PARAMETER IDENTIFICATION IN DYNAMIC ECONOMIC MODELS*

*Nikolay Iskrev***

1. INTRODUCTION

Parameter identification is a concept which every student of economics learns in their introductory econometrics class. The usual textbook treatment of identification leads one to think of identification as a technical issue relevant only to empirical work, and to regard identification problems as caused by either deficiencies of the available data, or of the statistical methodology used to estimate the models. In this note I will argue that the analysis of identification has an important economic modeling aspect, and that it may be very useful to researchers who are not interested in estimation. I will focus the discussion on the class of dynamic stochastic general equilibrium (DSGE) models which have become one of the main analytical tools of modern macroeconomics. The essence of my argument is that when the economic model supplies a complete characterization of the data generating process, parameter identification may be treated as a property of the underlying theoretical model. Parameters will be unidentifiable or weakly identified if the economic features they represent have no empirical relevance at all, or very little of it. This may occur either because those features are unimportant on their own, or because they are redundant given the other features represented in the model. These issues are particularly relevant to DSGE models, which are sometimes criticized of being too rich in features, and possibly overparameterized (Chari, Kehoe, and McGrattan, 2009).

A second reason why it is important to study identification is its econometric implications. The reliable estimation of a model is impossible unless its parameters are well identified. Again, this is crucial for DSGE models as their use for quantitative policy analysis often hinges upon having accurate parameter estimates.

Treating parameter identification as a property of the model means that we can study it without a reference to a particular data set. Such an a priori approach to identification is not always possible in econometrics since typically the relationship between the economic model and the observed data is known only partially. For instance, the degree of correlation between instruments and endogenous variables in the simple linear instrumental variables model depends on nuisance parameters which, in the absence of a fully-articulated economic model, have no structural interpretation. In contrast, when we are in a general equilibrium setting, as in the case of DSGE models, all reduced-form parameters become functions of structural parameters. In this setting we can study how the instruments' strength is determined by the properties of the underlying model.

* The author thanks the comments of João Sousa. The opinions expressed in the article are those of the author and do not necessarily coincide with those of Banco de Portugal or the Eurosystem. Any errors and omissions are the sole responsibility of the author.

** Banco de Portugal, Economics and Research Department.

In what follows I will use three examples, one purely statistical and two simple DSGE models, to illustrate the a priori analysis of identification and the kind of questions we can answer with its help. The presentation here is based on several papers: in Iskrev (2010a) it is explained how to determine if the parameters of a DSGE model are identified; Iskrev (2010a) shows how to evaluate the strength of identification of identified parameters; Iskrev (2010b) discusses the role of observables in the estimation of DSGE models.

2. A SIMPLE EXAMPLE

In this Section I use a simple model to discuss the problem of identification and to explain the main idea behind the *a priori* approach to identification analysis.

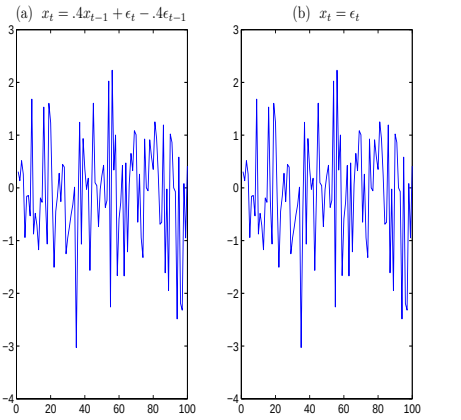
Consider the following autoregressive moving average (ARMA(1,1)) process:

$$x_t = \phi_1 x_{t-1} + \varepsilon_t - \phi_2 \varepsilon_{t-1}, \quad |\phi_1| < 1, |\phi_2| < 1, \quad \varepsilon_t \sim \mathbb{N}(0, \sigma^2) \tag{2.1}$$

Panel (a) of Chart 1 shows 100 observations generated by (2.1) with $\phi_1 = \phi_2 = 0.4,\ \sigma = 1$. Panel (b) shows the realizations of $\varepsilon_t,\ t = 1,...,T$ used to generate the observations for x_t. The two series x_t and ε_t are identical.

