed how to determine which parameters are not separately identifiable. Also, there are methods, such as the non-identification curves of Qu and Tkachenko (2012), that can reveal how changes in calibration values of fixed parameters affect the estimates of those free parameter that become identified as a result of calibration.[5] Thus, in principle, researchers already know how to investigate the impact of calibration when it is done in order to achieve identification. The contribution of the present paper is to show how to extend that analysis to situations where parameters are fixed even though they are identifiable.[6] It is, of course conceivable, that the set of calibrated parameters includes both parameters that can and parameters that cannot be identified. In that case, the vector $\boldsymbol{\theta}$ is assumed to include only the identifiable – calibrated and free – parameters.

2.4 A simple example

An illustration of the sensitivity and information gain measures for a two-parameter case is shown in Figure 1, where the joint distribution of $\boldsymbol{\theta} = [\theta_1, \theta_2]$ is Gaussian with both means equal to zero, variances equal to 1, and correlation coefficient equal to .9. Sensitivity in this case is equal to .9, which implies that a change of θ_2 from 0 to 1, i.e. a perturbation of one standard deviation, would shift the conditional mean of θ_1 by $.9 \times 1 = .9$. This represents an increase by .9 standard deviations. The conditional distribution of θ_1 is shown in the figure in green. In addition to the shift in the mean, we see also that the dispersion of the conditional distribution is smaller than that of the unconditional distribution. Using the measure of information gain introduced earlier, we have $\mathrm{IG}_{\theta_1}(\theta_2) = 100 \times \frac{(1-(1-.9^2))}{1} = 81\%$.

Some intuition for why in this example the value of θ_1 increases in response to a positive perturbation in the value of θ_2 can be gained by examining the local properties of the maximized likelihood function. Specifically, suppose that, instead of the mean of joint posterior distribution, the point $[0, 0]$ represents the unconstrained maximum of the log-likelihood function of $\boldsymbol{\theta}$. The inverse of the covariance matrix is the Fisher's information matrix, which has ones in the diagonal and $-.9$ in the off-diagonal positions. Since the information matrix is also the covariance matrix of the score vector, this implies that the correlation between the two elements of the score is $\mathrm{corr}(\partial\ell(\boldsymbol{\theta})/\partial\theta_1, \partial\ell(\boldsymbol{\theta})/\partial\theta_2) = -.9$. Therefore, the two parameters on average affect the log-likelihood function in the opposite direction and the effects are of similar magnitude. Since $\hat{\boldsymbol{\theta}} = [0, 0]$ is the mode of the log-likelihood, any perturbation in θ_2 away from 0 will lower the value of the log-likelihood

[5] The fact that this impact exists was first pointed out by Canova and Sala (2009).

[6] In practice, complete lack of identification is rarely used as a justification for calibration. It is far more common to see a weaker claim, such as parameters being "difficult" to identify or estimate.

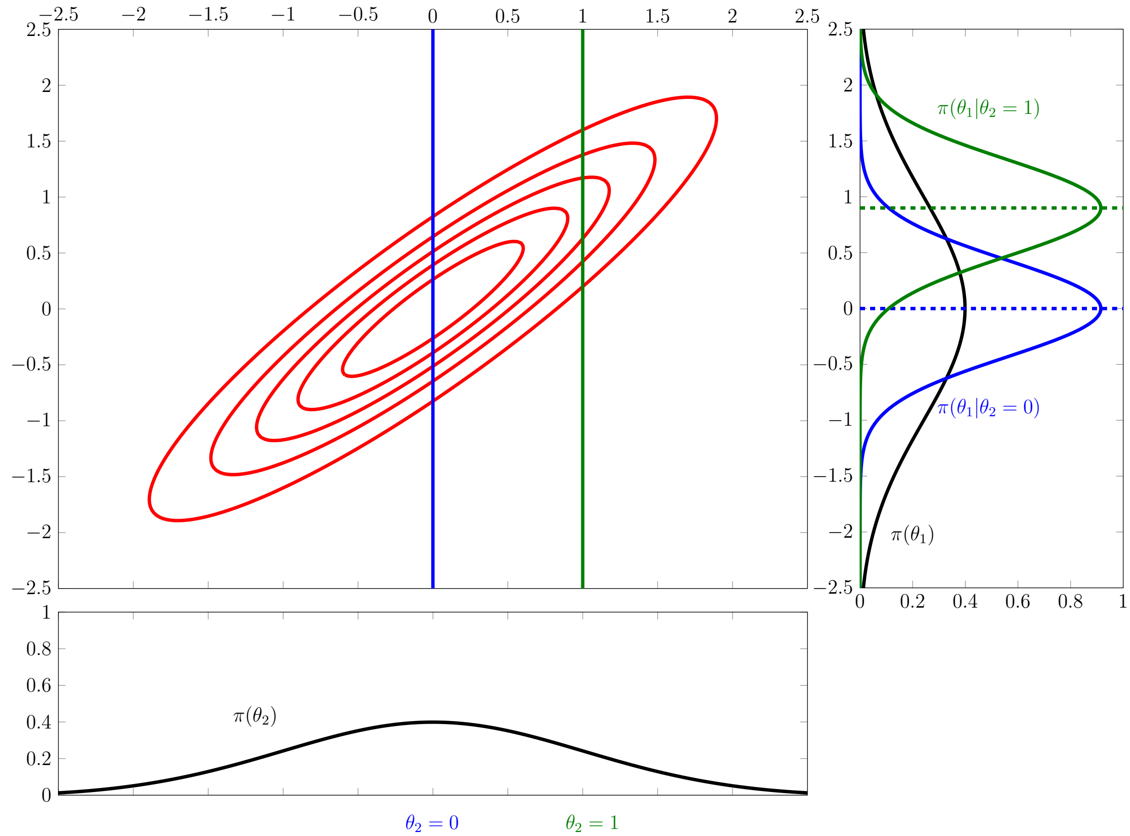

Figure 1: Two-parameter example. The figure shows how the conditional distribution of θ_1 depends on the value of θ_2.

distribution. To offset the effect of that change, θ_1 has to move in the same direction as θ_2. It is easy to show that, for small deviation $\triangle\theta_2$ in θ_2, the optimal change $\triangle\theta_1$ in θ_1 is given by:

$$\triangle\theta_1 = -\left(\frac{\partial^2\ell(\hat{\boldsymbol{\theta}})}{\partial\theta_1^2}\right)^{-1}\left(\frac{\partial^2\ell(\hat{\boldsymbol{\theta}})}{\partial\theta_1\partial\theta_2}\right)\triangle\theta_2 \tag{2.12}$$

This is the same expression as above except that in (2.10) the second derivatives of the log-likelihood function are replaced with their expected values. Hence, the sensitivity measure can be interpreted in terms of the required adjustment in the value of a free parameter in order to offset the effect of a perturbation in the value of a calibrated parameter. Note that the sign and size of that adjustment is determined by the correlation between the two elements of the score. If the correlation is positive, instead of negative as in the example above, the two parameters effect the log-likelihood function in the same direction, and therefore a positive perturbation in the value of the fixed parameter results in a negative change in the value of the estimated one. When the correlation is zero, there is no interaction between the two parameters, and fixing one of them has no effect on the marginal distribution of the other. Conversely, a correlation coefficient closer to one in absolute value results in both the sensitivity and information gain being larger than in the example above.[7]

This intuition extends to multi-parameter models: starting from the mode of the log-likelihood function, perturbation of one or more parameters away from their unrestricted optimal values can be partially offset by adjusting the remaining free parameters away from their unrestricted optimal values.[8] Since there are potentially many parameters that could be adjusted, the optimal size of the adjustment of each one depends on the full correlation structure, not just the pairwise correlations between free and calibrated parameters.

The preceding discussion also helps clarify the role of the asymptotic approximation in my analysis. As already noted earlier, the effect of calibration derives from the interdependence between free and calibrated parameters encoded in their joint posterior distribution. Using the asymptotic normal approximation to that distribution has two potentially important implications. First, it is a well-known fact that the correlation matrix fully characterized the dependence structure only in the case of elliptical distributions, such as the multivariate normal distribution. In the general case, correlation only captures the degree of linear dependence among variables. Therefore, while the sensitivity measure should deliver accurate predictions about the size of the effect of small perturbations in the calibration values, it might not do so for large perturbations. Second, using the asymptotic variance as a measure of estimation uncertainty might misrepresent the amount of information gained by calibration. Thus, the measures in this paper should not be relied upon to deliver numerically precise assessment of the exact impact of calibration on the estimation results. Instead, their main purpose is

[7]Remember that $\mathrm{corr}(\partial \ell(\boldsymbol{\theta})/\partial\theta_1, \partial \ell(\boldsymbol{\theta})/\partial\theta_2) = \pm 1$ means that the likelihood function is flat and θ_1 and θ_2 are not separately identifiable. Therefore, we can think of the case when the correlation is close to but smaller than one in absolute value as θ_1 and θ_2 being weakly identified. In that sense, fixing weakly identified parameters is associated with large values of the sensitivity and information gain measures for free parameters that would be weakly identified without calibration.

[8]The offset can only be partial unless the log-likelihood function is flat at the mode, i.e. the model is locally unidentified.

to identify the parameters whose estimates are more likely to have been affected by calibration, and to shed light on the nature of that effect. In particular, the measures should make accurate predictions about (1) which estimated parameters are more and which ones are less sensitive to changes in the value of a given fixed parameter, (2) the direction of the sensitivity, i.e. positive or negative, and (3) which parameters benefit more and which – less, in terms of estimation uncertainty, from the calibration of a given parameter. Some evidence in support of this claim in the context of a small scale DSGE model is presented in the Appendix, where the measures' predictions are compared to simulation-based results. In the next section the measures are used to assess the effect of calibration on the estimation results in two medium-scale models taken from the literature.

3 Applications

I now apply the proposed measures to investigate the effect of calibration in two estimated models: the medium-scale New Keynesian model of Smets and Wouters (2007), and the real business cycle model with news shocks of Schmitt-Grohé and Uribe (2012). In each case I take as given the division of the model parameters into freely-estimated and calibrated ones as well as the estimation results reported in those articles.

3.1 Smets and Wouters (2007)

The Smets and Wouters (2007) (hereafter SW) model is a medium-scale closed-economy New Keynesian model featuring price and wage rigidities, habit formation, capital accumulation, investment adjustment cost, variable capital utilization. The model is estimated with Bayesian methods using US data on output growth, consumption growth, investment growth, real wage growth, hours worked, inflation and the nominal interest rate. There are 41 parameters in the model 36 of which are estimated and the other 5 are calibrated. The calibrated parameters are: depreciation rate (δ), steady state wage mark-up (λ_w), exogenous spending-output ratio (g_y), and the curvature parameters of goods and labor market aggregators (ε_p and ε_w). The reasons SW give for calibrating these parameters are that δ and g_y are difficult to estimate with the observed series, while λ_w, ε_p and ε_w are not identified. As has been shown previously (see Iskrev (2010)), λ_w is in fact identified, while two pairs of parameters – (ξ_p, ε_p) and (ξ_w, ε_w) are not

Table 1: Calibrated parameters, SW (2007) model

	parameter	value
δ	depreciation rate	0.025
λ_w	steady state wage markup	1.50
g_y	exogenous spending-output ratio	0.18
ε_p	curvature of goods market aggregator	10.00
ε_w	curvature of labor market aggregator	10.00

separately identifiable. That is, in the linearized model ξ_p cannot be distinguished from ε_p and ξ_w cannot be distinguished from ε_w. This implies that the covariance matrix of the asymptotic posterior distribution of the full set of parameters is singular and the measures of information gains and sensitivity are not defined. Therefore, here I will study the effect of fixing 3 of the 5 parameters, namely δ, λ_w, and g_y, on the distribution of the 36 parameters which SW estimate, conditional on the curvature parameters of goods and labor market aggregators (ε_p and ε_w) being both fixed at 10, as in the original article.[9] I consider the same values for the calibrated parameters as in SW, shown in Table 1, while for the estimated parameters I take the posterior mean reported in the article – see Table 2. I use these values to compute the measures of sensitivity to and information gains from calibration.

The information gains due to calibration of δ, λ_w, and g_y are reported in panel (a) of Figure 2. The gains are zero or close to zero for 11 of the free parameters, and exceed 10% for 8 parameters. The largest information gains are with respect to the wage stickiness parameter ξ_w – almost 60%, and with respect to the elasticity of labor supply σ_c – about 40%. There are also significant gains of about 20% with respect to the discount factor $\bar{\beta}$ and the investment adjustment cost parameter φ.

