On the sources of information in the moment structure of dynamic macroeconomic models

Nikolay Iskrev

Abstract

What features of the data are the key sources of information about the parameters in structural macroeconomic models? As such models grow in size and complexity, the answer to this question has become increasingly difficult. This paper shows how to identify the main sources of parameter information across different parts of the moment structure of macroeconomic models. In particular, we propose a measure of the relative contribution of information by a given subset of moments with respect to any parameter of interest. The measure is trivial to compute even for large-scale models with many free parameters and observed variables. We illustrate our method with an application to a news-driven business cycle model developed by Schmitt-Grohé and Uribe (2012).

Keywords: DSGE models, Sources of Information, Generalized Method of Moments, News Shocks

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. The views expressed in this publication are those of its author and do not necessarily reflect the views of Banco de Portugal or the Eurosystem.
Av. Almirante Reis 71, 1150-012, Lisboa, Portugal, +351213130006, nikolay.iskrev (at) bportugal.pt, https://github.com/niskrev

1 Introduction

As macroeconomic models grow in size and complexity, it has become increasingly difficult to determine where in the data information about different model parameters comes from. This lack of transparency is one of the main reasons why the existing research on empirical dynamic stochastic general equilibrium (DSGE) models is viewed by some with a high degree of skepticism.[1]

In this paper we propose a new method for identifying the main sources of information about parameters in the moment structure of DSGE models. Thus, we address the question of where information comes from in terms of which particular moment or a group of moments contribute the most information with respect to different parameters. Starting with a complete set of available moments, as a representation of the total amount of available information, our method allows us to quantify the relative importance of any given subset of moments with respect to individual parameters of interest. In particular, we are able to compare the informativeness of moments of individual variables or of cross-moments of several variables, as well as to rank individual moments in terms of the amount of information they contribute with respect to a specific parameter.

We study the information content of moments by adopting the framework of the generalized method of moments (GMM). As shown in Hansen (1982), GMM estimators based on a set of valid population moment conditions are consistent and asymptotically normally distributed. Furthermore, from Chamberlain (1987) we know that optimally-weighted GMM estimators are asymptotically efficient in the class of estimators based on the same set of moment conditions. Therefore, the amount of information in a given set of moment conditions can be quantified using the asymptotic covariance matrix of the efficient GMM estimator. We use this property to measure and compare the information

[1] See e.g Blanchard (2017) and Romer (2016) for broader critical appraisals of the current DSGE framework.

content of different parts of the moment structure of DSGE models. Specifically, we consider a GMM estimator with population moment conditions defined as the difference between sample and model-implied moments of observed variables. We then measure the amount of information in a particular set of moments using the covariance matrix of the asymptotic distribution of the efficient GMM estimator based on those moments. By conditioning on different parts of the moment structure and comparing the resulting covariance matrices, we are able to assess the relative informativeness of different subsets of moments.

We emphasize that the purpose of our analysis is not to guide researchers in the selection of a small number of moments, which then to use for either estimation or calibration of a given model. DSGE models are typically estimated with full information approaches, either by maximum likelihood or by Bayesian methods. In either case, all sample information about the estimated parameters is contained in the likelihood function; our goal here is to determine which parts of the likelihood provide the most information about parameters of interest. To be concrete, a well-known property of linearized Gaussian models – the class of models we focus on, is that all relevant information is contained in the first and second-order moments of the observed variables. The question we address, therefore, is: what is the relative contribution of information by different moments – means, covariances and (cross) auto-covariances, with respect to any given model parameter?

To the best of our knowledge, ours is the first paper that shows how to formally analyze the sources of information in the moment structure of DSGE models. The issue itself is sometimes discussed in the empirical literature, largely in an informal manner. A notable example is Schorfheide (2008) who investigates the sources of information about the parameters of the New Keynesian Phillips curve. The author emphasizes the information advantages of using a full information likelihood-based system approach over single-equation estimation, and describes how information in contemporaneous and dynamic interactions among observed variables help identify the Phillips curve parameters. At the same time, in spite of the simplicity of the model, which is a stripped down three-equation New Keynesian system, the discussion is largely verbal without an attempt to quantify the relative importance of different information sources. A more recent example in the same vain is Ríos-Rull et al. (2012) who discuss the sources of information with respect to the aggregate labor supply elasticity parameter in a standard real business cycle model.

In terms of methodology, our paper is most closely related to Andrews et al. (2017a) who propose a local measure of sensitivity of parameter estimates to moments of the data. Similar to our approach, their measure is based on an asymptotic approximation of the mapping from moments to estimated parameters. Instead of quantifying the relative informativeness of the moments, however, the purpose of their measure is to identify the moments to which the parameter estimates are most sensitive, and which, if misspecified, could cause potentially large biases in the results. The second and more fundamental difference between our paper and Andrews et al. (2017a) is that their sensitivity measure is interpreted as a property of an estimator, while our measure of information content is a property of a model. Our approach takes a completely specified model as given,[2] and asks what that model implies about the information content of different parts of the moment structure of the model's observable variables. A crucial requirement for us to be able to answer this question is that, in addition to having valid GMM moment conditions, we also use the correct optimal weighting matrix. Any other suboptimal moment weighting scheme would no longer be about the model per se as it would not fully capture the true information content of the moments under consideration.[3] One of

[2]A model is completely specified if we know everything necessary to simulate data from that model.

[3]For instance, it is easy to come up with examples where suboptimal weighting matrices would imply that using a larger number of valid moments is asymptotically less efficient than using fewer moments,

the contributions of our paper is to show how to evaluate the true optimal weighting matrix for a wide range of DSGE models. In contrast, Andrews et al. (2017a) take an estimator, i.e. a set of moment conditions and a weighting matrix, as given, and ask how perturbations in the moment conditions, due to misspecification in some dimensions, are translated by the estimator into asymptotic bias in the parameter estimates. Hence, their moment sensitivity analysis can be applied in a wide range of estimation problems, while our approach is only feasible in the context of fully specified models, such as DSGE models.

To be clear, one does not have to accept that a particular model is correctly specified for our analysis to be useful.[4] Its purpose is not to criticize models or evaluate their plausibility, but to understand what a given model implies about the main sources of information with respect to its parameters. We believe that conducting and reporting the results from this analysis will benefit both researchers estimating DSGE models and the readers of such research, by improving their understanding and increasing the transparency of estimated structural models. In that respect, our analysis is related to, and can be thought of as an extension of the local identification analysis of DSGE models (see, inter alia, Canova and Sala (2009), Iskrev (2010), Komunjer and Ng (2011), Qu and Tkachenko (2012)). In addition to checking whether the parameters of DSGE model are locally identified, we ask what the main sources of the information are. Similarly to the local identification analysis, the model itself, as well as the point in the parameter space where the analysis is conducted, are taken as given, and the answers one obtains are conditional on the model and the values.

Our paper is also related to the literature on redundant moments in GMM. Breusch et al. (1999) define as redundant moment conditions which contribute no information with

see Meng and Xie (2014).

[4]See Inoue et al. (2019) for a recent work on detecting misspecification in DSGE models.

respect to the estimated parameters. More specifically, a set of moment conditions is called redundant relative to another set of moment conditions if adding the former to the latter does not increase the asymptotic efficiency of the GMM estimator. Similarly, moment conditions which contribute no additional information with respect to a subset of the estimated parameters are called partially redundant. Redundancy and partial redundancy are binary concepts and the purpose of detecting redundant moment conditions is to achieve a better finite sample behavior of the GMM estimator by removing such moments.[5] Instead, our goal here is to determine the relative information content of different moments. Therefore, we are interested not just in whether there is an increase in efficiency as a result of adding a set of moments, but also in the magnitude of the efficiency gain. Note that, like the notion of redundancy, our measure of the amount of information contributed by a set of moments is conditional on what other moments are being utilized. This means that if, for instance, two sets of moments contain similar information, the relative contribution of either one of them is small. In other words, each set of moments contributes a small amount of unique information. We show an example of this phenomenon in our application where the relative contribution of information by moments of investment is shown to be much larger when consumption is not observed than when it is observed.

The remainder of the paper is organized as follows. Section 2 introduces the general setup, and reviews some standard properties of GMM estimators. It also defines our measure of the amount of information contributed by a given set of moments. The proposed measure is applied in Section 3 to identify the main sources of information with respect to parameters of a medium-scale DSGE model. The model is a real business cycle model with news shocks taken from Schmitt-Grohé and Uribe (2012). We study the

[5]See Andrews et al. (2017b) for simulation evidence on the effect of using redundant moment conditions in instrumental variables estimation.

contribution of information by moments across several dimensions, namely: lag structure, single and pairs of observed variables, as well as individual moments. Section 4 contains some concluding remarks.

2 Methodology

This section describes our approach to measuring the information content of moments with respect to DSGE model parameters. First, we introduce the class of linearized Gaussian models and derive their moment structure. Then, we review some standard properties of GMM estimators and show how to use them to quantify the contribution of information by a set of moments with respect to parameters of interest.

2.1 Moment structure of linear Gaussian DSGE models

A linearized DSGE model can be expressed as a linear state space system with a state transition and measurement equations given by:[6]

$$\boldsymbol{s}_t = \boldsymbol{\Phi}_1(\boldsymbol{\theta})\boldsymbol{s}_{t-1} + \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t \tag{2.1}$$

$$\boldsymbol{y}_t = \boldsymbol{\Psi}_0(\boldsymbol{\theta}) + \boldsymbol{\Psi}_1(\boldsymbol{\theta})\boldsymbol{s}_t \tag{2.2}$$

where $\boldsymbol{y}_t$ is a $n_{\boldsymbol{y}}$ vector of observed variables, $\boldsymbol{s}_t$ is a $n_{\boldsymbol{s}}$ vector of state variables, $\boldsymbol{\varepsilon}_t$ is a $n_{\boldsymbol{\varepsilon}}$ vector of exogenous shocks, $\boldsymbol{\Phi}_1$ is $n_{\boldsymbol{s}} \times n_{\boldsymbol{s}}$ matrix, $\boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}$ is a $n_{\boldsymbol{s}} \times n_{\boldsymbol{\varepsilon}}$ matrix, $\boldsymbol{\Psi}_0$ is $n_{\boldsymbol{y}}$ vector, and $\boldsymbol{\Psi}_1$ is a $n_{\boldsymbol{y}} \times n_{\boldsymbol{s}}$ matrix. In general, the coefficient matrices in (2.1) and (2.2) are functions of the structural parameters of the model, collected in the $n_{\boldsymbol{\theta}}$ vector $\boldsymbol{\theta}$.