To better understand how individual calibrated parameters contribute to the total information gains, in panels (b), (c), and (d) of the same figure I report the size of the gains from fixing only one of the three parameters at a time, either δ, λ_w, or g_y, respectively, while keeping the other two parameters free. This exercise shows that most of the larger gains – those with respect to ξ_w, σ_c, φ, and $\bar{\beta}$, are due to information obtained from knowing the value of λ_w alone. Knowing the value of δ provides significant

[9]Since lack of identification implies infinite variance of the asymptotic marginal posterior distribution, in the case of ξ_p and ξ_w we have information gains of 100% due to fixing ε_p and ε_w, respectively.

Table 2: Estimated parameters, SW (2007) model

	parameter	value
ρ_{ga}	productivity shock in government spending	0.52
$\bar{l}$	steady state hours	0.54
$\bar{\pi}$	steady state inflation	0.79
$\bar{\beta}$	normalized discount factor $^{(a)}$	0.17
μ_w	MA wage markup	0.84
μ_p	MA price markup	0.70
α	capital share	0.19
ψ	capacity utilization cost	0.55
φ	investment adjustment cost	5.74
σ_c	elasticity of intertemporal substitution	1.38
λ	habit	0.71
Φ	fixed cost in production	1.60
ι_w	wage indexation	0.59
ξ_w	wage stickiness	0.70
ι_p	price indexation	0.24
ξ_p	price stickiness	0.65
σ_l	elasticity of labor supply	1.84
r_π	monetary policy response to inflation	2.05
$r_{\triangle y}$	monetary policy response to change in output gap	0.22
r_y	monetary policy response to output gap	0.09
ρ	interest rate smoothing	0.81
ρ_a	AR productivity shock	0.96
ρ_b	AR risk premium shock	0.22
ρ_g	AR government spending shock	0.98
ρ_I	AR investment specific shock	0.71
ρ_r	AR monetary policy shock	0.15
ρ_p	AR price markup shock	0.89
ρ_w	AR wage markup shock	0.97
γ	trend growth rate	0.43
σ_a	standard deviation productivity shock	0.46
σ_b	standard deviation risk premium shock	0.24
σ_g	standard deviation government spending shock	0.53
σ_I	standard deviation investment specific shock	0.45
σ_r	standard deviation monetary policy shock	0.25
σ_p	standard deviation price markup shock	0.14
σ_w	standard deviation wage markup shock	0.24

Note: The values are of the mean of the posterior distribution of the Smets and Wouters (2007) model.
(a) $\bar{\beta} = 100(\beta^{-1} - 1)$ where β is the usual discount factor.

amount of information with respect to α, ψ, and ρ_a. The least informative of the three calibrated parameters is g_y, which nonetheless contributes a substantial amount of information with respect to Φ, σ_g and ψ.

Turning to the sensitivity of the parameter estimates to changes in the calibration values, Figure 3 plots the values of the sensitivity measure. To make the values comparable, I scale sensitivity by the standard deviations of the parameters so that the displayed values show the change, in terms of standard deviations of the respective parameter, to a one standard deviation increase in the value of each calibrated parameter.[10] The results closely mirror those in Figure 2. The largest impact is on the estimate of ξ_w, which drops by 0.9 standard deviations as a result of one standard deviation increase in λ_w. Perturbing the value of λ_w also has a significant impact on the values of σ_c, φ, and $\bar{\beta}$, raising by more than .6 standard deviations the values of the first two parameters and reducing by almost .6 standard deviations the value of $\bar{\beta}$. As before, the strongest impact from a change in δ is on α, ψ, and ρ_a, all of which decrease by about 0.5 standard deviations as a result of a one standard deviation increase in δ. In the case of g_y, the impact is again most pronounced with respect to Φ, ψ, and σ_g, whose values decline by between .3 and .4 standard deviations due to a one standard deviation increase in g_y.

Note that unlike the computation of the information gains with respect to a single parameter in panels (b), (c) and (d) of Figure 2, the sensitivity measures in Figure 3 are computed assuming that all calibrated parameters remain fixed, and only one of them is perturbed at a time. In particular, when one of the calibrated parameters is perturbed only the free parameters are allowed to respond, while the other two calibrated parameters remain fixed. This was not the case in Figure 2. The distinction may be important, particularly when there is a strong interdependence among the calibrated parameters. For instance, if λ_w and g_w are free to adjust when δ is perturbed, there may be a much smaller response of the other free parameters since some of the effect of changing δ could be offset by changes in λ_w and g_w. On the other hand, if the calibrated parameters are close to independent, changing one of them would lead to a small or no change in the other two, even if those were allowed to adjust. In Figure A1 of the Appendix I show the sensitivities when only one of the three calibrated parameters is fixed at a time. The results are very similar to those in Figure 3, implying that there is only weak interdependence among λ_w, δ and g_w.

[10]Both standard deviations are computed using the asymptotic covariance matrix of the unrestricted model.

In the Appendix I also report pairwise conditional information gains and pairwise conditional sensitivity values, where for each pair of parameters the conditioning is on *all* remaining 37 parameters. The pairwise conditional gains (see Figure A2) show how much information about a given parameter θ_i is gained if another parameter θ_j is fixed, conditional on knowing all parameters except these two. There are some marked differences, especially between the conditional and unconditional gains from fixing λ_w (compare panel (c) in Figure 2 with panel (b) in Figure A2). Note that the gains with respect to ξ_w are very large both conditionally and unconditionally. However, the conditional information gains with respect to μ_w, σ_l, ρ_w, and σ_w are much larger than the unconditional gains for those parameters. In contrast, the unconditional gains with respect to $\bar{\beta}$ and σ_c are significantly larger than the conditional ones.

These findings underscore the fact that in a multiparameter setting the effect of calibration cannot be easily discerned using simple bivariate relationships between individual calibrated and free parameters. Intuitively, one might expect that the effect will be greater for parameters which, in the model, are functionally closely related to some calibrated parameters. As the example in Section 2.4 reveals, in a bivariate setting strong correlation between the scores $\partial\ell(\boldsymbol{\theta})/\partial\theta_i$ and $\partial\ell(\boldsymbol{\theta})/\partial\theta_j$, which reflects similar functional roles of θ_i and θ_j, would cause fixing one of the two parameters to have a large impact on the conditional distribution of the other. With more than two parameters, the negative of $\mathrm{corr}(\partial\ell(\boldsymbol{\theta})/\partial\theta_i, \partial\ell(\boldsymbol{\theta})/\partial\theta_j)$ represents the conditional correlation between θ_i and θ_j, given the remaining model parameters.[11] Differences between the conditional and the marginal correlation structures can lead to very different conditional and unconditional information gains, as in the case of the gains due to fixing λ_w. Consider Figure 4 where I show two sets of parameters that are strongly related to λ_w. In particular, panel (a) displays a conditional correlation network of all parameters connected with λ_w, while panel (b) shows a marginal correlation network of the parameters connected with λ_w. In both cases, to make the graphs more readable, I show only links between parameters whose correlation is greater or equal to .4 in absolute value. The full set of marginal and conditional correlations can be found in Figure 5. It can be seen that μ_w, σ_l, ρ_w, and σ_w are strongly conditionally correlated with both ξ_w and λ_w, as well as among each other. This explains the large pairwise conditional information gains in panel (b) of Figure A2, where the gains from fixing λ_w are conditional on all other parameters, and in particular

[11] This follows from the fact that the covariance matrix of the scores is the precision matrix of the asymptotic posterior distribution, and thus it encodes the conditional correlations between pairs of parameters given the remaining parameters (see Cramér (1946)).

ξ_w, also being fixed. At the same time, the marginal correlations between λ_w and those four parameters are too week to show in the graph in panel (b). This is mainly due to the fact that, because of their functional similarity in the model, λ_w and ξ_w are very strongly correlated both conditionally and unconditionally. As a result, fixing λ_w while keeping ξ_w free provides very little information with respect to μ_w, σ_l, ρ_w, and σ_w. On the other hand, the marginal correlations of λ_w with σ_c and $\bar{\beta}$ are strong, in spite of the very weak conditional correlations. This implies that these two parameters benefit from fixing λ_w only indirectly – through other free parameters which are more closely linked to λ_w and whose uncertainty is impacted directly as a result of fixing that parameter. In the conditional case those parameters are already known and thus fixing λ_w contributes little (in the case of σ_c) or no (in the case of $\bar{\beta}$) additional information.

The differences between conditional and unconditional sensitivities can be explained in a similar fashion. As can be seen by comparing Figures A1 and A3 of the Appendix, the conditional sensitivities tend to be significantly larger than the unconditional ones. This is because in the conditional case only one parameter at a time is free to adjust so as to optimally offset the effect of changing the value of a given calibrated parameter. In the case of the unconditional sensitivities, all free parameters are allowed to move and thus the magnitudes of the optimal adjustments tend to be smaller.

Figure 2: Information gains from calibration. Panel (a) shows the gains from knowing the values of all calibrated parameters. Panels (b), (c), and (d) show the gains from knowing only one parameter at a time.

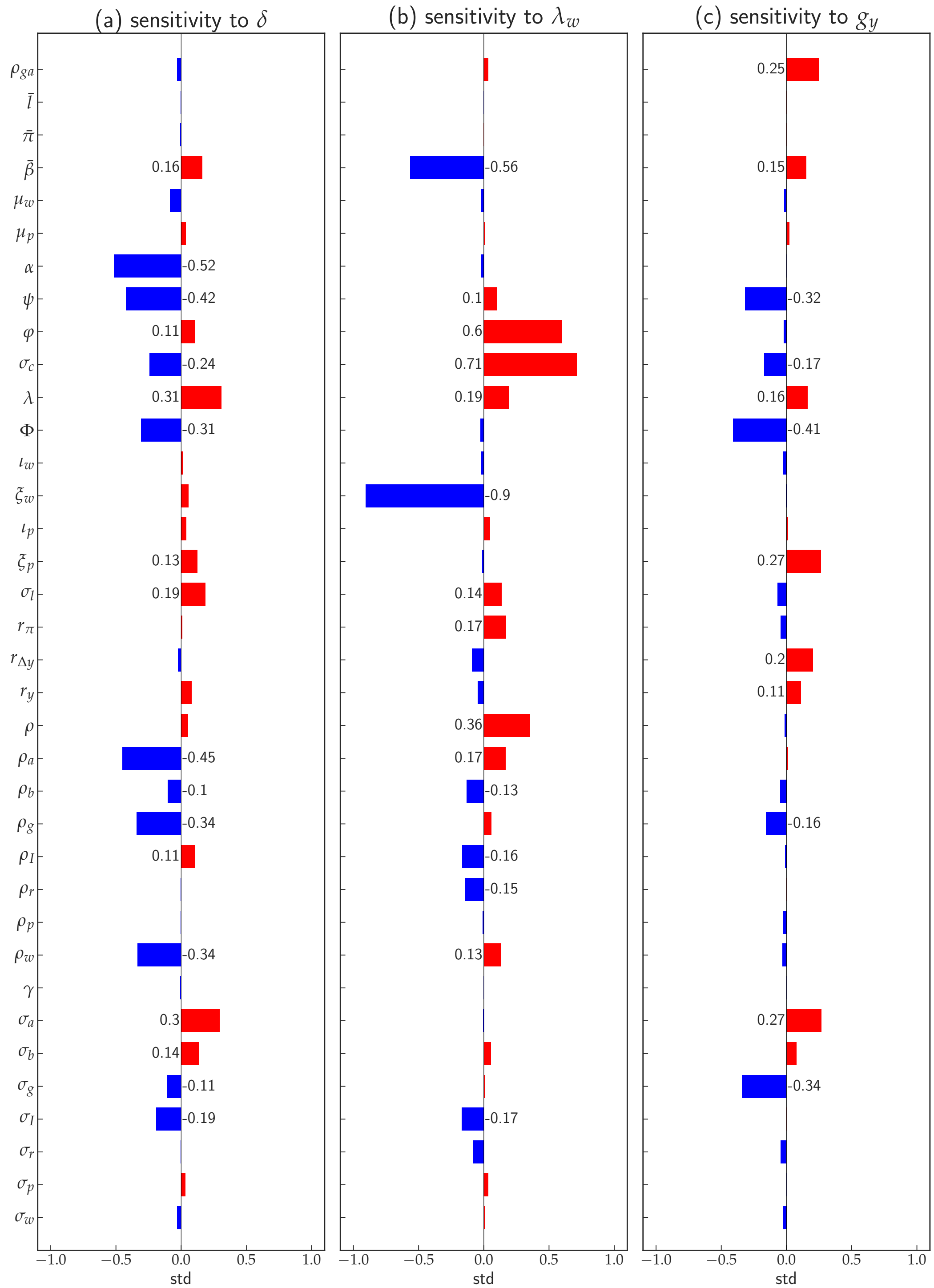

Figure 3: Sensitivity to changes in the calibrated parameters. Each panel shows the effect of a one-standard-deviation increase in the respective parameter on the value of each free parameter, in units of standard deviations.

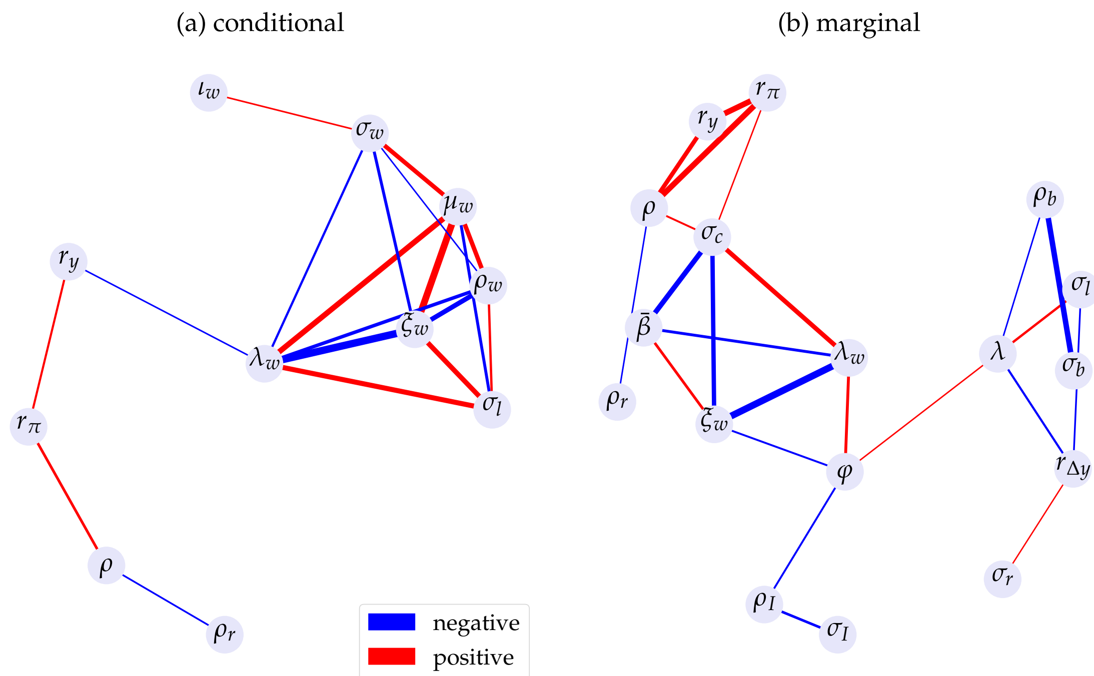

Figure 4: Conditional and marginal correlation networks of parameters connected with λ_w. Both graphs show only edges between parameters whose conditional (panel (a)) or marginal (panel (b)) correlations are greater than or equal to .4 in absolute value. The lines thickness is proportional to the strength of correlation, and the color depends on its sign.

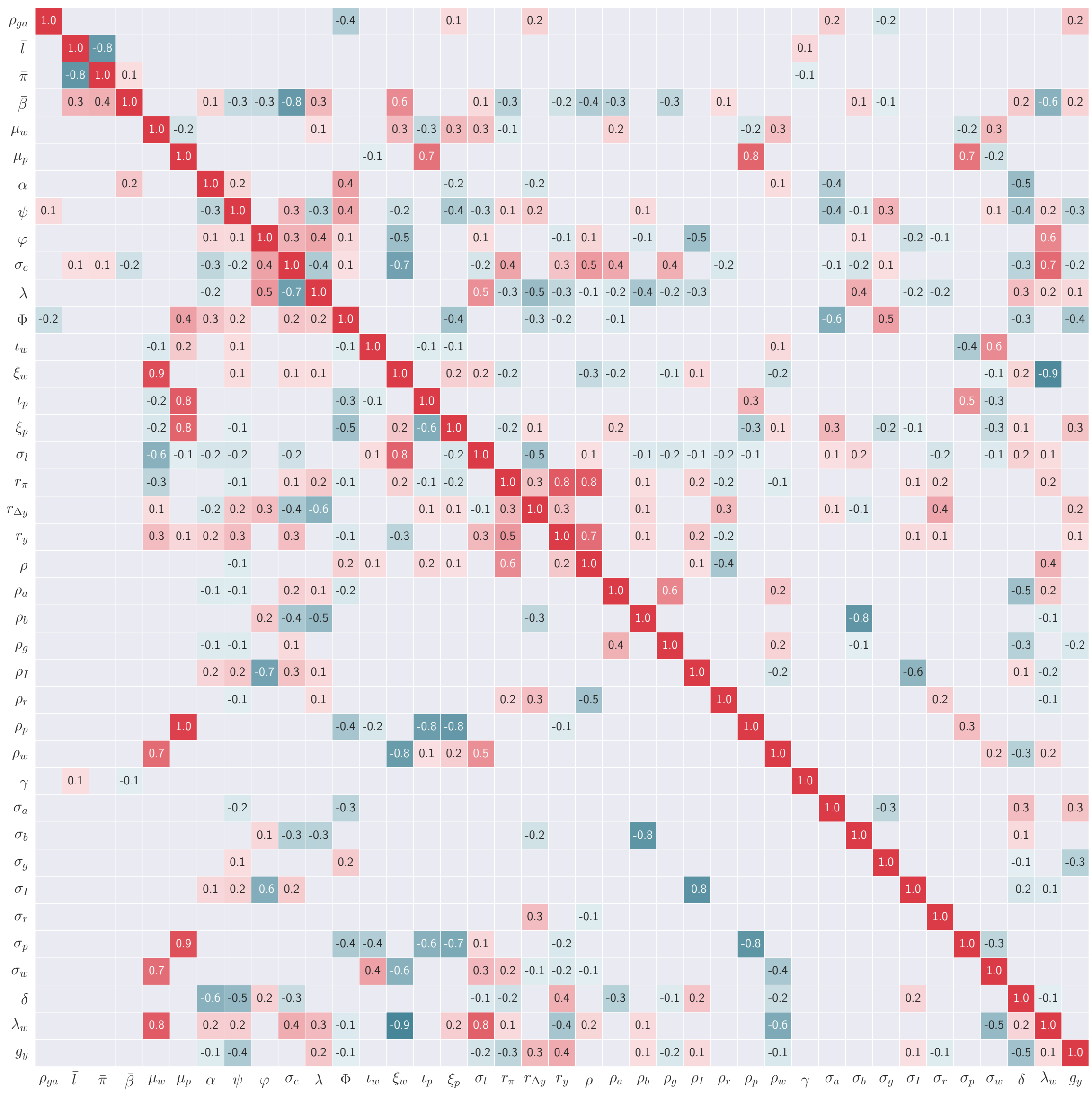

Figure 5: Parameter correlations in the SW model. The lower triangle of the matrix shows the conditional correlation coefficients between each pair of parameters. The upper triangle shows the marginal correlation coefficients. The values are obtained from the joint asymptotic posterior distribution of the parameters evaluated at the posterior mean in SW. Correlation coefficients smaller than .1 in absolute value are not displayed.

Table 3: Calibrated parameters, SGU (2012) model

	parameter	value
α_k	Capital share	0.225
α_h	Labor share	0.675
δ_0	Steady-state depreciation rate	0.025
β	Subjective discount factor	0.99
h_{ss}	Steady-state hours	0.2
μ	Steady-state wage markup	1.15
μ^a	Steady-state gross growth rate of price of investment	0.9957
μ^y	Steady-state gross per capita GDP growth rate	1.0045
σ	Intertemporal elasticity of substitution	1
g_y	Steady-state share of government consumption in GDP	0.2

3.2 Schmitt-Grohé and Uribe (2012)

The Schmitt-Grohé and Uribe (2012) (hereafter SGU) model is a medium-scale closed-economy real business cycle model augmented with real rigidities in consumption, investment, capital utilization, and wage setting. The model has seven fundamental shocks: to neutral productivity (stationary and non-stationary), to investment-specific productivity (stationary and non-stationary), government spending, wage markups and preferences. Each one of the seven shocks is driven by three independent innovations, two anticipated and one unanticipated. More concretely, the process governing shock x_t is given by

$$\ln(x_t/x) = \rho_x \ln(x_{t-1}/x) + \sigma_x^0 \varepsilon_{x,t}^0 + \sigma_x^4 \varepsilon_{x,t-4}^4 + \sigma_x^8 \varepsilon_{x,t-8}^8, \tag{3.1}$$

where $\varepsilon_{x,t}^j$ for $j = 0, 4, 8$ are independent standard normal random variables. The anticipated innovations $\varepsilon_{x,t-4}^4$ and $\varepsilon_{x,t-8}^8$ are known to agents in periods $t-4$ and $t-8$, respectively. Thus, they can be interpreted as news shocks.

The model has 45 parameters, the following 10 of which are calibrated: capital and labor shares (α_k and α_h), steady-state depreciation rate (δ_0), subjective discount factor (β), steady-state hours (h_{ss}), steady-state wage markup (μ), steady-state growth rate of price of investment (μ^a), steady-state gross per capita GDP growth rate (μ^y), intertemporal elasticity of substitution (σ), and steady-state share of government consumption in GDP (g_y). The values of these parameters are listed in Table 3.

Table 4: Estimated parameters, SGU (2012) model

	parameter	value
θ	Frisch elasticity of labor supply	5.39
γ	wealth elasticity of labor supply	0.00
κ	investment adjustment cost	25.07
δ_2/δ_1	capacity utilization cost	0.44
b	habit in consumption	0.94
ρ_{xg}	smoothness of trend in government spending	0.74
ρ_z	AR stationary neutral productivity	0.96
ρ_{μ^a}	AR nonstationary investment-specific productivity	0.48
ρ_g	AR governement spending	0.96
ρ_{μ^x}	AR nonstationary neutral productivity	0.77
ρ_μ	AR wage markup	0.98
ρ_ζ	AR preference	0.10
ρ_{z^I}	AR stationary investment-specific productivity	0.21
$\sigma^0_{\mu^a}$	std. dev. nonstationary investment-specific productivity 0	0.16
$\sigma^4_{\mu^a}$	std. dev. nonstationary investment-specific productivity 4	0.20
$\sigma^8_{\mu^a}$	std. dev. nonstationary investment-specific productivity 8	0.19
$\sigma^0_{\mu^x}$	std. dev. nonstationary neutral productivity 0	0.45
$\sigma^4_{\mu^x}$	std. dev. nonstationary neutral productivity 4	0.12
$\sigma^8_{\mu^x}$	std. dev. nonstationary neutral productivity 8	0.12
$\sigma^0_{z^I}$	std. dev. stationary investment-specific productivity 0	34.81
$\sigma^4_{z^I}$	std. dev. stationary investment-specific productivity 4	11.99
$\sigma^8_{z^I}$	std. dev. stationary investment-specific productivity 8	14.91
σ^0_z	std. dev. stationary neutral productivity 0	0.62
σ^4_z	std. dev. stationary neutral productivity 4	0.11
σ^8_z	std. dev. stationary neutral productivity 8	0.11
σ^0_μ	std. dev. wage markup 0	1.51
σ^4_μ	std. dev. wage markup 4	3.93
σ^8_μ	std. dev. wage markup 8	3.20
σ^0_g	std. dev. government spending 0	0.53
σ^4_g	std. dev. governement spending 4	0.69
σ^8_g	std. dev. governement spending 8	0.43
σ^0_ζ	std. dev. preference 0	2.83
σ^4_ζ	std. dev. preference 4	2.76
σ^8_ζ	std. dev. preference 8	5.34
$\sigma^{me}_{g^y}$	std. dev. measurement error in output	0.30

Note: Maximum likelihood estimates of Schmitt-Grohé and Uribe (2012)

The remaining 35 parameters are estimated using both Bayesian methods and by maximum likelihood with US data on the growth rates of output, consumption, investment, government expenditure, the relative price of investment, total factor productivity, and hours worked. The analysis here uses the maximum likelihood estimates reported in SGU and reproduced in Table 4. Alternative results based on the median of the posterior distribution are available upon request.