Assume that the structural shocks are drawn from a Gaussian distribution, i.e. $\boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{I}_{n_{\boldsymbol{\varepsilon}}}\right)$. The linear state space structure of the model implies that the autocovariance

[6] See Fernández-Villaverde et al. (2016) for more details on the results presented here.

matrix $\boldsymbol{\Gamma}_{\boldsymbol{yy}}(h) = \operatorname{cov}(\boldsymbol{y}_t, \boldsymbol{y}_{t-h})$ can be computed as

$$\boldsymbol{\Gamma}_{\boldsymbol{yy}}(h) = \boldsymbol{\Psi}_1(\boldsymbol{\theta})\boldsymbol{\Gamma}_{\boldsymbol{ss}}(h)\boldsymbol{\Psi}_1(\boldsymbol{\theta})' \tag{2.3}$$

where, for $h = 0$, $\boldsymbol{\Gamma}_{\boldsymbol{ss}}(h)$ solves the equation

$$\boldsymbol{\Gamma}_{\boldsymbol{ss}}(0) = \boldsymbol{\Phi}_1(\boldsymbol{\theta})\boldsymbol{\Gamma}_{\boldsymbol{ss}}(0)\boldsymbol{\Phi}_1(\boldsymbol{\theta})' + \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta})\boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta})' \tag{2.4}$$

and, for $h \neq 0$, the autocovariances of $\boldsymbol{s}_t$ are computed using

$$\boldsymbol{\Gamma}_{\boldsymbol{ss}}(h) = \boldsymbol{\Phi}_1^h(\boldsymbol{\theta})\boldsymbol{\Gamma}_{\boldsymbol{ss}}(0) \tag{2.5}$$

Furthermore, the unconditional mean $\boldsymbol{\mu} = \mathrm{E}\,\boldsymbol{y}_t$ of $\boldsymbol{y}_t$ is given by

$$\boldsymbol{\mu} = \boldsymbol{\Psi}_0(\boldsymbol{\theta}) \tag{2.6}$$

The linearity of the model and Gaussianity of the exogenous shocks $\boldsymbol{\varepsilon}_t$ imply that the joint probability distribution of a $T \times n_{\boldsymbol{y}}$ vector of observations $\boldsymbol{Y}_T = [\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_T']'$ is also Gaussian. This means that the probability distribution is fully characterized by the first and second order moments of $\boldsymbol{Y}_T$. Hence, all relevant model-implied information about $\boldsymbol{\theta}$ is contained in the mean $\boldsymbol{\mu}$ and the sequence of autocovariances $\boldsymbol{\Gamma}_{\boldsymbol{yy}}(0), \boldsymbol{\Gamma}_{\boldsymbol{yy}}(1), \ldots, \boldsymbol{\Gamma}_{\boldsymbol{yy}}(T-1)$ of $\boldsymbol{y}_t$. Our objective is to determine which of these moments are most useful in terms of the information they provide with respect to different model parameters. To do that we adopt a generalized method of moments framework, where theoretical first and second-order moments are matched to their empirical counterparts. This approach is motivated by the result that a GMM estimator with optimal weighting matrix is asymptotically efficient among all the consistent estimators utilizing the same set of moment conditions. Efficiency implies that the GMM estimator utilizes the included moments in an optimal fashion, maximizing their information content to achieve the lowest possible estimation uncertainty. As we explain next, we use this property to determine the amount of information contributed by different parts of the moment structure.

2.2 GMM and information in the moment structure of $\boldsymbol{Y}_T$

We consider a GMM estimator $\hat{\boldsymbol{\theta}}_T(p)$ which matches the model-implied value of a vector of moments

$$\boldsymbol{m}(\boldsymbol{\theta}, p) = [\boldsymbol{\mu}', \mathrm{vech}(\boldsymbol{\Gamma}_{\boldsymbol{yy}}(0))', \mathrm{vec}(\boldsymbol{\Gamma}_{\boldsymbol{yy}}(1))', ..., \mathrm{vec}(\boldsymbol{\Gamma}_{\boldsymbol{yy}}(p))']',$$

where $p \leq T-1$, to their sample counterparts, collected in the vector $\hat{\boldsymbol{m}}_T(p)$. The estimator is defined as the solution to the following optimization problem:

$$\hat{\boldsymbol{\theta}}_T(\boldsymbol{W}, p) = \underset{\boldsymbol{\theta}}{\mathrm{argmin}} \Big\{ (\boldsymbol{m}(\boldsymbol{\theta}, p) - \hat{\boldsymbol{m}}_T(p))' \, \boldsymbol{W}_T \, (\boldsymbol{m}(\boldsymbol{\theta}, p) - \hat{\boldsymbol{m}}_T(p)) \Big\}, \tag{2.7}$$

where $\boldsymbol{W}_T$ is a positive definite and possibly random weighting matrix converging in probability to a positive definite matrix $\boldsymbol{W}$. Under the regularity condition in Hansen (1982), $\hat{\boldsymbol{\theta}}_T(p)$ is consistent and asymptotically normally distributed with

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T(p) - \boldsymbol{\theta}_0\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \boldsymbol{V}_{\boldsymbol{\theta}}(\boldsymbol{W}, p)\right). \tag{2.8}$$

For a given weighting matrix $\boldsymbol{W}$ the asymptotic covariance matrix is:[7]

$$\boldsymbol{V}_{\boldsymbol{\theta}}(\boldsymbol{W}, p) = (\boldsymbol{M}(p)'\boldsymbol{W}\boldsymbol{M}(p))^{-1}\boldsymbol{M}(p)'\boldsymbol{W}\boldsymbol{V}_{\boldsymbol{m}}(p)\boldsymbol{W}(p)\boldsymbol{M}(p)(\boldsymbol{M}(p)'\boldsymbol{W}\boldsymbol{M}(p))^{-1} \tag{2.9}$$

where $\boldsymbol{M}(p) = \partial \boldsymbol{m}(\boldsymbol{\theta}_0, p)/\partial \boldsymbol{\theta}'$ is the Jacobian matrix of the moment conditions with respect to $\boldsymbol{\theta}$, and $\boldsymbol{V}_{\boldsymbol{m}}(p)$ is the asymptotic covariance matrix of the moment condition,

$$\sqrt{T}\left(\hat{\boldsymbol{m}}_T(p) - \boldsymbol{m}(\boldsymbol{\theta}_0, p)\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \boldsymbol{V}_{\boldsymbol{m}}(p)\right) \tag{2.10}$$

For a given set of moments in $\boldsymbol{m}(\boldsymbol{\theta}, p)$, the efficiency of the GMM estimator depends on the choice of weighting matrix $\boldsymbol{W}$. As shown by Hansen (1982) and Chamberlain

[7]Note that the sensitivity measure of Andrews et al. (2017a) is defined as

$$\Lambda = -(\boldsymbol{M}(p)'\boldsymbol{W}\boldsymbol{M}(p))^{-1}\boldsymbol{M}(p)'\boldsymbol{W}$$

(1987), the GMM estimator is asymptotically efficient among the estimators using the same set of moment conditions when $\boldsymbol{W}$ is equal to the inverse of $\boldsymbol{V_m}(p)$. In that case, the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}_T(p)$ is

$$\boldsymbol{V}^*_{\boldsymbol{\theta}}(p) = \Big(\boldsymbol{M}(p)'\boldsymbol{V_m}(p)^{-1}\boldsymbol{M}(p)\Big)^{-1} \tag{2.11}$$

Asymptotic efficiency implies that, for large T, the estimation uncertainty of any consistent estimator of $\boldsymbol{\theta}$ is greater than or equal to that of $\hat{\boldsymbol{\theta}}_T$. In order to evaluate $\boldsymbol{V}^*_{\boldsymbol{\theta}}(p)$ we need the asymptotic covariance matrix of the sample moments $\boldsymbol{V_m}(p)$. Note that this matrix is block diagonal with blocks $\boldsymbol{V_\mu} = \lim_{T\to\infty} T\,\mathrm{cov}\,(\hat{\boldsymbol{\mu}}_T - \boldsymbol{\mu})$, and $\boldsymbol{V_\gamma}(p) = \lim_{T\to\infty} T\,\mathrm{cov}\,(\hat{\boldsymbol{\gamma}}_T(p) - \boldsymbol{\gamma}(p))$ corresponding to the mean vector $\boldsymbol{\mu}$ and the vector of second order moments $\boldsymbol{\gamma}(p) = [\mathrm{vech}(\boldsymbol{\Gamma_{yy}}(0))', \mathrm{vec}(\boldsymbol{\Gamma_{yy}}(1))', ..., \mathrm{vec}(\boldsymbol{\Gamma_{yy}}(p))']'$. It is well known (see e.g. Fuller (1996)) that $\boldsymbol{V_\mu}$ is equal to 2π times the spectral density of $\boldsymbol{y}_t$ evaluated at frequency zero. Therefore, using (2.1)–(2.2) we have

$$\boldsymbol{V_\mu} = \sum_{h=-\infty}^{\infty} \boldsymbol{\Gamma}(h) = \boldsymbol{\Psi}_1\left(\boldsymbol{I}_m - \boldsymbol{\Phi}_1\right)^{-1}\boldsymbol{\Phi_\varepsilon}\boldsymbol{\Phi}'_{\boldsymbol{\varepsilon}}\left(\boldsymbol{I}_m - \boldsymbol{\Phi}'_1\right)^{-1}\boldsymbol{\Psi}'_1 \tag{2.12}$$