Checking the rank condition for identification shows that the steady-state hours parameter h_{ss}, which SGU calibrate, is not identified. Therefore, in my analysis I consider only the remaining nine calibrated parameters. In addition, unlike SGU who use de-meaned data, I assume that information from both the mean and the covariance structure of the seven observed variables is used. This is important since most of the calibrated parameters are related to the steady state of the model and thus information from the mean is important for their identification.

Figure 6 presents the information gains due to calibration. As in Section 3.1, I report the gains from fixing all nine parameters (panel (a)), and the individual information gains from fixing only one parameter at a time (panels (b) to (f)). I do not report individual information gains from the calibration of α_k, μ^a, μ^y and g_y since they are less than 1% for all parameters. The total information gains are less than 1% for 3 of the free parameters, and exceed 10% in the case of 7 parameters. The largest gains are about 50% – with respect to the consumption habit parameter b, and between 35% and 42% for the parameters of the investment adjustment cost (κ), capacity utilization cost (δ_2/δ_1), and the unanticipated innovations to the stationary investment-specific productivity shock ($\sigma^0_{z_I}$). There are also relatively large information gains of around 15% with respect to the Frisch elasticity of labor supply parameter (θ), and the volatility parameters of two of the innovations to the wage markup shock (σ^0_μ and σ^8_μ). Panels (b) to (f) of the same figure help identify the main sources of the overall information gains. The bulk of information with respect to b comes from fixing the value of σ, while δ_0 is the most informative calibrated parameter with respect to κ, δ_2/δ_1, $\sigma^0_{z_I}$ and θ. Fixing the value of μ contributes the most to reduce the uncertainty about σ^0_μ and σ^8_μ, although δ_0 is the most informative parameter to calibrate with respect to σ^4_μ. The calibration of β helps reduce the uncertainty of κ, $\sigma^0_{z_I}$, δ_2/δ_1, and b, while that of α_h is only marginally informative with respect to a few parameters, most notably b.

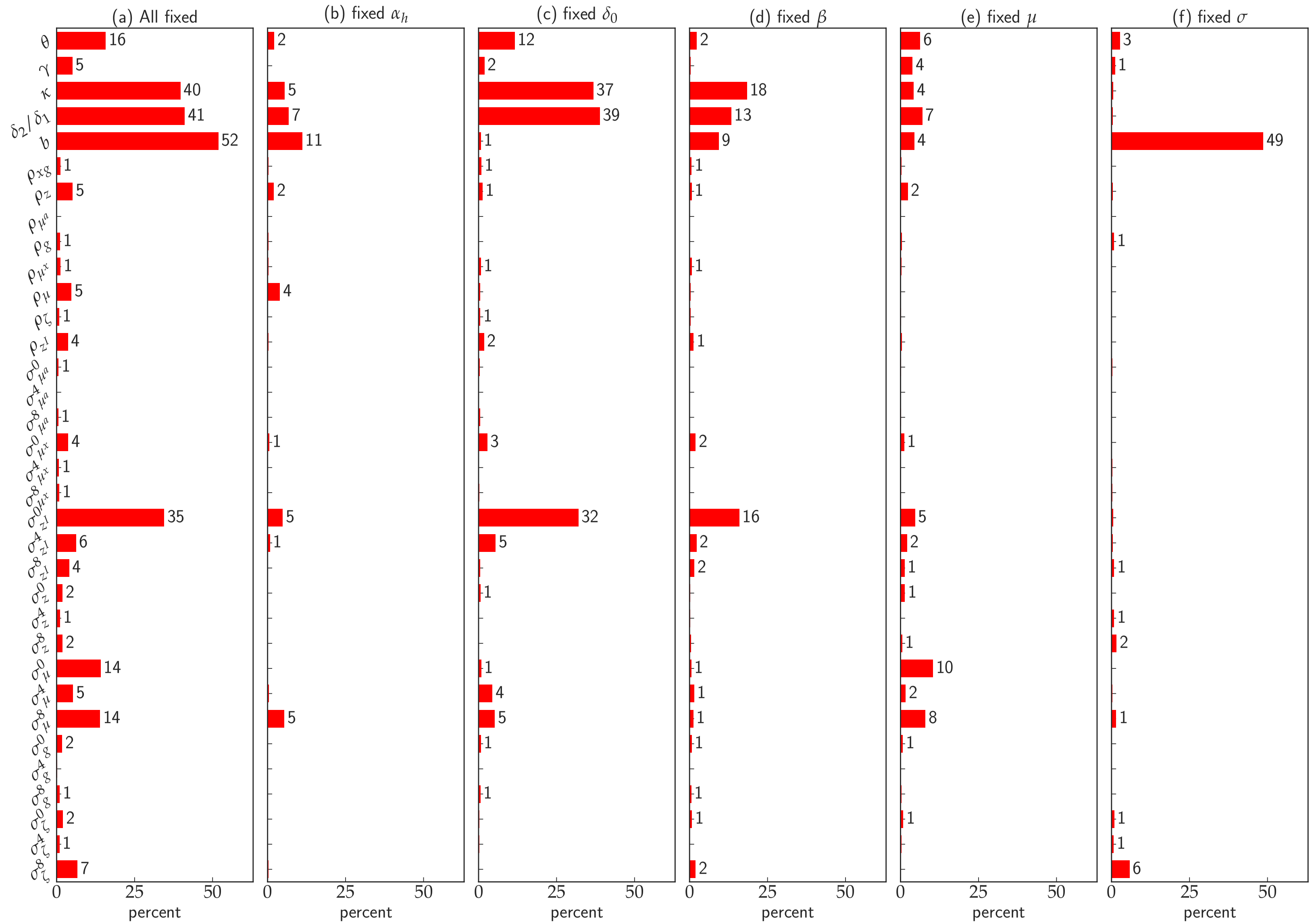

Figure 6: Information gains from calibration. Panel (a) shows the gains from knowing the values of all calibrated parameters $(\alpha_k, \alpha_h, \delta_0, \beta, \mu, \mu^a, \mu^y, \sigma, g_y)$. Panels (b), (c), and (d) show the gains from knowing only one parameter at a time. All gains from fixing α_k, μ^a, μ^y, or g_y are less than 1% and therefore are not displayed.

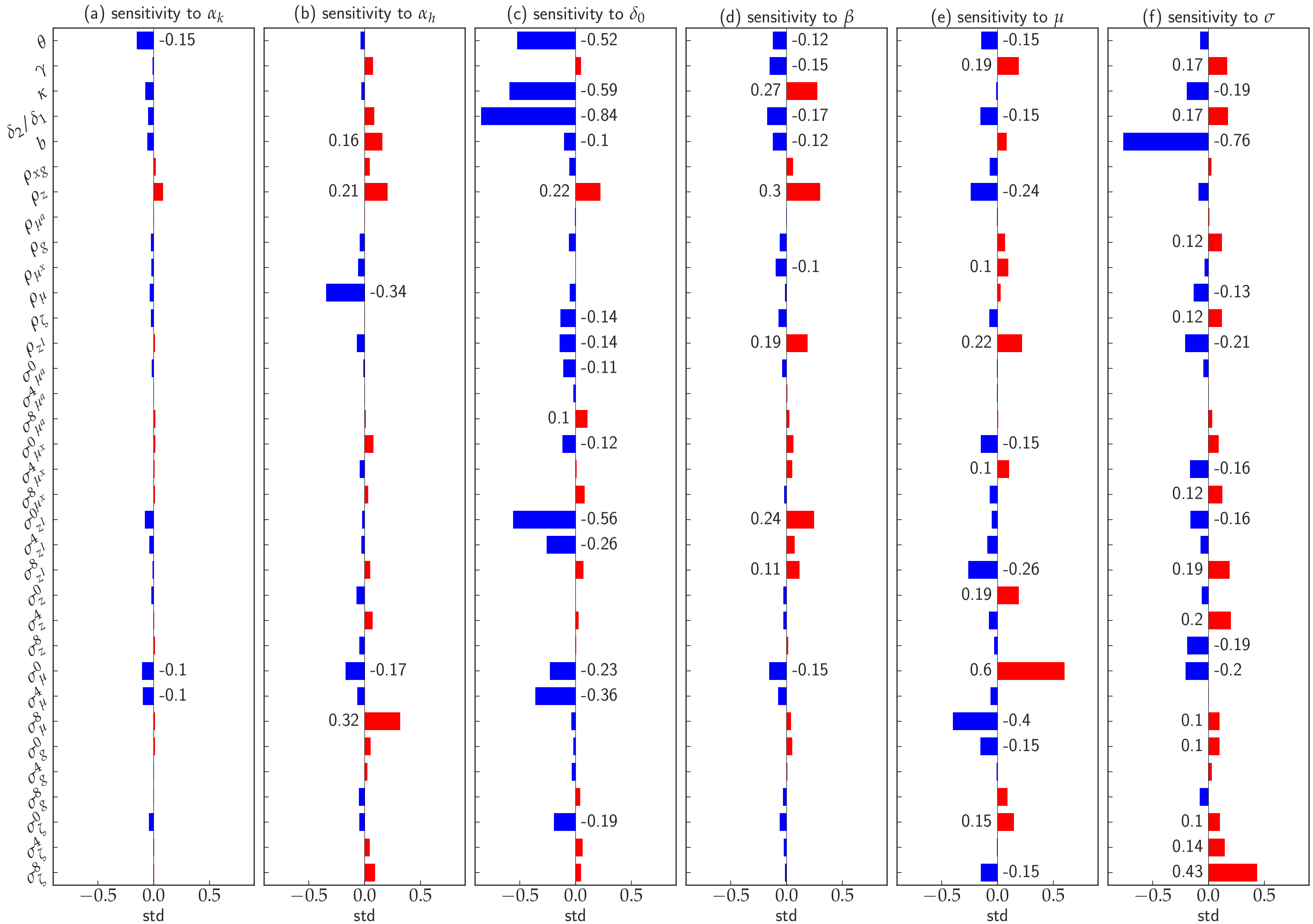

Figure 7: Sensitivity to changes in the calibrated parameters. Each panel shows the effect of a one-standard-deviation increase in the respective parameter on the value of each free parameter, in units of standard deviations.

Sensitivity to calibration results are presented in Figure 7. As before, I scale the sensitivity measure so that the values represent the change, in terms of standard deviations of each free parameter, as a result of a one standard deviation increase in the value of a given calibrated parameter. Again, I do not show sensitivity results with respect to μ^a, μ^y and g_y as they are always smaller than 0.1 in absolute value. Similar to the results in Figure 6, the largest sensitivities are with respect to δ_0. In particular, θ, κ, δ_2/δ_1, and $\sigma^0_{z_I}$ all decrease by more than 0.5 standard deviations as a result of one standard deviation increase in δ_0. In the case of δ_2/δ_1 the sensitivity is more than 0.8 in absolute value. Two additional parameters – b and σ^0_μ, show sensitivity greater than 0.5 in absolute value – with respect to σ and μ, respectively.

Similarly to the SW model, drawing conclusions about the effect of calibration on the basis of simple pairwise relationships between fixed and free parameters can be misleading. For instance, conditionally on the remaining parameters, fixing the intertemporal elasticity of substitution (σ) can have a very large impact on the wealth elasticity of labor supply (γ) and the preference shock parameters (σ^0_ζ, σ^4_ζ, σ^8_ζ). This can be seen in Figures A5 and A6 of the Appendix, which show conditional information gain and sensitivity values. On the other hand, from Figures 6 and 7 we see that, when the other parameters are free to adjust, the effect of fixing σ is very small for γ, σ^0_ζ, σ^4_ζ, and σ^8_ζ . As before, this and other similar observations can be explained with the large, in some cases, differences between conditional and marginal correlation patterns of the parameters' impact on the log-likelihood function. This can be observed in Figure 8.

The main conclusion to draw from the findings in this section is that the estimation results in both the SW and SGU models are influenced by calibration. Not all free parameters are affected and in many cases the consequences are found to be minor. However, for some parameters the consequences are quite large. This does not necessarily mean that the results in the these papers are invalid. Assuming that the calibration values are well-justified, the point estimates obtained as a result are consistent with those choices. At the same time, as I have shown, the estimation uncertainty is still likely to be mis-represented since it is unrealistic to accept that the fixed parameters are indeed known with certainty.[12] To the extent that model predictions depend on such parameters, it is important to be aware of the possible underestimation of that uncertainty. More generally, researchers clearly do not always agree on whether and how

[12]Note that even in the case of the least controversial calibration values – those of parameters that are directly related to observed long-run ratios – one has to account for estimation uncertainty associated with those sample moments.

to calibrate. Providing readers with information about the potential consequences of calibration should help increase the transparency and improve the credibility of estimated structural models.

4 Conclusion

Estimation of structural macroeconomic models often assumes the complete knowledge of some of their parameters. Whether or not this is a reasonable assumption to make is perhaps an open question. However, it is important to bear in mind that, even when it is well justified, calibration can have a substantial impact on the estimation results stemming from parameter interdependence, which is common feature of macroeconomic models. It is therefore appropriate that researchers who estimate such models by combining calibration and estimation, discuss not only the reasons for and methods of calibration, but also the impact it may have on their results.

In this paper I propose two new measures that can be used to shed light on the consequences of calibration. The first one shows how much information is introduced with respect to each freely estimated parameter as a result of calibration of one or more model parameters. The second measures the sensitivity of different parameter estimates to perturbations in the values of the calibrated parameters. By design, our measures capture the main ways in which calibration could influence estimation – by changing the location and reducing the spread of the marginal posterior distributions of the estimated parameters. Providing readers with information about these effects is important in recognition of the fact that there may be disagreements among researchers both in terms of whether certain parameters can reasonably be assumed to be known, and regarding what their values should be.

The main advantage of the proposed measures is that they are easy to interpret and simple to compute without requiring additional estimation effort. This makes them straightforward to incorporate into the standard estimation output reported in empirical DSGE studies. At the same time, they also have the limitation of being local and hence valid only in the neighborhood of the original calibration values and parameter estimates. Needless to say, the measures are not appropriate to use as a substitute for a full-scale re-estimation of a model under alternative calibration assumptions.

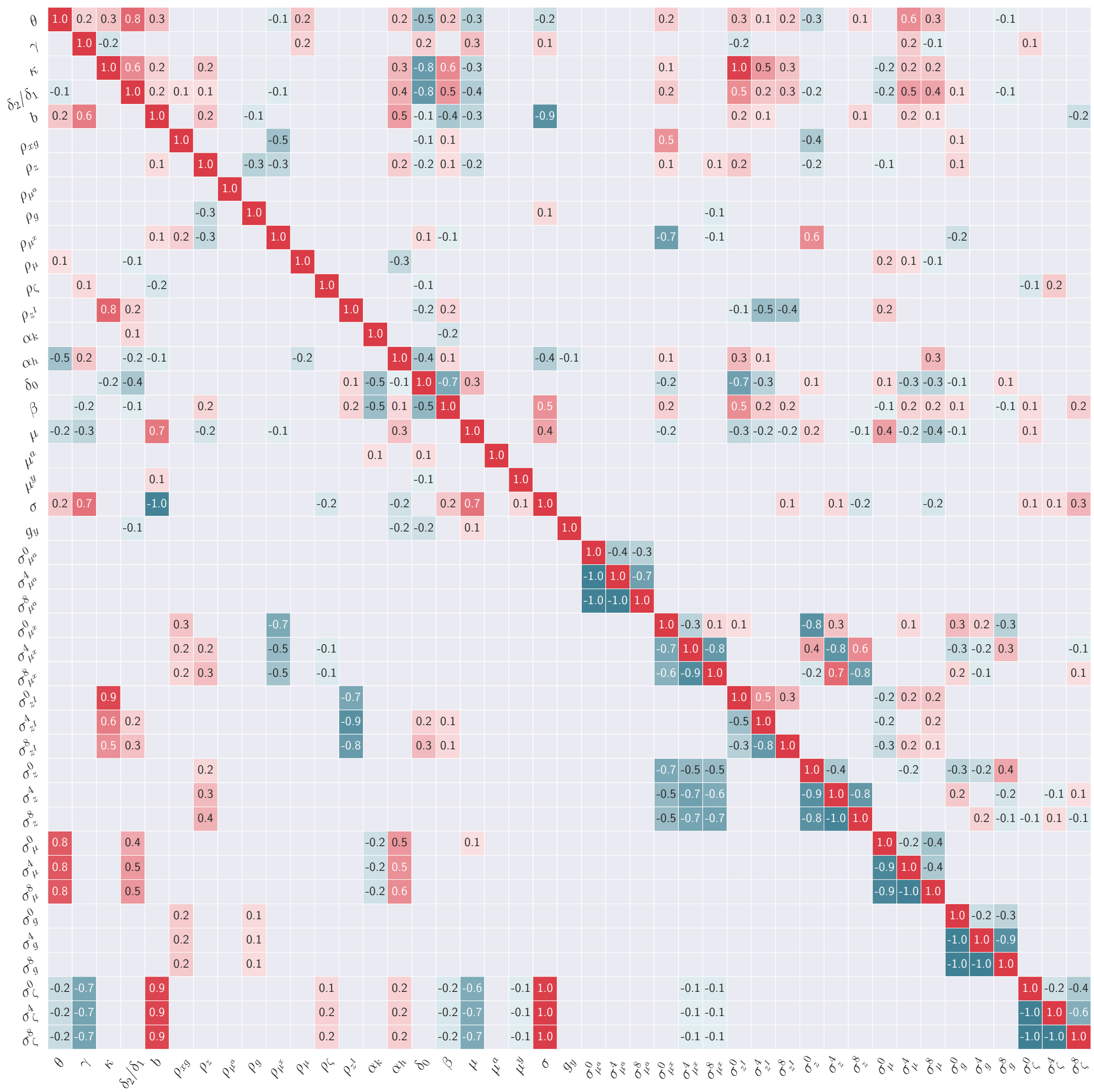

Figure 8: Parameter correlations in the SGU model. The lower triangle of the matrix shows the conditional correlation coefficients between each pair of parameters. The upper triangle shows the marginal correlation coefficients. The values are obtained from the joint asymptotic posterior distribution of the parameters evaluated at the MLE in SGU. Correlation coefficients smaller than .1 in absolute value are not displayed. Off-diagonal values of -1 or 1 are due to rounding errors.

References

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017): "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 132, 1553–1592.

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

Chen, C.-F. (1985): "On asymptotic normality of limiting density functions with Bayesian implications," *Journal of the Royal Statistical Society. Series B (Methodological)*, 540–546.

Cramér, H. (1946): *Mathematical Methods of Statistics*, Princeton University Press.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, and F. Schorfheide (2016): "Solution and estimation methods for DSGE models," in *Handbook of Macroeconomics*, Elsevier, vol. 2, 527–724.

Heyde, C. and I. Johnstone (1979): "On asymptotic posterior normality for stochastic processes," *Journal of the Royal Statistical Society. Series B (Methodological)*, 184–189.

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

Kim, J.-Y. (1998): "Large sample properties of posterior densities, Bayesian information criterion and the likelihood principle in nonstationary time series models," *Econometrica*, 359–380.

Magnus, J. R. and K. M. Abadir (2005): *Matrix Algebra*, Cambridge University Press.

Müller, U. K. (2012): "Measuring prior sensitivity and prior informativeness in large Bayesian models," *Journal of Monetary Economics*, 59, 581–597.

Phillips, P. C. and W. Ploberger (1996): "An asymtotic theory of Bayesian inference for time series," *Econometrica: Journal of the Econometric Society*, 381–412.

Qu, Z. and D. Tkachenko (2012): "Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models," *Quantitative Economics*, 3, 95–132.

Schmitt-Grohé, S. and M. Uribe (2012): "What's News in Business Cycles," *Econometrica*, 80, 2733–2764.

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

Walker, A. (1969): "On the asymptotic behaviour of posterior distributions," *Journal of the Royal Statistical Society. Series B (Methodological)*, 80–88.

Appendix

A Figures

A.1 Smets and Wouters (2007) model

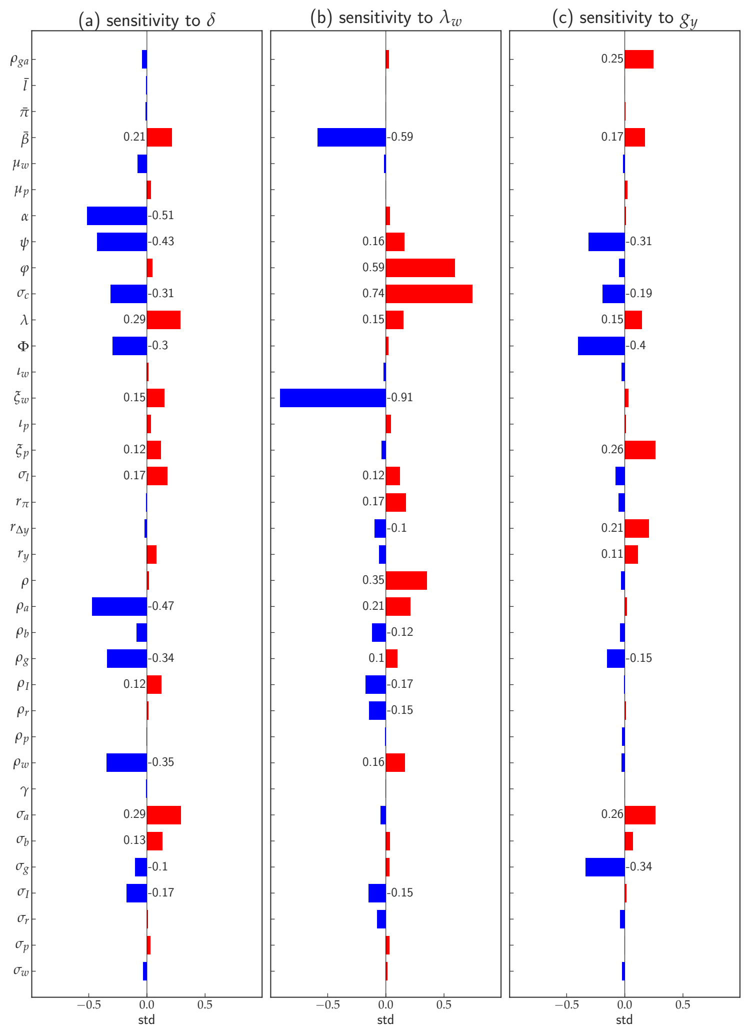

Figure A1: Sensitivity to changes in the calibrated parameters. Each panel shows the effect of a one-standard-deviation increase in the respective parameter on the value of each free parameter, in units of standard deviations. Only one parameter is held fixed at a time.