To compute the asymptotic covariance matrix of $\boldsymbol{\gamma}(p)$, we use the multivariate version of the so-called Bartlett formula (see Bartlett (1955)). In particular, Su and Lund (2012) show that, if $\gamma_{i,j}(q)$ is the autocovariance at lag q between the $i-$th and the $j-$th components of $\boldsymbol{y}_t$, then the joint asymptotic distribution of any two sample autocovariances $\hat{\gamma}_{a,b}(q_1)$ and $\hat{\gamma}_{c,d}(q_2)$ is given by

$$\sqrt{T}\left(\begin{pmatrix}\hat{\gamma}_{a,b}(q_1)\\ \hat{\gamma}_{c,d}(q_2)\end{pmatrix} - \begin{pmatrix}\gamma_{a,b}(q_1)\\ \gamma_{c,d}(q_2)\end{pmatrix}\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix}\omega_{q_1,q_1} & \omega_{q_1,q_2}\\ \omega_{q_2,q_1} & \omega_{q_2,q_2}\end{pmatrix}\right), \tag{2.13}$$

where

$$\omega_{k,l} \quad = \quad \sum_{h=-\infty}^{\infty}\Big(\gamma_{a,c}(h)\gamma_{b,d}(h-k+l) + \gamma_{a,d}(h+l)\gamma_{b,c}(h-k)\Big) \tag{2.14}$$

Using (2.13), we can construct the full asymptotic covariance matrix of $\boldsymbol{\gamma}$. Unlike $\boldsymbol{V_\mu}$, which can be evaluated in closed form using (2.12), in general $\boldsymbol{V_\gamma}(p)$ has to be approximated by truncating the infinite sum in (2.14). However, since the terms involved in the summation are very easy to compute, for any lag, with the formula in (2.3), matrix $\boldsymbol{V}_{\boldsymbol{\gamma}}(p)$ can be approximated arbitrarily well using the the expression in (2.13). Note that the expressions in (2.13)–(2.14) are valid only for Gaussian models. In the case of non-Gaussian models one needs to account for the non-zero higher order moments of the distribution. Details on how to compute the asymptotic covariance matrix of $\boldsymbol{\gamma}$ in the case of non-Gaussian models can be found in Su and Lund (2012).

Using these results, we are able to evaluate the asymptotic covariance matrix of the GMM estimator for a given set of moment conditions and any weighting matrix $\boldsymbol{W}$. Suppose that we want to know how much additional information about $\boldsymbol{\theta}$ is contributed by auto and cross-autocovariances at lag $p+1$, given the information contained in the mean and autocovariances at lags from 0 to p. The size of this contribution can be determined by comparing $\boldsymbol{V}_{\boldsymbol{\theta}}^*(p+1)$ to $\boldsymbol{V}_{\boldsymbol{\theta}}^*(p)$. Alternatively (and equivalently), we could compare $\boldsymbol{V}_{\boldsymbol{\theta}}^*(p+1)$ and $\boldsymbol{V}_{\boldsymbol{\theta}}(\boldsymbol{W}, p+1)$, where the weighting matrix $\boldsymbol{W}$ is constructed so as to weight the moments in $\boldsymbol{m}(p)$ optimally – using $\boldsymbol{V}_{\boldsymbol{m}}(p)$, while placing zero weights on the elements of $\operatorname{vec}(\boldsymbol{\Gamma}_{yy}(p+1))$. More generally, starting with a *complete* set of moments, say $\boldsymbol{m}(P)$, we can measure the contribution of information by any subset of moments $\bar{\boldsymbol{m}} \subset \boldsymbol{m}(P)$ by constructing a weighting matrix which weights optimally the moments in the complement of $\bar{\boldsymbol{m}}$, $\bar{\boldsymbol{m}}^c = \boldsymbol{m}(P) \setminus \bar{\boldsymbol{m}}$, and puts zero weights on the ones in $\bar{\boldsymbol{m}}$. The difference between $\boldsymbol{V}_{\boldsymbol{\theta}}(\boldsymbol{W}, P)$ and $\boldsymbol{V}_{\boldsymbol{\theta}}^*(P)$ shows the marginal contribution of information about $\boldsymbol{\theta}$ that is due to $\bar{\boldsymbol{m}}$.

In practice, we are often more interested in the contribution of information by a set of moments with respect to individual elements of $\boldsymbol{\theta}$. Therefore, we define a measure of efficiency gains that uses the diagonal elements of the asymptotic covariance matrices of efficient GMM estimators based on different sets of moments. Specifically, let $\operatorname{std}(\theta_i|\bar{\boldsymbol{m}}^c)$ and $\operatorname{std}(\theta_i|\boldsymbol{m}(P))$ be the square roots of the diagonal elements of $\boldsymbol{V}_{\boldsymbol{\theta}}(\boldsymbol{W}, P)$ and $\boldsymbol{V}_{\boldsymbol{\theta}}^*(P)$, respectively. The efficiency gain (EG) due to $\bar{\boldsymbol{m}}$ with respect to parameter θ_i is defined as

$$\mathrm{EG}_{\theta_i}(\bar{\boldsymbol{m}}|\bar{\boldsymbol{m}}^c) = \left(\frac{\mathrm{std}(\theta_i|\bar{\boldsymbol{m}}^c) - \mathrm{std}(\theta_i|\boldsymbol{m}(P))}{\mathrm{std}(\theta_i|\bar{\boldsymbol{m}}^c)}\right) \times 100. \tag{2.15}$$

Note that (2.15) is a measure of conditional gain, i.e. the increase in efficiency of the GMM estimator of θ_i due to information in $\bar{\boldsymbol{m}}$ given the information already in $\bar{\boldsymbol{m}}^c$.

Using the above approach allows us to compare the efficiency of GMM estimators that use different subsets of moments from the moment structure of $\boldsymbol{Y}_T$. In addition, we may also be interested in comparing the efficiency of the GMM estimator and the full information maximum likelihood (ML) estimator. The ML estimator $\hat{\boldsymbol{\theta}}_T^{ml}$ solves

$$\hat{\boldsymbol{\theta}}_T^{ml} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, \ell(\boldsymbol{\theta}|\boldsymbol{Y}_T) \tag{2.16}$$

where $\ell(\boldsymbol{\theta}|\boldsymbol{Y}_T)$ is the logarithm of the likelihood function of $\boldsymbol{Y}_T$. The asymptotic distribution of $\hat{\boldsymbol{\theta}}_T^{ml}$ is

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T^{ml} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathcal{I}_{\boldsymbol{\theta}}^{-1}\right) \tag{2.17}$$

where $\mathcal{I}_{\boldsymbol{\theta}}$ is the asymptotic Fisher information matrix evaluated at the true value of $\boldsymbol{\theta}$. A key property of the MLE is that it is asymptotically efficient among all consistent estimators of $\boldsymbol{\theta}$. Therefore, the asymptotic covariance matrix of the MLE is a lower bound on the uncertainty of the parameter estimates, which can be reached asymptotically using all available sample information. As Carrasco and Florens (2014) show, under certain conditions, the optimal GMM estimator is asymptotically as efficient as the MLE. This suggests that, as we increase the number of lagged autocovariances used by the optimal GMM estimator, its asymptotic covariance matrix should converge to the covariance matrix of MLE.

3 An Application

In this section, we apply our method to a DSGE model with news shocks estimated in Schmitt-Grohé and Uribe (2012). This is a closed economy real business cycle model augmented with real rigidities in consumption, investment, capital utilization, and wage setting. The details of the model are given in the Appendix. We take as given the estimation setup and results of the original article, i.e. assume the same free parameters and observed variables, and use the estimation results reported there in our analysis.[8] Our objective is show what how our approach can be used to shed light on the question of what features of the data, i.e. moments of the observables, inform the estimates of key parameters of that model.

SGU estimate the model using quarterly U.S. data for the period between 1955:Q2 and 2006:Q4. The variables they use are: the growth rates of per capita real GDP (y_t), real consumption (c_t), real investment (i_t), and real government expenditure (g_t) the growth rates of the relative price of investment (a_t) and of total factor productivity (tfp_t), and hours worked (h_t). All series are demeaned, which implies that only information from second-order moments is used in estimation. Note, however, that this does not lead to a loss of information with respect to the estimated parameters since all parameters for which the first-order moments are informative are calibrated prior to estimation. In other words, the means of the observed variables are redundant with respect to the set of freely estimated parameters. Thus, in the following analysis we consider only moments from the covariance structure of the seven observed variables.

There are 34 free parameters, 6 of which are related to preferences and technology, 7 are autoregressive coefficients of shocks, and the remaining 21 parameters are standard

[8] In our analysis we use the point values from the ML estimation reported in Schmitt-Grohé and Uribe (2012) (see Table II of the article). Using the mean or the mode from the posterior distribution instead does not change the results substantially. These results are available upon request.

deviations of anticipated and unanticipated innovations to shocks.[9] Since the purpose of this section is only to illustrate how our approach works, here we focus on a subset of the estimated parameters. The parameters we consider are (1) standard deviations of the 6 innovations to the permanent $(\sigma^0_{\mu^a}, \sigma^4_{\mu^a}, \sigma^8_{\mu^a})$ and transitory $(\sigma^0_{z^I}, \sigma^4_{z^I}, \sigma^8_{z^I})$ investment-specific productivity shocks, and (2) all parameters that are not related to shocks – Frisch elasticity of labor supply (θ), wealth elasticity of labor supply (γ), investment adjustment cost (κ), capacity utilization cost (δ_2/δ_1), habit (b), smoothness of trend in government spending (ρ_{xg}). Results about the remaining parameters are presented in an on-line appendix.