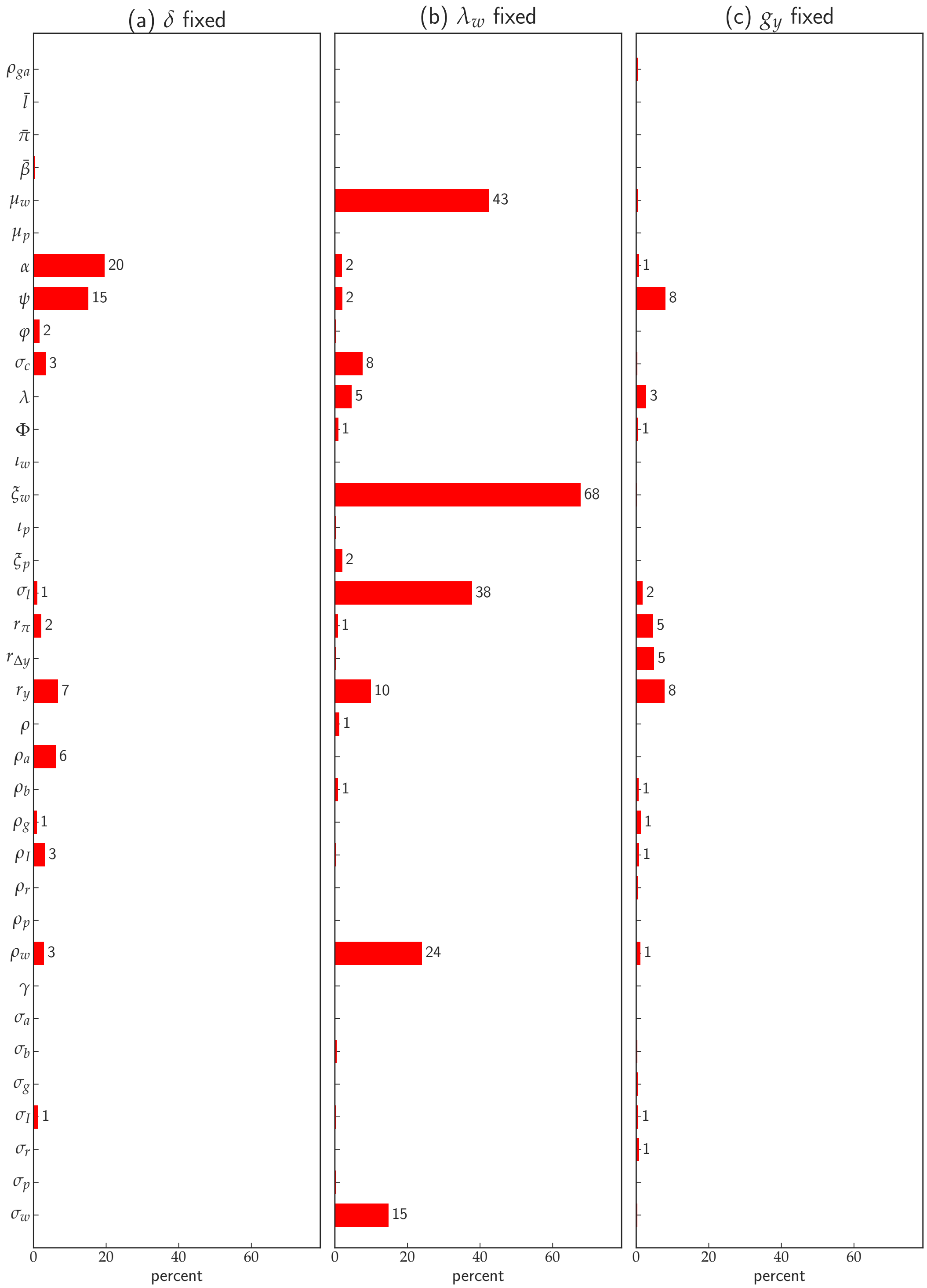

Figure A2: Pairwise conditional information gains. The values show the reduction of uncertainty about a parameter from knowing either the value of δ, λ, or g_y, and conditinal on knowing *all* other parameters.

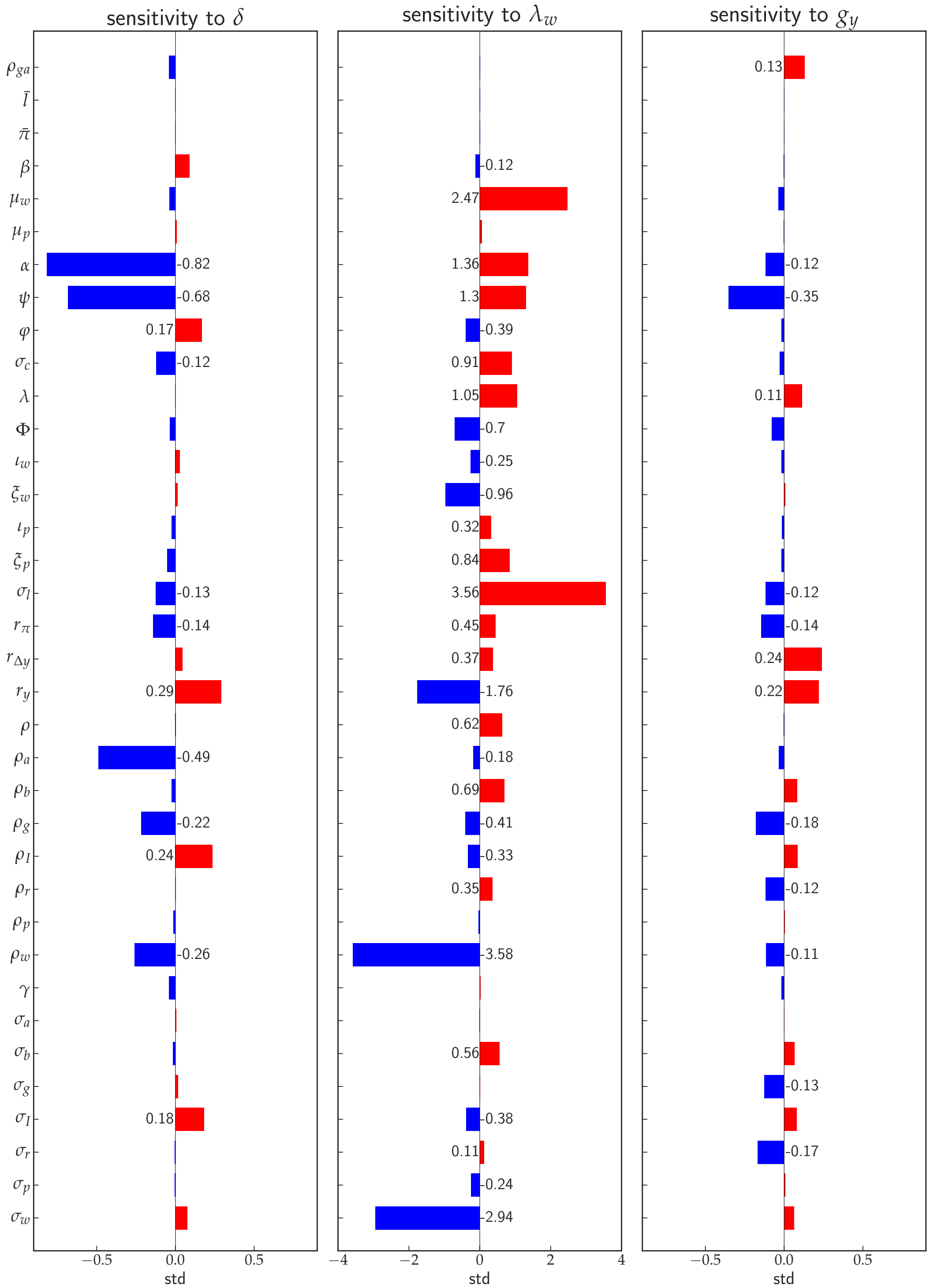

Figure A3: Pairwise conditional sensitivities. The values shows the effect, in units of standard deviations, of a one-standard-deviation increase in the value of δ, λ, or g_y on the value of each free parameter, assuming all remaining parameters are known and remain fixed.

A.2 Schmitt-Grohé and Uribe (2012) model

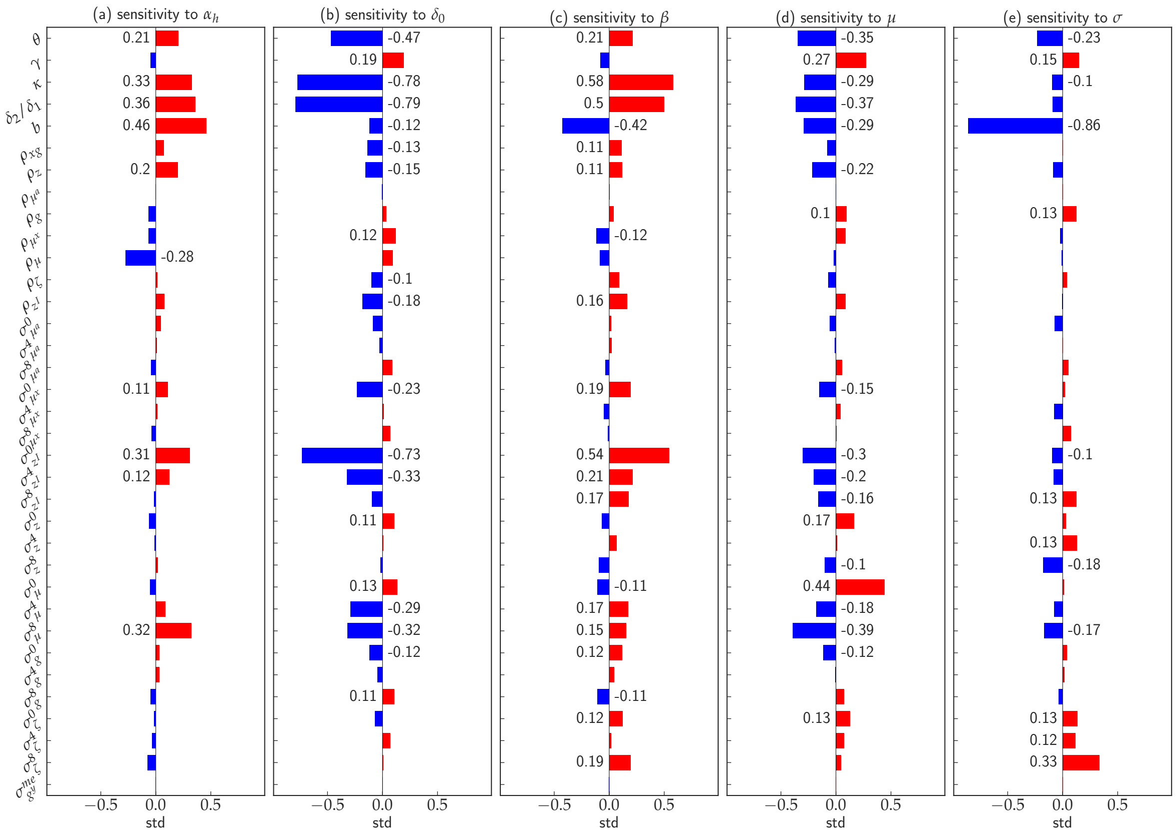

Figure A4: Sensitivity to changes in the calibrated parameters. Each panel shows the effect of a one-standard-deviation increase in the respective parameter on the value of each free parameter, in units of standard deviations. Only one parameter is held fixed at a time.

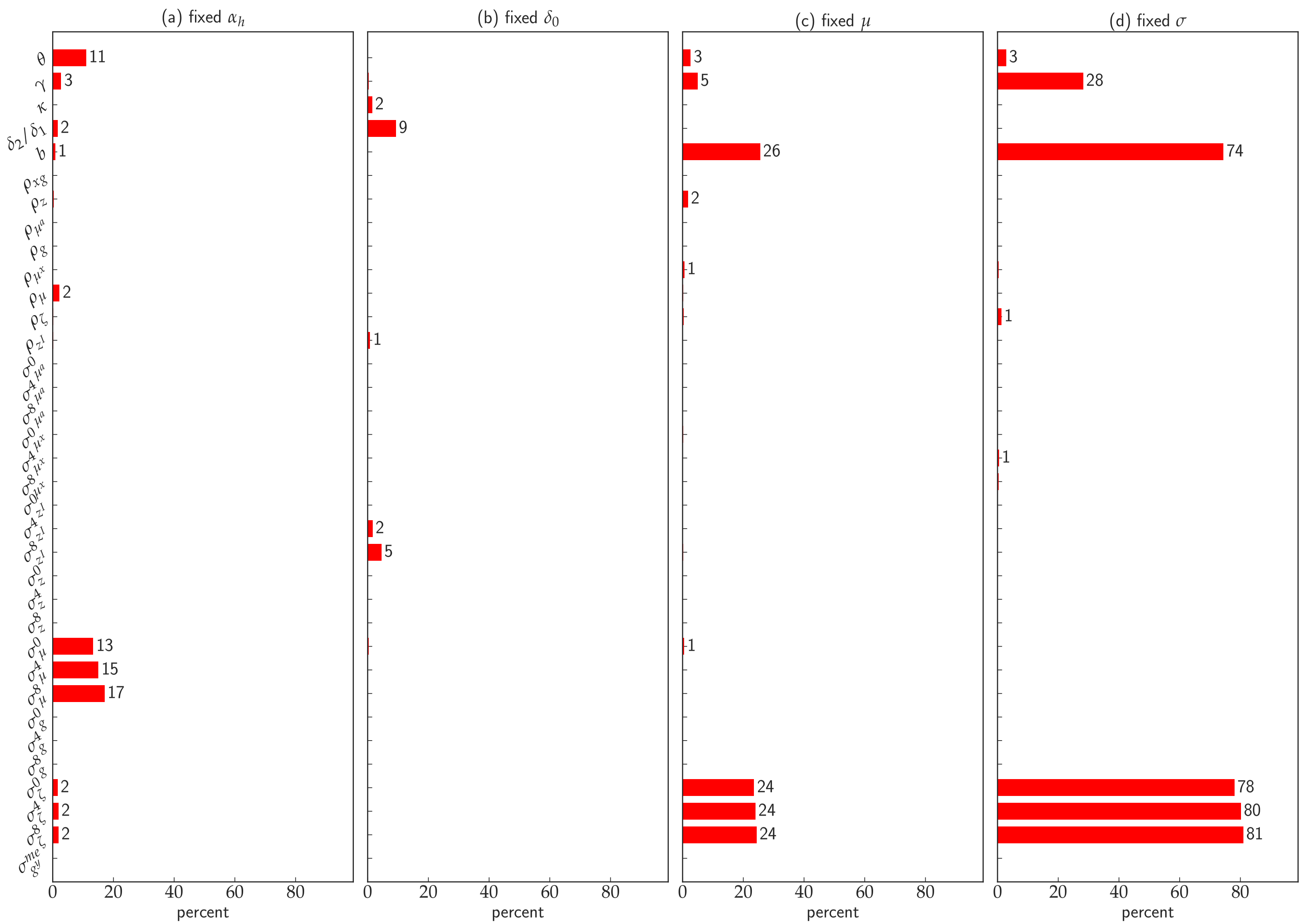

Figure A5: Pairwise conditional information gains. The values show the reduction of uncertainty about a parameter from knowing either the value of δ, λ, or g_y, and conditinal on knowing *all* other parameters.