In general, one could divide the moment structure into many different groups of moments. In our analysis we consider four such groups: moments at different lags, moments of different variables, cross-moments of different pairs of variables, and individual moments.

3.1 Lag structure

Here we analyze how information about $\boldsymbol{\theta}$ accumulates as we add more lagged autocovariances to the set of included moments. Our purpose is two-fold. First, to determine whether certain lagged autocovariances play a crucial role with respect to some parameters, in particular those of the news shocks. Second, to compare the efficiency of the optimal GMM estimator to that of the ML estimator, which, as noted earlier, achieves the asymptotic Cramér-Rao (CR) bound. We do that by computing the asymptotic standard deviations of the efficient GMM estimators and of the ML estimator using the results of Section 2. The relative efficiency of GMM with a given number of lagged autocovariance and MLE is measured with the ratio of the two standard deviations. Figure 1 present results for the investment-specific productivity shock parameters. The

[9] In addition to the structural parameters SGU estimate one measurement error parameter.

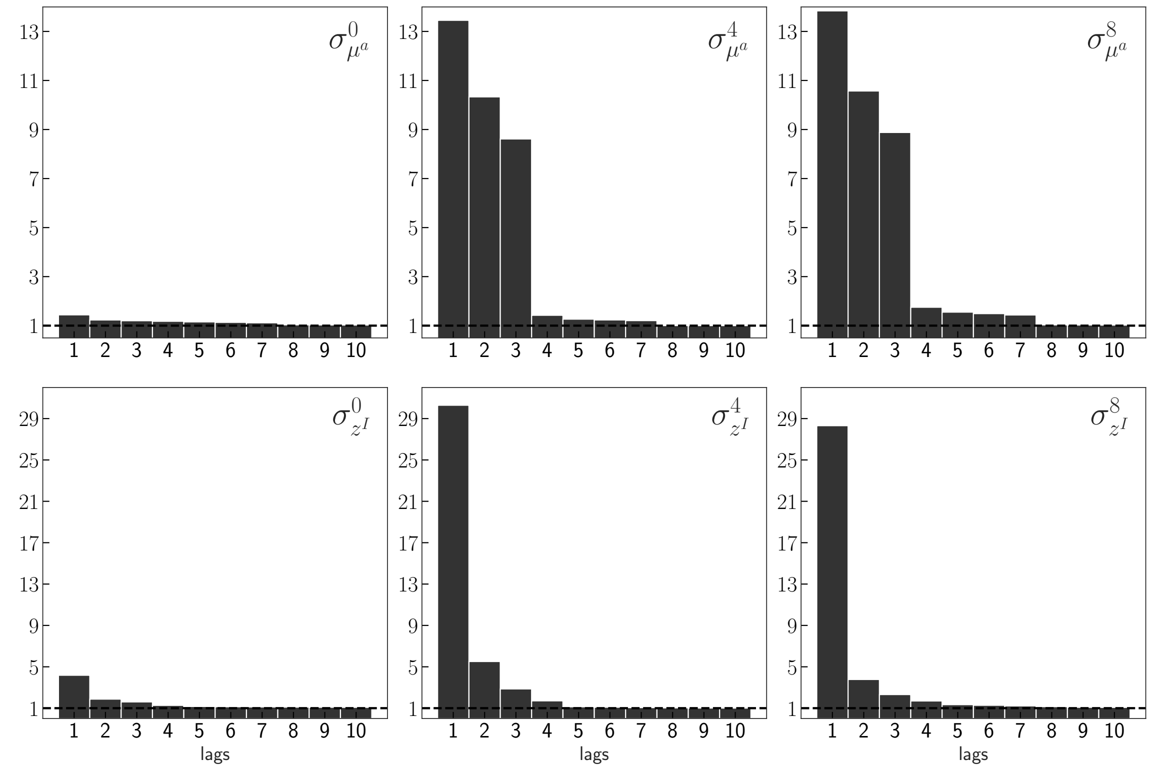

Figure 1: Investment-specific productivity shocks parameters. The figure shows the asymptotic efficiency of GMM estimator for different number of lags relative to MLE.

bars in each bar plot show the ratios of the asymptotic standard deviations of the efficient GMM estimator based on all covariances and autocovariances up to lags from 1 to 10, divided by the asymptotic standard deviation of the MLE.[10] Several results are worth noting: First, as more lags are added the standard deviations of the GMM estimator decrease monotonically towards the standard deviations of MLE. This is not a general property of GMM estimators, i.e. one that holds for any choice of a weighting matrix. As explained in Meng and Xie (2014) adding more information does not guarantee that the estimation uncertainty will decrease, unless the information is used in an optimal

[10]In principle, we could also consider a GMM estimator which uses only the contemporaneous covariances. However, this gives us only 28 moment conditions with which we cannot identify the 34 free parameters. For identification of all free parameters it is sufficient to use the covariance and first order autocovariance matrices of the observed variables.

fashion, as in the case of MLE or GMM estimators with optimal weighting matrices.[11] In the case of the μ^a shock parameters, by lag 10 the GMM standard deviations are about 2% or less larger than their CR bounds. In the case of the z^I shock the GMM standard deviations are about 8% or less larger than the CR bounds. Second, it takes a larger number of lags for the GMM standard deviations of the news shocks parameters to decline substantially and come close to the respective CR bonds. For instance, by lag 3 the GMM standard deviations of $\sigma^0_{\mu^a}$ and $\sigma^0_{z^I}$ are about 20% and 60% larger than the CR bonds, while in the case of $(\sigma^4_{\mu^a}, \sigma^8_{\mu^a})$ and $(\sigma^4_{z^I}, \sigma^8_{z^I})$ the standard deviations are more than 760% and 130% greater than the respective bounds. Third, there is a significant and distinct impact of adding autocovariances at lags 4 and, to a lesser extent, 8 on the GMM standard deviations of the μ^a news shock parameters but not in the case of the z^I news parameters. For both $\sigma^4_{\mu^a}$ and $\sigma^8_{\mu^a}$ the standard deviations drop by more than 80% when autocovariances up to lag 4 are used, compared to using only autocovariances up to lag 3. Increasing the number of lags from 7 to 8 results in a decline by 16% and 28%, for $\sigma^4_{\mu^a}$ and $\sigma^8_{\mu^a}$, respectively. Although much smaller, the effect of using 8 instead of 7 lags is notable as it is larger than the cumulative impact of adding lags 5, 6 and 7, compared to using only lags up to 4. In the case of the z^I news shock parameters, increasing the number of lags from 3 to 4 reduces the GMM standard deviations by 40% (for $\sigma^4_{z^I}$) and 27% (for $\sigma^8_{z^I}$), while increasing the number of lags from 7 to 8 reduces the standard deviations by 2% (for $\sigma^4_{z^I}$) and 6% (for $\sigma^8_{z^I}$).

Figure 2 shows results for the 6 structural parameters that are not related to shocks. Again, as more lagged autocovariances are used, the efficient GMM estimator converges to MLE in terms of efficiency. However, in the case of θ and especially δ_2/δ_1 the convergence is quite slow. By lag 10 the GMM standard deviation of θ is more than 30% larger than

[11] In addition, as noted earlier, convergence in efficiency to MLE occurs only under certain conditions, namely that the true score belongs to the closure of the linear space spanned by the moment conditions of the GMM estimator, see Carrasco and Florens (2014).

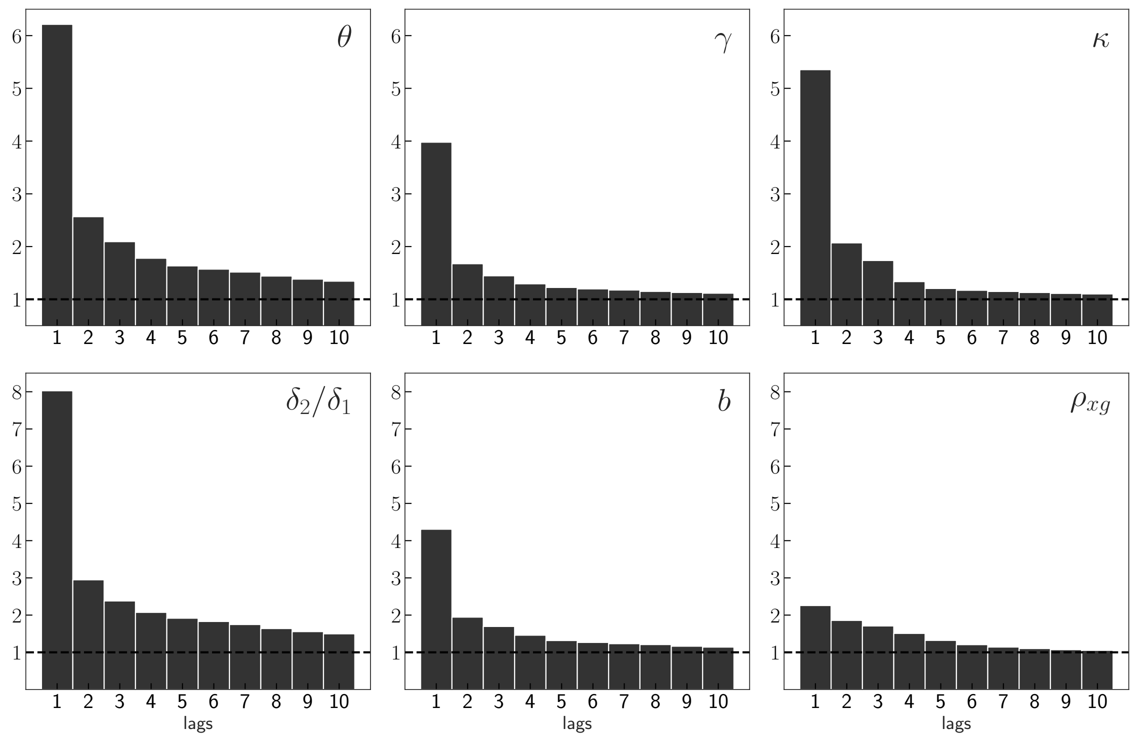

Figure 2: Structural parameters. See note to Figure 1.

the MLE standard deviation, and in the case of δ_2/δ_1 the relative inefficiency is almost 50%. For the remaining 4 parameters the GMM standard deviations are between 14% and 5% larger than the respective MLE standard deviations. The largest in relative terms drop in estimation uncertainty in all cases except ρ_{xg} occurs when the set of autocovariances at lag 2 is added to the moment conditions of GMM estimator.