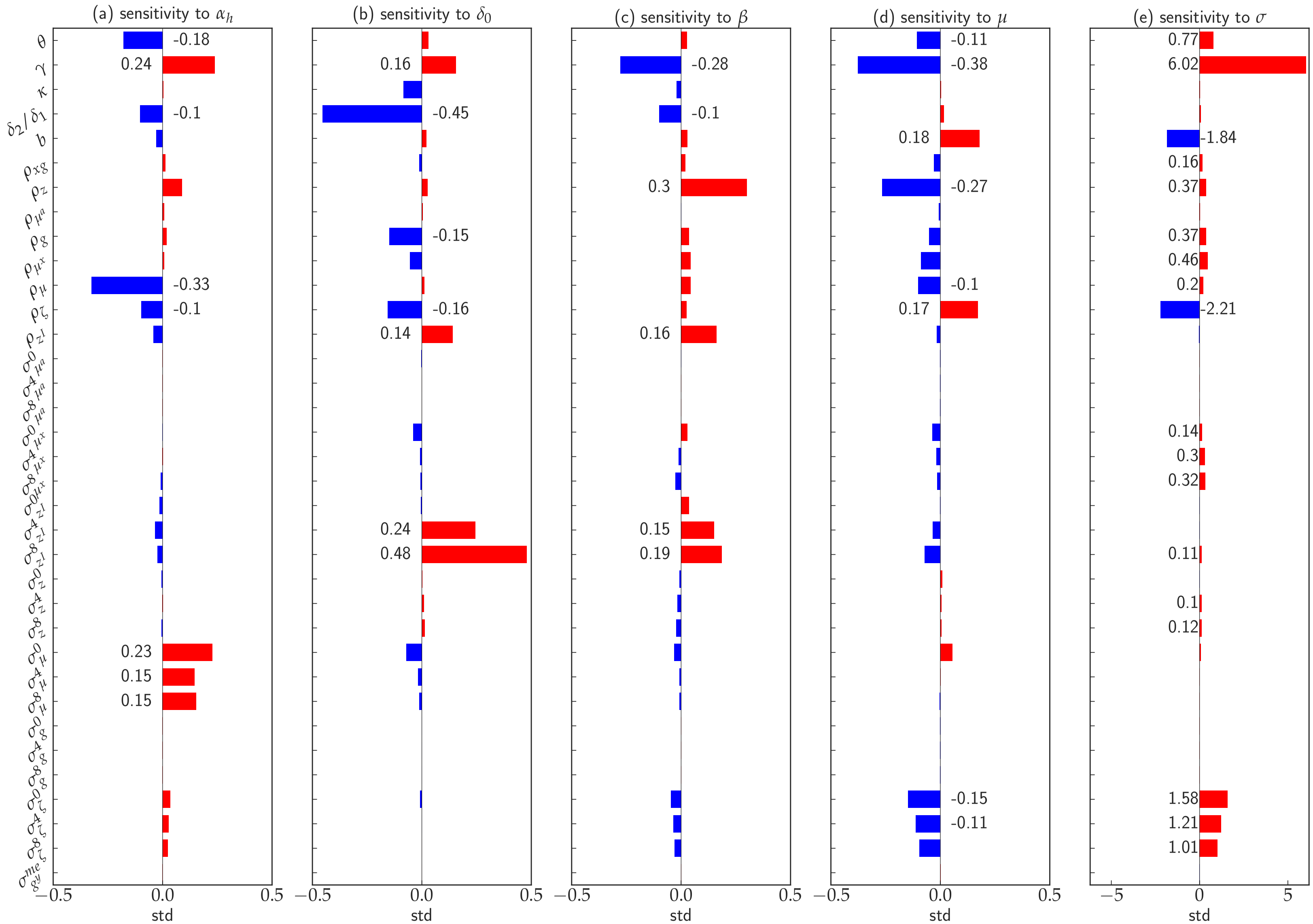

Figure A6: Pairwise conditional sensitivities. The values shows the effect, in units of standard deviations, of a one-standard-deviation increase in the value of δ, λ, or g_y on the value of each free parameter, assuming all remaining parameters are known and remain fixed.

B Simulations

Model

In this appendix I compare my measures' predictions about the effect of calibration to actual outcomes from estimating a model under different calibrations assumptions using simulated data. To that end I use a simple New Keynesian DSGE model taken from Fernández-Villaverde et al. (2016). In its log-linearized form the model consists of the following equations:

$$\hat{x}_t = \mathrm{E}_t\, \hat{x}_{t+1} - \left(\hat{R}_t - \mathrm{E}_t\, \hat{\pi}_{t+1}\right) + \mathrm{E}_t\, z_{t+1} \tag{B.1}$$

$$\hat{\pi}_t = \beta\, \mathrm{E}_t\, \hat{\pi}_{t+1} + \kappa_p \left(\hat{w}_t + \lambda_t\right) \tag{B.2}$$

$$\hat{w}_t = (1+\nu)\, \hat{x}_t + \phi_t \tag{B.3}$$

$$\hat{R}_t = \psi \hat{\pi}_t + \sigma_R \epsilon_{R,t}, \qquad \epsilon_{R,t} \sim \mathcal{N}(0,1) \tag{B.4}$$

$$z_t = \rho_z z_{t-1} + \sigma_z \epsilon_{z,t}, \qquad \epsilon_{z,t} \sim \mathcal{N}(0,1) \tag{B.5}$$

$$\lambda_t = \rho_\lambda \lambda_{t-1} + \sigma_\lambda \epsilon_{\lambda,t}, \qquad \epsilon_{\lambda,t} \sim \mathcal{N}(0,1) \tag{B.6}$$

$$\phi_t = \rho_\phi \phi_{t-1} + \sigma_\phi \epsilon_{\phi,t}, \qquad \epsilon_{\phi,t} \sim \mathcal{N}(0,1) \tag{B.7}$$

where $\psi = 1/\beta$, and $\kappa_p = \frac{1-\zeta_p\beta}{\zeta_p}$. The variables $\hat{x}_t$, $\hat{\pi}_t$, $\hat{w}_t$ and $\hat{R}_t$ are the log-deviations of output, inflation, wages and the nominal interest rate from their respective steady state values. There are four shocks in the model: a technology growth shock z_t, a preference shock ϕ_t, a price markup shock λ_t, and a monetary policy shock $\epsilon_{R,t}$.

Fernández-Villaverde et al. (2016) use this stylized model to show how to solve and estimate DSGE models, and how to evaluate their properties and performance. Among other things, they use the model to study the sampling distribution of the maximum likelihood estimator of ζ_p when the model is correctly specified and the other deep parameters are assumed to be known and fixed at their true values (see Section 11.1.2). They do that by repeatedly generating random samples for four observed variable – output growth ($\hat{x}_t - \hat{x}_{t-1} + z_t + \log(\gamma)$), the labor share ($\hat{w}_t + \log(w^*)$), inflation rates ($\pi_t + \log(\pi^*)$) and net interest rates ($\hat{R}_t + \log(\pi^*\gamma/\beta)$), and evaluating the ML estimator for each one of them. Here I extend their analysis in two ways. First, I assume that one of the fixed parameters is mis-calibrated, i.e. during estimation it is held fixed at a value different from the one used to generate the random samples. This is done for each calibrated parameter, one at a time, with both positive and negative calibration errors. The goal of this exercise is to see how sensitive the estimate of ζ_p is to errors in the calibration of the fixed parameters. Second, I simulate and estimate the model assuming that one of the originally calibrated parameters is free and its value has to be estimated together with that of ζ_p. Again, this is done for each parameter, one at a time, and the purpose of the exercise is to find out how the sampling uncertainty about ζ_p is affected by uncertainty about the true values of other parameters. The true values of the parameters are shown in Table B1 and are the same as in Fernández-Villaverde et al. (2016). Since there are at most 2 free parameters at a time, finding the ML estimator for

Table B1: Parameter values

	parameter	value
β	discount factor	0.9901
γ	growth rate of technology	exp(0.005)
λ	steady-state price markup	0.15
π^*	target inflation rate	exp(0.005)
ζ_p	Calvo probability	0.65
ν	Frisch elasticity parameter*	0
ρ_ϕ	AR preference shock	0.94
ρ_λ	AR price markup shock	0.88
ρ_z	AR technology growth shock	0.13
σ_ϕ	standard deviation preference shock	0.01
σ_λ	standard deviation price markup shock	0.01
σ_z	standard deviation technology growth shock	0.01
σ_R	standard deviation monetary policy shock	0.01

* The Frisch labor supply elasticity is $1/(1+\nu)$

each sample is done by a grid search. This virtually guarantees that a global maximum is always found.

The first exercise requires making a choice about the size of the calibration error. One possibility is to use the standard deviation of the marginal asymptotic posterior distribution of each parameter, e.g. an error of plus or minus one standard deviation. The sampling distributions of $\hat{\zeta}_p$ for that case are displayed in Figure B1. The sample size is assumed to be 200. The results show that the estimates of ζ_p are most sensitive to calibration errors in ρ_λ and σ_λ, while errors in ν, ρ_ϕ and β, have small but no-zero impact. One-standard-deviation errors in the remaining parameters have no impact, conditional, in each case, on the other parameters being fixed at their true values. In the case of ρ_λ, σ_λ, ρ_ϕ and β, positive errors shift the distribution of $\hat{\zeta}_p$ to the right, i.e. lead to a positive bias in the estimates, while negative errors result in negative bias. Calibration error in ν implies an estimation error for ζ_p with the opposite sign.

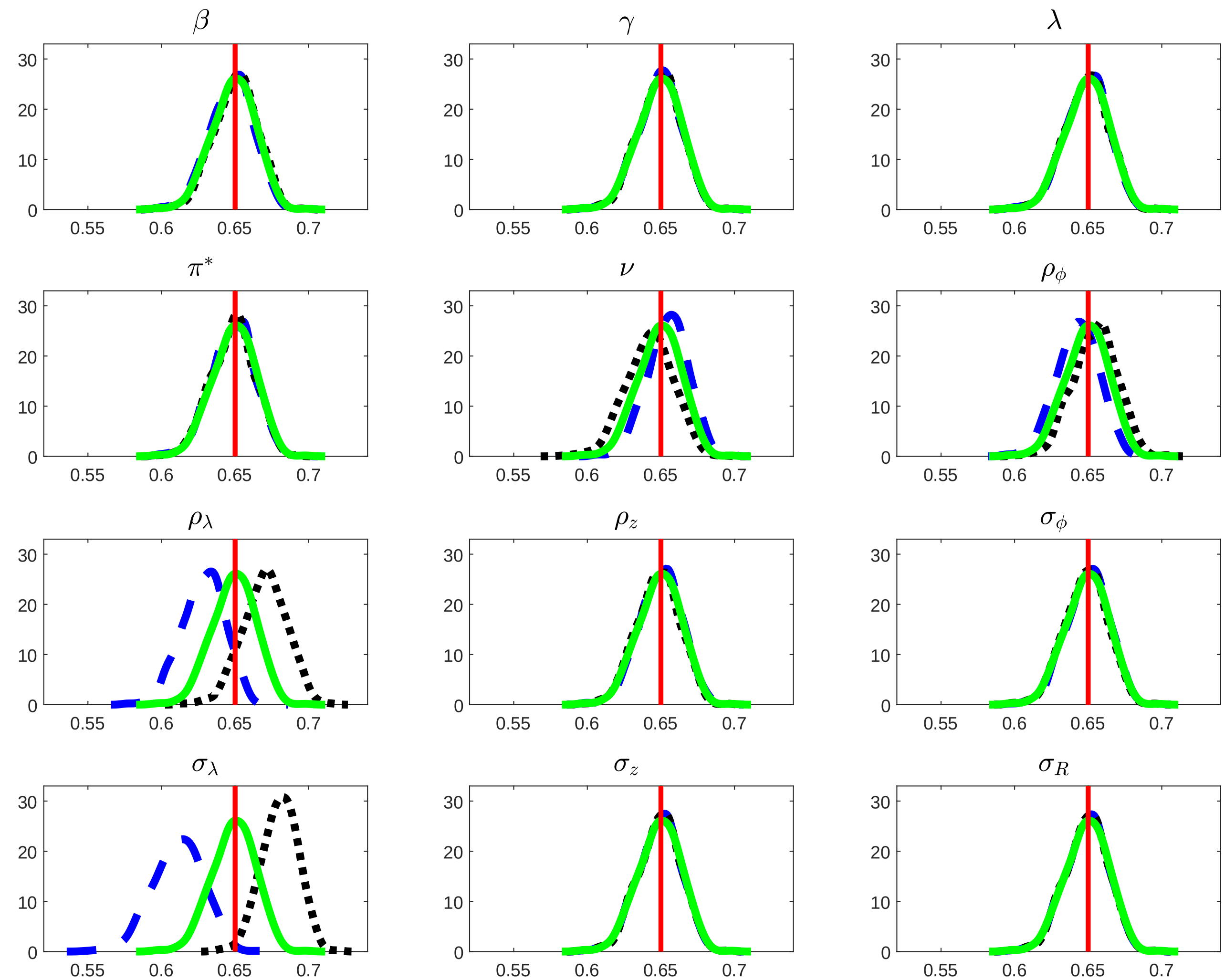

Figure B1: Sensitivity of the MLE of ζ_p to mis-calibration of the fixed parameters by 1 std. The figure shows the densities of the sampling distribution of $\hat{\zeta}_p$ for a given parameter being fixed at a value plus (black) or minus (blue) 1 std. away from the true value of that parameter. The sampling density without mis-calibration is shown in green. The red vertical line shows the true value of ζ_p.

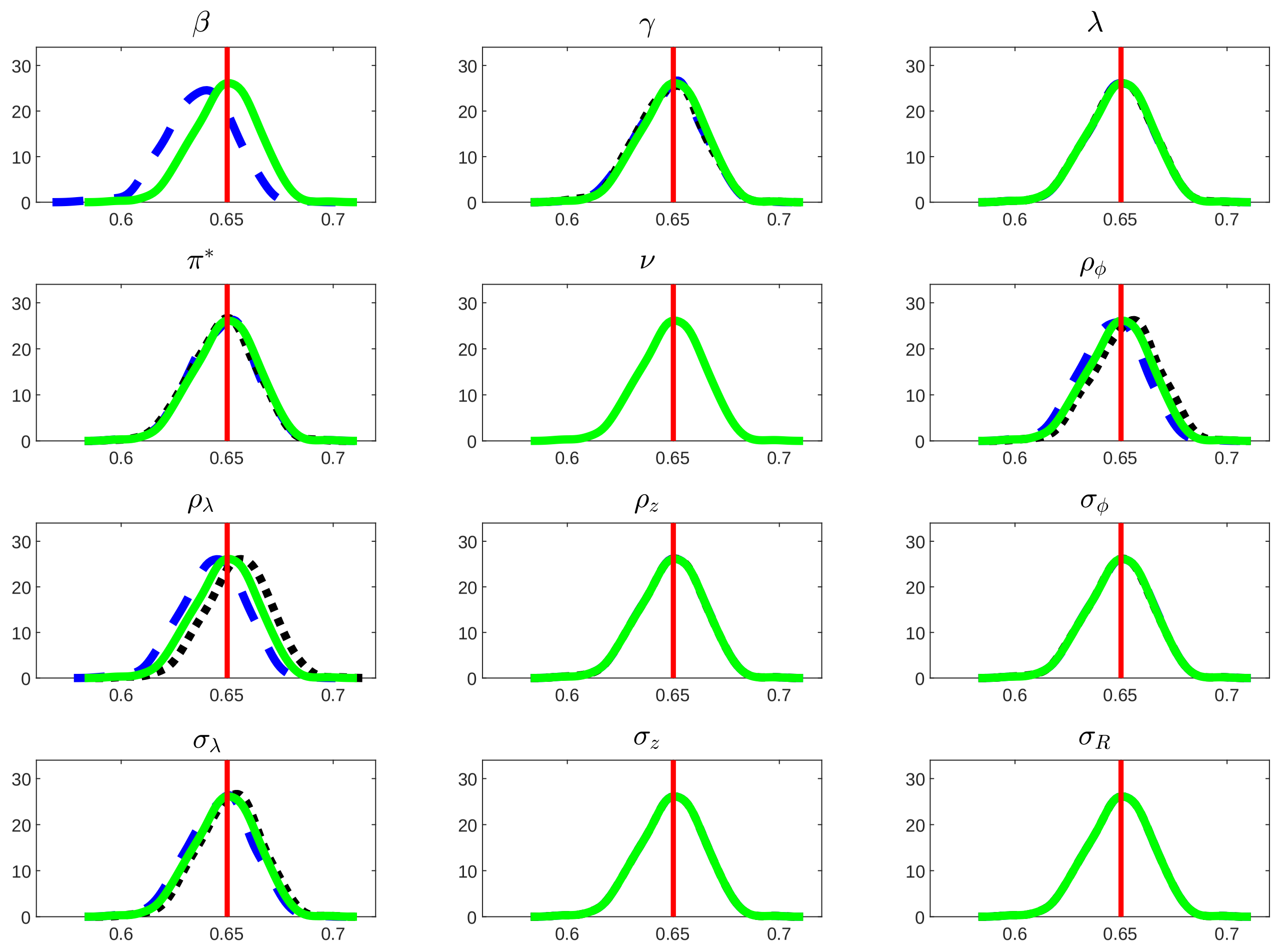

Figure B2: Sensitivity of the MLE of ζ_p to mis-calibration of the fixed parameters by 1%. The figure shows the densities of the sampling distribution of $\hat{\zeta}_p$ for a given parameter being fixed at a value plus (black) or minus (blue) 1% away from the true value of that parameter. In the case of β only mis-calibration by -1% is shown. The sampling density without mis-calibration is shown in green. This is the only density shown for ν. The red vertical line shows the true value of ζ_p.

Table B2: Sensitivity of $\hat{\zeta}_p$ to calibration errors

mis-	1 std.				1%			
calibrated	simulation		analytical		simulation		analytical	
parameter	$-$	$+$	$-$	$+$	$-$	$+$	$-$	$+$
β	-0.04	0.03	-0.03	0.03	-1.8	n.a	-1.8	1.8
γ	0	0	0	0	0	0	0	0
λ	0	0	0	0	0	0	0	0
π^*	0	0	0	0	0	0	0	0
ν	0.15	-0.19	0.17	-0.17	n.a	n.a	n.a	n.a
ρ_ϕ	-0.17	0.15	-0.16	0.16	-0.5	0.5	-0.5	0.5
ρ_λ	-0.57	0.58	-0.58	0.58	-0.9	0.9	-0.9	0.9
ρ_z	0.01	-0.03	0.02	-0.02	0.0	0	0	0
σ_ϕ	0.01	-0.02	0.01	-0.01	0.0	0	0	0
σ_λ	-1.06	0.87	-0.97	0.97	-0.4	0.3	-0.3	0.3
σ_z	0	0	0	0	0	0	0	0
σ_R	0	0	0	0	0	0	0	0

See description in the text. The sample size is 200.

Another possibility is to introduce calibration errors that are proportional to the true value of each parameter, for instance, 1% of the true value. Figure B2 shows the sampling distributions of $\hat{\zeta}_p$ for that case, again for samples of size 200. Note that, since the true values of ν is 0, no calibration errors are shown for that parameter. Also, only the case of a negative error is displayed for β since adding 1% to the true value results in $\beta = 1$ which is not permissable. However, another simulation with proportional errors smaller than 1% shows that the sampling distribution for positive error is nearly identical to the one for negative error, but shifted to right of the distribution of $\hat{\zeta}_p$ when there is no calibration error. As before, calibration errors in ρ_λ, σ_λ, ρ_ϕ and β have non-zero impact on the distribution of $\hat{\zeta}_p$. The size of the impact is different, in particular, now errors in β have the largest effect, followed by ρ_λ, σ_λ and ρ_ϕ. Such a discrepancy is to be expected because estimation uncertainty, represented here by the asymptotic standard deviation, is affected by multiple factors not just the parameter value. This is not an issue in practice since one would be interested in measuring and comparing how a calibration error of a *given* size in a *given* fixed parameter affects different estimated parameters. The goal here is to assess the accuracy of the predictions delivered by the analytical measures introduced in the paper by comparing them to the simulation results. Table B2 does that, showing, under the label "simulation", the errors in $\hat{\zeta}_p$ induced by an error in each fixed parameter, and under "analytical" - the values of the analytical sensitivity measure. The errors in $\hat{\zeta}_p$ are defined as the normalized difference between the mean of

Table B3: Information gains due to calibration (%)

parameter	simulation	analytical
β	1	0
γ	0	0
λ	0	0
π^*	1	0
ν	2	4
ρ_ϕ	5	5
ρ_λ	30	25
ρ_z	1	0
σ_ϕ	0	0
σ_λ	43	39
σ_z	0	0
σ_R	0	0

The information gain is measured as $100 \times (x - y)/x$ where x is the standard deviation of the distribution of $\hat{\zeta}_p$ when ζ_p is one of two estimated parameters and y is the standard deviation when only ζ_p is estimated. The sample size is 200.

the sampling distribution of $\hat{\zeta}_p$ when there is error in calibration and the mean without miscalibration. As explained above, the normalization is done in two ways: dividing by the standard deviation of $\hat{\zeta}_p$ – under the label "1 std." and dividing by the true value of ζ_p and multiplying by 100 – under "1%". Therefore, the units of the errors are standard deviations of ζ_p for a 1 standard deviation error in a fixed parameter, and percent deviations in ζ_p for a 1 percent error in a fixed parameter. These normalizations make the simulation results comparable to the values of the analytical measure in the table. As can be seen, the analytical measures are very accurate not only in terms of their qualitative predictions – about the sign and relative magnitude of the effect of the errors in calibration, but also in terms of the predicted size of the effect. The main discrepancy between the analytical and simulation results is that the analytical measure is by construction symmetric, while the simulation results in some case show some asymmetry in the effect of positive and negative calibration errors.

In the second exercise I compare the sampling uncertainty of $\hat{\zeta}_p$ when there is another parameter to estimate to the uncertainty when ζ_p is the only free parameter. The reduction of uncertainty, measured my the standard deviation, as a percent of the uncertainty with two free parameters corresponds to the measure of (conditional) information gain introduced in the paper. Table B3 shows the simulation results and the values obtained with the analytical measure of information gain. As above, the sample size is assumed to be 200. Again, the results are very similar, not only qualitatively but also numerically. The largest gain for ζ_p (price stickiness) come from fixing either one of

Table B4: Sensitivity and information gains for sample size of 80

parameter	sensitivity (std)				information gain (%)	
	simulation		analytical		simulation	analytical
	$-$	$+$	$-$	$+$		
β	-0.04	0.04	-0.03	0.03	0	0
γ	0	0	0	0	0	0
λ	0	1	0	0	0	0
π^*	0	0	0	0	0	0
ν	0.16	-0.18	0.17	-0.17	5	4
ρ_ϕ	-0.17	0.15	-0.16	0.16	4	5
ρ_λ	-0.57	0.59	-0.57	0.57	24	25
ρ_z	0.02	-0.03	0.02	-0.02	0	0
σ_ϕ	0.02	-0.01	0.01	-0.01	0	0
σ_λ	-1.05	0.87	-0.97	0.97	38	39
σ_z	0	0	0	0	0	0
σ_R	0	0	0	0	0	0

The table shows sensitivity of $\hat{\zeta}_p$ to calibration errors (for 1 std.) and the information gains due to calibration. The sample size is 80.

the two parameters of the price markup process. This is in line with the earlier result about the sensitivity to calibration for calibration errors which are proportional to the intrinsic uncertainty of the fixed parameters.

Table B4 compares simulation-based results and analytical predictions when the sample size is assumed to be 80. In the case of the sensitivity measures only sensitivity to errors of 1 standard deviation are reported. There only very small changes compared to the earlier results, and the analytical predictions remain very close to the simulation-based outcomes.