3.2 Observed variables

Next, we consider the contribution of information by moments of each observed variables. There are two ways to define the information content of a variable: (1) as information only from the own moments of that variable, e.g. the variance and all autocovariances of y_t, and (2) as information from all moments of a variable, e.g. the variance, autocovariances, and all covariances and cross-autocovariances between y_t and the other 6 variables. As

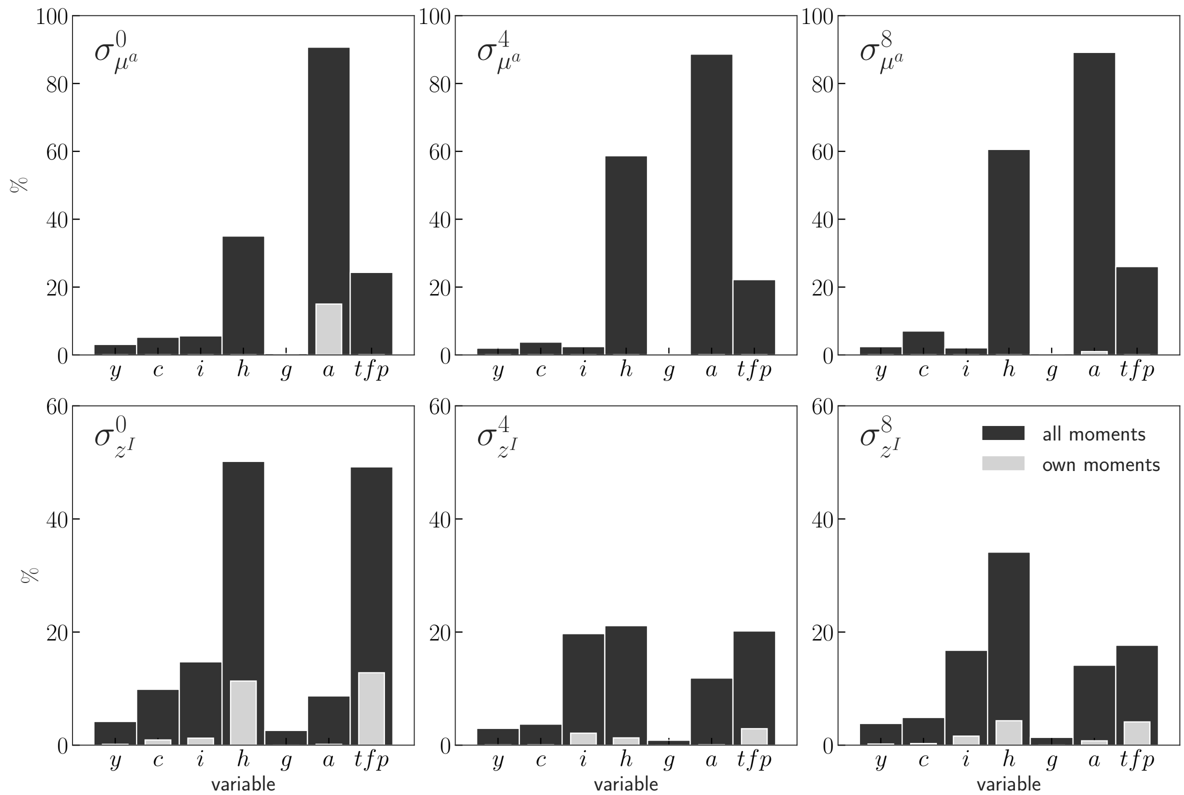

Figure 3: Nonstationary (μ^a) and stationary (z^I) investment-specific productivity shocks parameters. The figure shows the efficiency gains due to moments of each observed variable.

explained in Section 2.2, we measure the amount of information contributed by the (own or all) moments of a variable as the per cent reduction in estimation uncertainty, i.e. the asymptotic standard deviation of the efficient GMM estimator, as a result of including these moments in the set of moment conditions. Thus, we compare the asymptotic covariance matrix of a GMM estimator based on a subset of moments to the covariance matrix of an estimator that uses the complete set of moments. To find the latter, we use the asymptotic covariance matrix of an efficient GMM estimator using all variances and autocovariances up to lag 100. It is unnecessary to include autocovariance at a higher order as the values of the standard deviations with up to 100 lags are already very close to the MLE asymptotic standard deviations.

Figure 3 presents results for the parameters of the investment-specific productivity shocks. The bars in grey represent the efficiency gains due to the own moments of each variable; the bars in black show the gains from all moments. The upper part of the figure shows clearly that the relative price of investment (a_t) is a paramount source of unique information with respect to the μ^a shock parameters. The efficiency gains are about 90%. Among the other variables, hours (h_t) and to a lesser extent tfp_t also play significant role. Note that bars with heights close to zero indicate that the moments of the respective variable contain little unique information, and therefore adding them to the moment conditions results in only small increase of estimation precision. Thus, the results suggest that there is relatively little unique information with respect to the μ^a shock parameters in all moments of y_t, c_t, i_t, and g_t, and in the own moments of all variables, including a_t.[12] From the lower part of the figure we see that there is not a single variable whose moments are as important for the z^I shock parameters as a_t is with respect to the μ^a parameters. The most significant sources of unique information are the moments of h_t and tfp_t, especially with respect to $\sigma^0_{z^I}$, while i_t and a_t are about as important in terms of the parameters of the news components of that shock.

Figure 4 shows results for the 6 non-shock parameters. Again, we see that h_t, and to a lesser extent tfp_t, are the main sources of unique information with respect to 5 of these parameters. The only exception is ρ_{xg} for which government spending (g_t) replaces h_t in terms of importance. The moments of h_t are particularly important for θ and γ, whose asymptotic standard deviations are reduced by between 70% and 80% due to information in those moments. Among the other variables, only c_t plays a non-negligible role for most parameters.

One robust conclusion we can draw from the results in Figures 3 and 4 is that the bulk of the unique information about the parameters we analyze comes from cross-moments,

[12] There is about 15% efficiency gain with respect to $\sigma^0_{\mu^a}$ from including the own moments of a_t, and between 4% and 7% gain due to all moments of c_t with respect to the μ^a parameters.

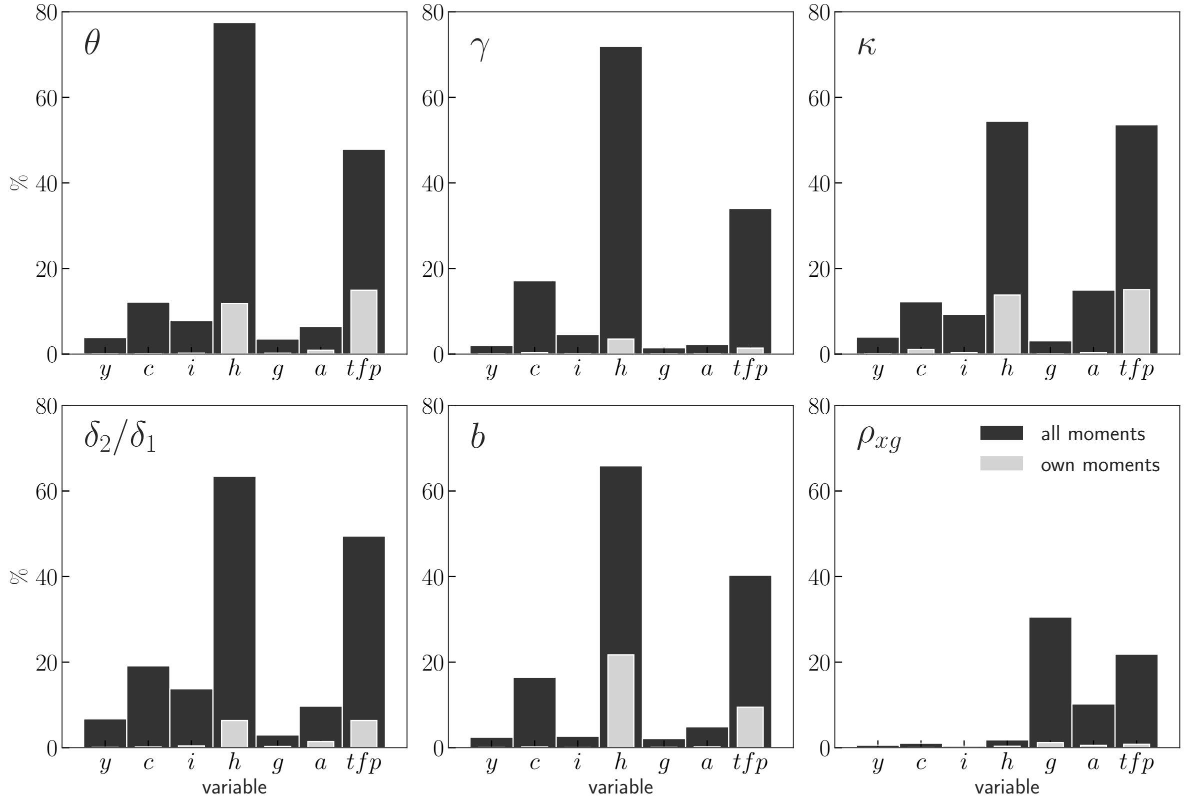

Figure 4: Structural parameters. See note to Figure 3.

and not the own moments of individual variables. We next ask which particular groups of cross-moments are the most significant sources of information.

3.3 Main groups of covariances

Since we have seven observed variables, the covariance structure of $\boldsymbol{Y}_T$ consists of 21 main groups of covariances, which we define as follows

$$\mathrm{cov}(z_1, z_2) = [\mathrm{cov}(z_{1,t-T+1}, z_{2,t}), \ldots, \mathrm{cov}(z_{1,t}, z_{2,t}), \ldots, \mathrm{cov}(z_{1,t}, z_{2,t-T+1})],$$

for $z_1, z_2 \in \{y, c, i, g, a, tfp, h\}$, $z_1 \neq z_2$. For example, all cross-(auto) covariances between y and c are collected in $\mathrm{cov}(y, c)$. As in the previous section, the contribution of information by a group of covariances is measured as the per cent reduction in the

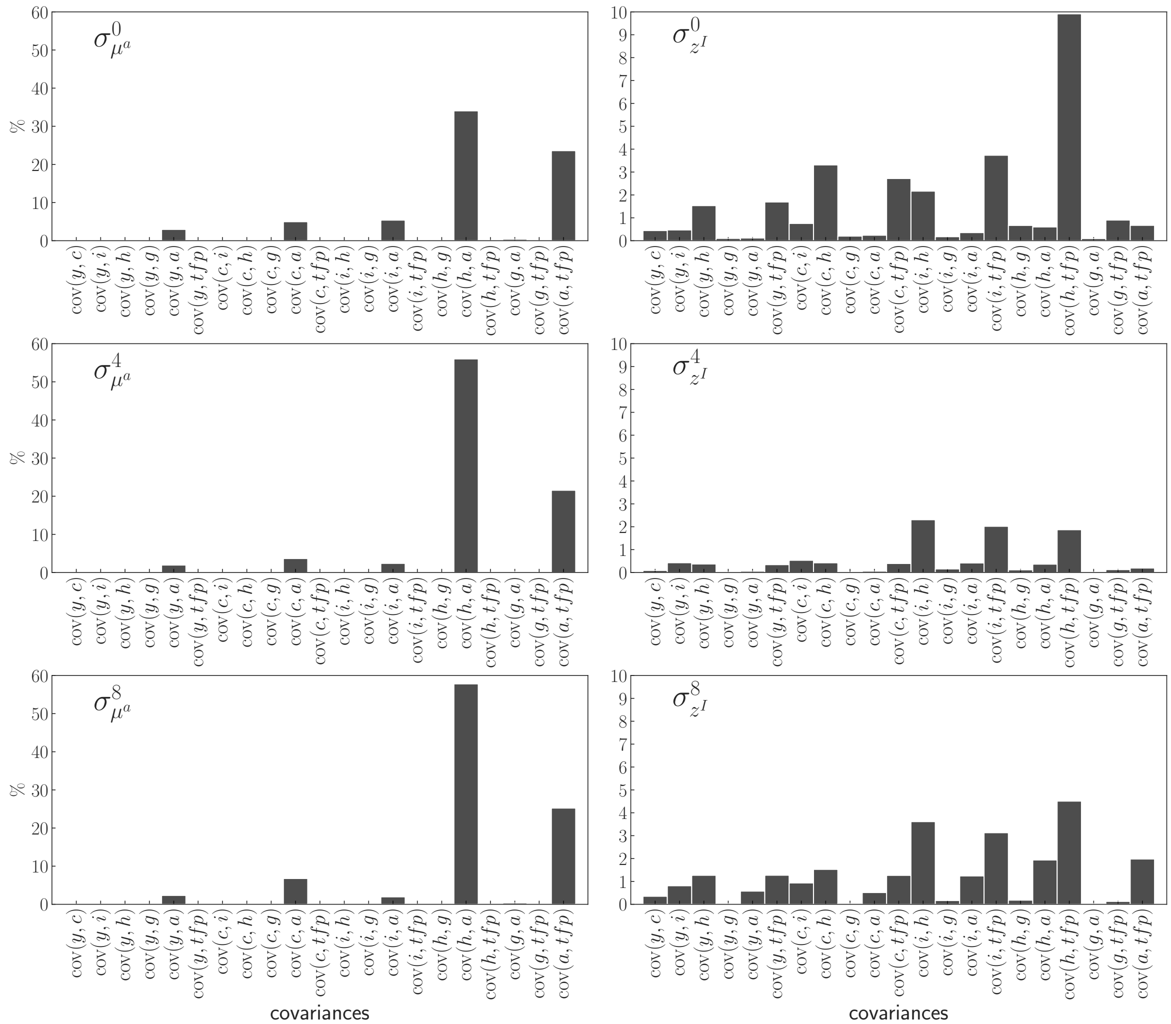

Figure 5: Nonstationary (μ^a) and stationary (z^I) investment-specific productivity shocks parameters. The figure shows the efficiency gains due the covariances in each group.

asymptotic standard deviations of the optimal GMM estimator when the moments of that group are included in the set of moment conditions, relative to the asymptotic standard deviations without those moments.

Figure 5 presents results for the investment-specific productivity shocks parameters. We see that virtually all information about the μ^a shock parameters comes from covariances of a_t and other variables, and most of it is from either the covariances between a_t and h_t, or the covariances between a_t and tfp_t. The covariances between a_t and h_t, in particular, are extremely informative with respect to the standard deviations of the anticipated innovations to that shock. There is close to 60% reduction in the estimation uncertainty of $\sigma^4_{\mu^a}$ and $\sigma^8_{\mu^a}$ due to the moments in $\text{cov}(h, a)$, compared to about 20% reduction due to $\text{cov}(tfp, a)$. The cross-moments of h_t and tfp_t are the single most important source of information about the parameter of the unanticipated innovations to the z^I shock, followed by covariances of these two variables with i_t and c_t. Similarly, the moments in $\text{cov}(h, tfp)$, $\text{cov}(i, tfp)$, and $\text{cov}(i, h)$ provide the largest contributions of information with respect to the parameters of the anticipated innovations to the z^I shock. However, consistent with the earlier observations, the information about the parameters of that shock is much more dispersed across different moments compared to the μ^a shock.

Figure 6 presents results for the 6 parameters unrelated to shocks. The covariances between h_t and tfp_t are the most informative group of cross-moments for four of these parameters, namely θ, κ, δ_2/δ_1, and b. In addition, the cross-moments of h_t and tfp_t with c are also a significant source of information about b, while larger number of moments, including the covariances between h_t and tfp_t with c_t, i_t, and a_t provide about as much unique information about δ_2/δ_1 as $\text{cov}(h, tfp)$. For the remaining two parameters, the moments in $\text{cov}(c, h)$ are by far the most important ones for γ, while $\text{cov}(g, tfp)$ and $\text{cov}(g, a)$ are the most informative groups of moments with respect to ρ_{xg}.

The results in this section show that in most cases there are only a few groups of cross-moments which contribute the bulk of information with respect to estimated parameters. Next, we examine whether there are particular individual moments which on their own contribute significant portions of that information.

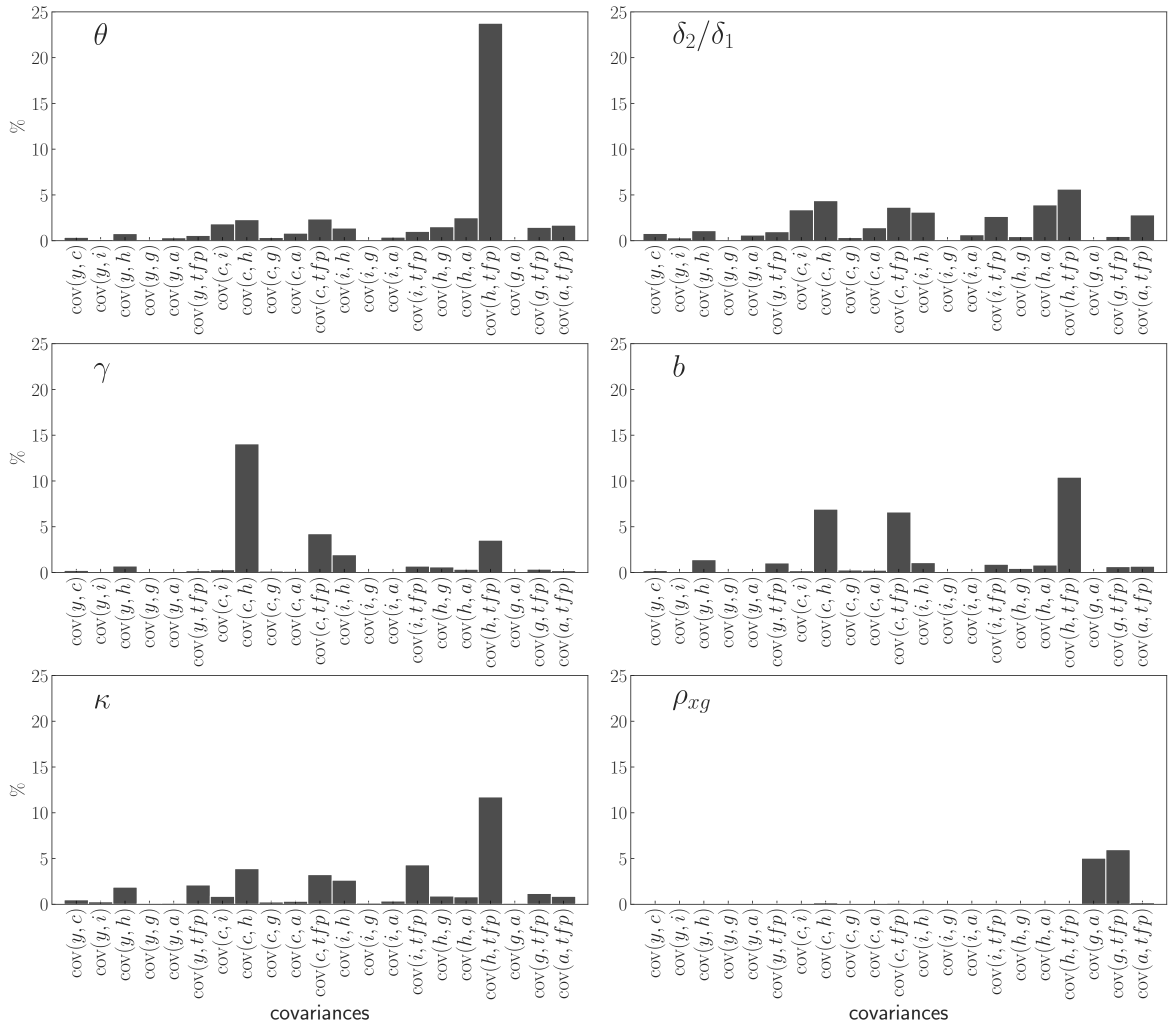

Figure 6: Structural parameters. See note to Figure 5.

3.4 Individual moments

So far we have been comparing information contributed by different subsets of moments. The purpose of this section is to determine which individual moments are most informative with respect to the parameters we study. Note that with T observations, the covariance structure of our model contains $49 \times T - 21$ individual variances, covariances, and cross auto-covariances. To find the most informative ones, we evaluate, in the same way as before, the contribution of each moment and present in Figures 7 and 8 the top 10 most informative moments for each parameter.

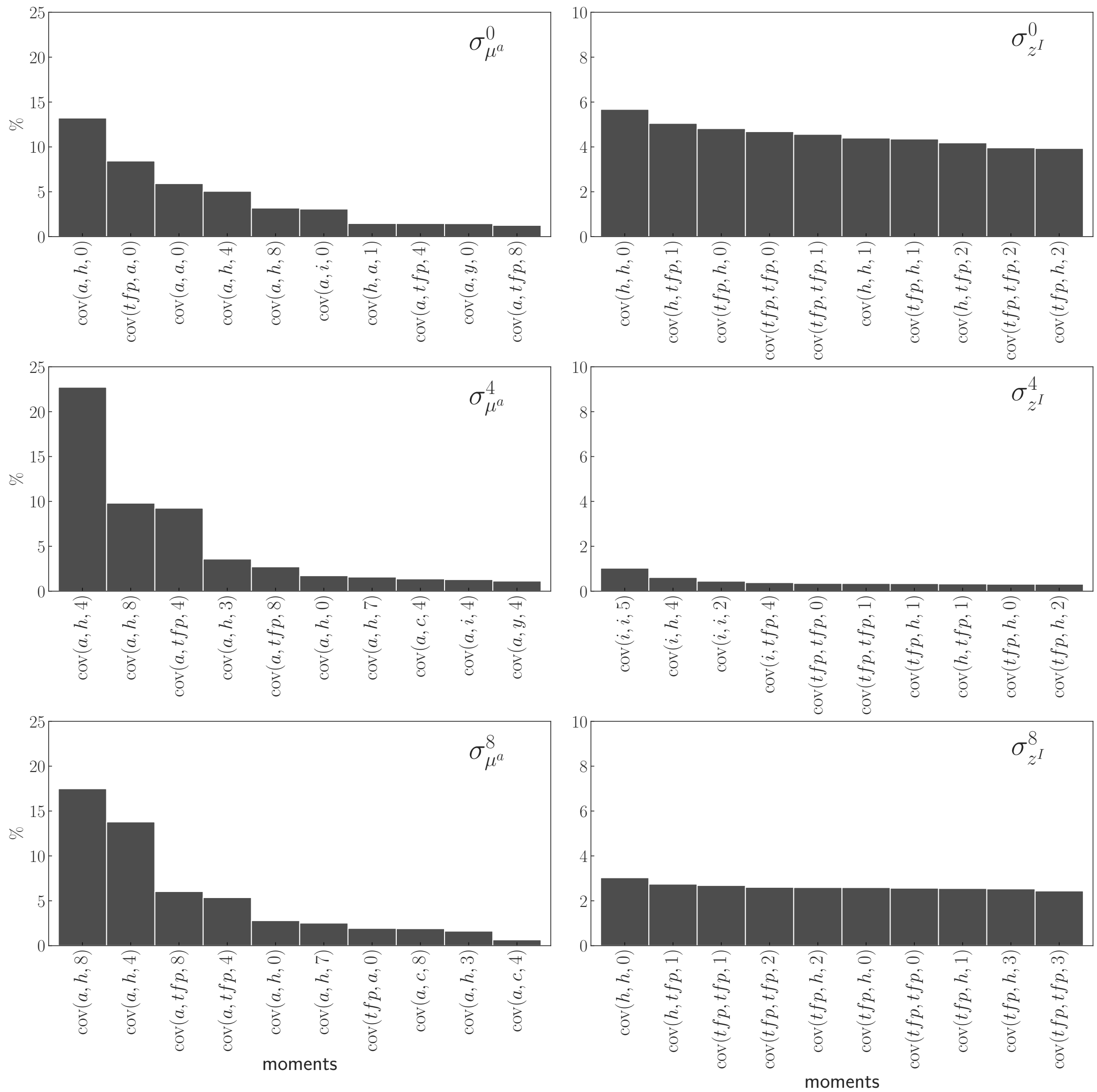

Figure 7: Nonstationary (μ^a) and stationary (z^I) investment-specific productivity shocks parameters. The figure shows the 10 most informative moments with respect to each parameter.

Starting with the μ^a shock parameters, we see in Figure 7 that there is a clear relationship between the type of innovation on one hand, and the order of the most informative autocovariances. In particular, the most informative second moment with

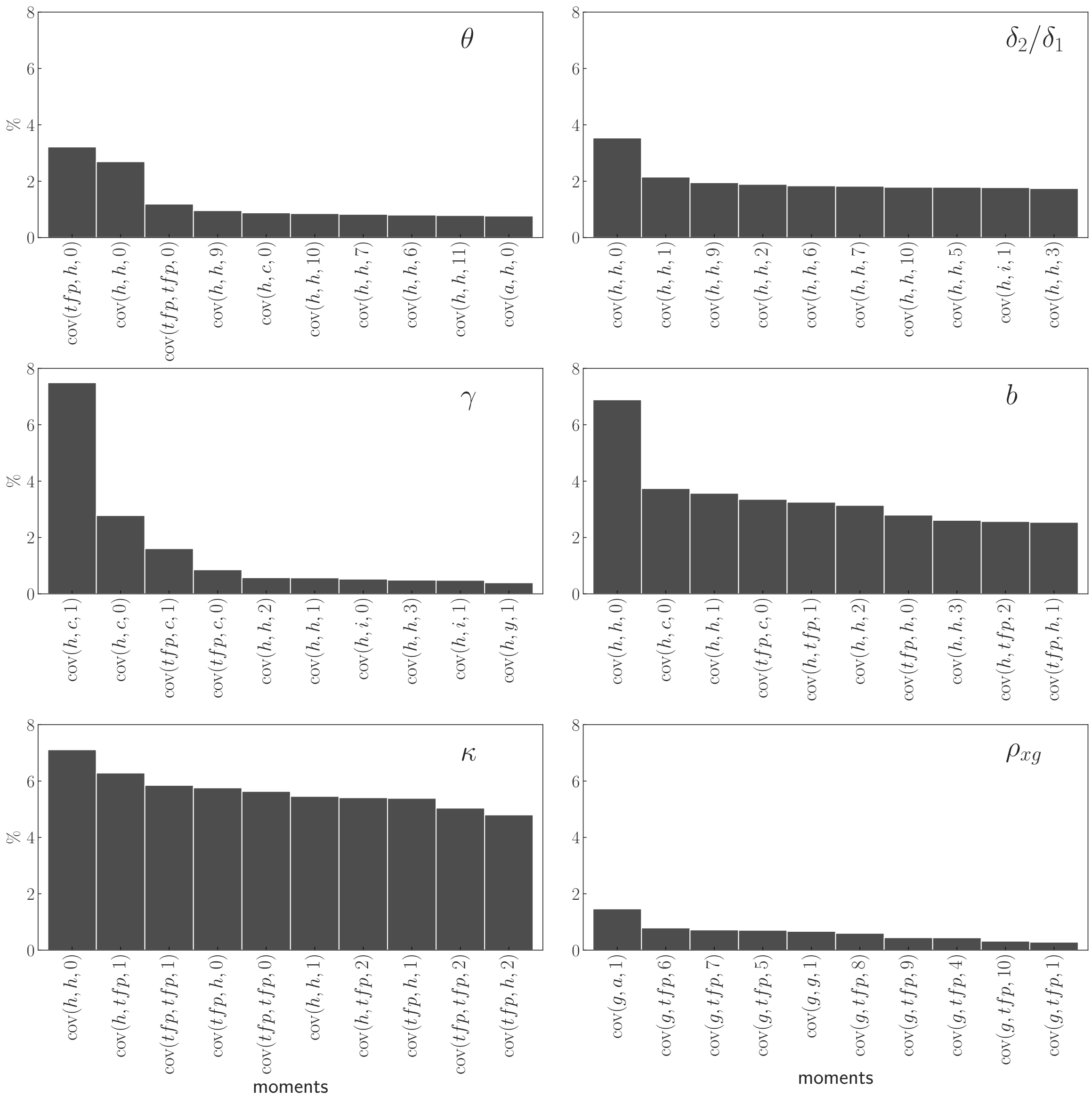

Figure 8: Structural parameters. See the note to Figure 7

respect to the unanticipated shock parameter is $\text{cov}(a_t, h_t)$, while $\text{cov}(a_t, h_{t-4})$ and $\text{cov}(a_t, h_{t-8})$ contribute the largest amounts of information about $\sigma^4_{\mu^a}$ and $\sigma^8_{\mu^a}$, respectively. These covariances alone contribute between 13% and 23% of the unique information about the μ^a shock parameters. Furthermore, the first three most important moments in the case of $\sigma^0_{\mu^a}$ are all contemporaneous covariances – the variance of a_t and the covariances of a_t with h_t and tfp_t, while the three most informative moments with respect to the news shock parameters are cross-autocovariances of a_t with h_t and tfp_t of order either 4 or 8. Interestingly, we do not find a similarly well-defined relationship in the case of the z^I shock parameters. The variance of h_t and the first order autocovariance between h_t and tfp_t rank as the first and second most informative moments both for $\sigma^0_{z^I}$ and $\sigma^8_{z^I}$. In the case of $\sigma^4_{z^I}$, the first two moments are the autocovariance of order 5 of i_t and the corss-autocovariance of order 4 between i_t and h_t. Furthermore, the contributions of individual moments are much smaller and more uniform in size for the parameters of the z^I shock compared to the μ^a shock. In particular, the largest contribution of individual moments – of $\text{cov}(h_t, h_t)$ with respect to $\sigma^0_{z^I}$, is less than 6%, while the most informative moment with respect to $\sigma^4_{z^I}$ contributes only about 1%.

The most informative second moments with respect to the 6 non-shock parameters are displayed in Figure 8. Again, (cross) moments of h_t tend to be among the moments that contribute the most information. Specifically, the contemporaneous covariance between h_t and tfp_t and the first order cross-autocovariance between h_t and c_t are the moments with largest contributions with respect to θ and γ, while the variance of h_t is the most informative moment for κ, δ_2/δ_1, and b. Lastly, in the case of ρ_{xg} the moment with largest contribution of information is the first order cross-autocovariance between a_t and g_t. Similar to Figure 7, there is a considerable variation in the amounts of information contributed by the most informative moments. In particular, note that the contribution of $\text{cov}(h_t, c_{t-1})$ – the most informative moment with respect to γ, is more than twice as large as the contribution of the second most informative moment, and more than seven times as large as the contribution of the fourth most informative moment. At the same time, in the case of κ the size of the contributions decreases very slowly, from close to 7% to about 5% by the tenth most informative moment.

3.5 Discussion of the results

Several of our results merit further discussion. One surprising finding is the relatively minor role variables like investment or consumption seem to play as sources of information.[13] In particular, one might expect investment to be a major source of information for parameters such as the investment adjustment cost parameter (κ), or the standard deviations of the innovations to the investment-specific productivity shock. Similarly, consumption can be expected to be the main source of information for the consumption habit parameter (b). The reason why we do not observed this in our results is that the information these two variables provide is to a large extent not unique. That is, most of the information in either investment or consumption is also contained in the remaining 6 variables, when all of them are observed together. Consequently, observing or not investment, for instance, has a relatively minor impact on the total amount of information. One way to confirm this is to evaluate the contribution of investment when consumption is not observed. Figure 9 shows the contributions of the remaining 6 variables with respect to the parameters of the z^I shock. Investment is now the most informative variable for these parameters, with efficiency gains between 60% and 80%. When consumption is not among the observables, investment becomes also the most informative variable with respect to the investment adjustment cost parameter, with efficiency gain of about 70%. Similarly, we find that if investment is not observed, consumption becomes the most informative variable with respect to the habit persistence parameter, with efficiency gain of more than 70%.[14] Note that similar results obtain if other variables, in particular output or government spending, are excluded instead of

[13] As can be seen in the appendix, the moments of consumption are a major source of information with respect to the parameters of the preference shock, and especially the standard deviations of the news components of that shock.

[14] It should be noted that hours remain very informative with respect to b, with efficiency gain only a little smaller than that of consumption. These results for κ and b are available upon request.

consumption or investment. The reason why little information is lost when any one of these four variables is removed can be traced to the resource constraint in the model, which implies a tight relationship among these variables.

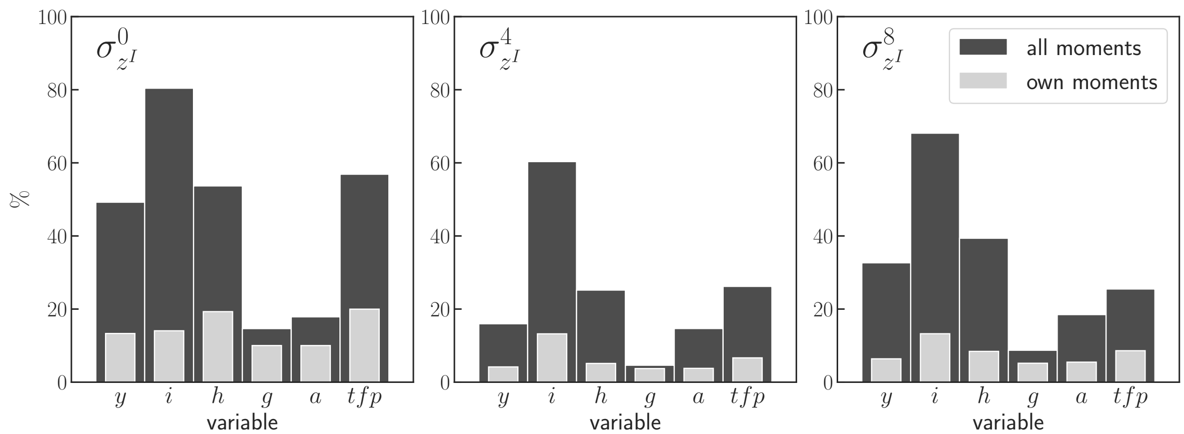

Figure 9: Stationary investment-specific productivity shocks parameters. The figure shows the efficiency gains due the moments of each observed variable.

The fact that information from variables like investment or consumption is, to a certain degree, redundant, given the information from the other observables, could also explain some of the differences in the results for the μ^a shock parameters, on one hand, and most of other parameters, on the other. The relative price of investment, as well as h_t and tfp_t, contain mostly non-redundant information. Because of that, the efficiency gains from moments of a_t and the covariances of a_t with h_t and tfp_t are much larger than the efficiency gains with respect to other parameters. In contrast, information about the z^I shock parameters is dispersed among covariances of h_t and tfp_t with variables other than a_t, each contributing relatively small amount of unique information. The same holds for other shock parameters, e.g. the neutral productivity and the markup shock parameters, the results for which are presented in the appendix. In both cases the main sources of information are either covariances of tfp_t or tfp_t and h_t with variables other than a_t. Consequently, the efficiency gains from groups of covariances or individual moments are relatively small.

Another result worth highlighting is the overall importance of hours worked and TFP as sources of information. Moments of either h_t or tfp_t are the most significant source of information in the case of 21 of the 34 free model parameters, often with the other variable being the second most important one. In addition to θ, γ, κ, and δ_2/δ_1, and the z^I shock parameters, the moments of tfp_t are the main source of information with respect to the neutral productivity shock parameters, and especially the standard deviations of the non-stationary component of that shock. Similarly, h_t is the most informative variable with respect to the parameters of the wage-markup shock.

4 Conclusion

Modern DSGE models are often too large and complex to permit researchers to determine the most informative features of the data by reasoning alone. This paper has proposed a new method for identifying the main sources of parameter information in the moment structure of DSGE models. The method is based on comparing the relative contribution of information by different groups of moments of variables used in estimation. Our measure of information is derived using the asymptotic distribution of the efficient GMM estimator and is easy to compute even for large-scale models. We have demonstrated the usefulness of our approach with an application to a news-driven DSGE model estimated with US data. We believe that conducting and reporting the results from such analysis would benefit both researchers estimating DSGE models and readers of such research, by improving their understanding and increasing the transparency of estimated DSGE models.

References

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

Andrews, M., O. Elamin, A. R. Hall, K. Kyriakoulis, and M. Sutton (2017b): "Inference in the presence of redundant moment conditions and the impact of government health expenditure on health outcomes in England," *Econometric Reviews*, 36, 23–41.

Bartlett, M. (1955): *An introduction to stochastic processes: with special reference to methods and applications*, University Press.

Blanchard, O. (2017): "Do DSGE Models Have a Future?" in *DSGE Models in the Conduct of Policy: Use as intended*, ed. by R. Gürkaynak and C. Tille, CEPR Press, 93–100.

Breusch, T., H. Qian, P. Schmidt, and D. Wyhowski (1999): "Redundancy of moment conditions," *Journal of econometrics*, 91, 89–111.

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

Carrasco, M. and J.-P. Florens (2014): "On the asymptotic efficiency of GMM," *Econometric Theory*, 30, 372–406.

Chamberlain, G. (1987): "Asymptotic efficiency in estimation with conditional moment restrictions," *Journal of Econometrics*, 34, 305–334.

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

Fuller, W. A. (1996): *Introduction to Statistical Time Series*, vol. 230, John Wiley & Sons.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–54, available at http://ideas.repec.org/a/ecm/emetrp/v50y1982i4p1029-54.html.

Inoue, A., C.-H. Kuo, and B. Rossi (2019): "Identifying the sources of model misspecification," *Journal of Monetary Economics, forthcoming.*

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

Komunjer, I. and S. Ng (2011): “Dynamic Identification of Dynamic Stochastic General Equilibrium Models,” *Econometrica*, 79, 1995–2032.

Meng, X.-L. and X. Xie (2014): “I got more data, my model is more refined, but my estimator is getting worse! Am I just dumb?” *Econometric Reviews*, 33, 218–250.

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.

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

Romer, P. (2016): “The trouble with macroeconomics,” *The American Economist*, 20, 1–20.

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

Schorfheide, F. (2008): “DSGE model-based estimation of the New Keynesian Phillips curve,” *Economic Quarterly*, 397–433.

Su, N. and R. Lund (2012): “Multivariate versions of Bartlettś formula,” *Journal of Multivariate Analysis*, 105, 18–31.