Chart 1

OBSERVATIONAL EQUIVALENCE WITH AN ARMA(1,1) PROCESS



Source: Author's calculations.

This example illustrates what in econometrics is called observational equivalence: there are two values of the vector of parameters $\boldsymbol{\theta} = [\phi_1, \phi_2, \sigma]'$, $\boldsymbol{\theta}_1 = [.4, .4, 1]'$ and $\boldsymbol{\theta}_2 = [0, 0, 1]'$, which can produce the same observations for x_t. In fact, in the ARMA(1,1) model there are infinitely many such values; as long as σ is kept the same, and ϕ_1 is equal to ϕ_2, the realizations of x_t would be indistinguishable from those of ε_t.

The reason for this observational equivalence is easy to understand if we consider the autocovari-

ance function (ACF), which for an ARMA(1,1) process is given by:

$$\gamma_0 = (x_t^2) = \frac{(1+\phi_2^2 - 2\phi_1\phi_2)\sigma^2}{1-\phi_1^2}$$

$$\gamma_1 = (x_t x_{t-1}) = \frac{(\phi_1 - \phi_2)(1-\phi_1\phi_2)\sigma^2}{1-\phi_1^2} \qquad (2.2)$$

$$\gamma_h = (x_t x_{t-h}) = \phi_1 \gamma_{h-1}, \quad h \geq 2$$

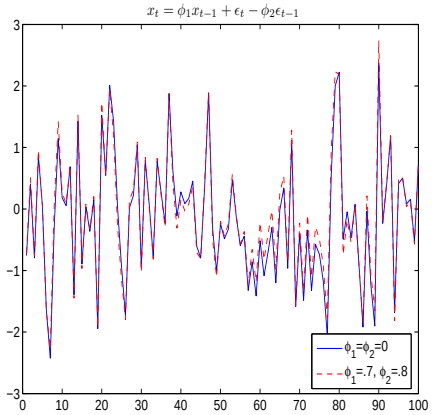
From the definition it is clear that $\phi_1 = \phi_2$ is equivalent to $\gamma_0 = \sigma^2,\ \gamma_k = 0,\ k \geq 1$. Therefore, when the autoregressive and moving average coefficients are equal, the ACF of the ARMA(1,1) process x_t is identical to that of the white noise process ε_t. This implies that we cannot distinguish data generated from ARMA(1,1) process with arbitrary $\phi_1 = \phi_2$ from data generated from ARMA(1,1) process with $\phi_1 = \phi_2 = 0$.

Now consider Chart 2, which shows two series of 100 observations generated by (2.1) with $\boldsymbol{\theta}_1 = [0,0,1]'$ (solid line) and $\boldsymbol{\theta}_2 = [.7,.8,1]'$ (dashed line), using the same realizations of ε_t. Clearly, the two series are very similar, though not identical. In this case we have an example of near observational equivalence: data generated from ARMA(1,1) model with $\phi_1 \approx \phi_2$ is difficult to distinguish from data generated by the model with arbitrary $\phi_1 = \phi_2$ and the same value of σ.

How can we detect observational equivalence (lack of identification) and near observational equivalence (weak identification)? A powerful result, due to Rothenberg (1971), provides a general necessary and sufficient condition for identification, namely, that the information matrix is non-singular.

Chart 2

NEAR OBSERVATIONAL EQUIVALENCE WITH AN ARMA(1,1) PROCESS



Source: Author's calculations.

As Rothenberg (1971) points out, the information matrix "is a measure of the amount of information about the unknown parameters available in the sample". A parameter is unidentified when there is no information about it in the sample, or if the existing information is insufficient to distinguish that parameter from other parameters in the model. Both cases result in a singular information matrix.

In the case of the ARMA (1,1) model, the information matrix is given by:

$$\mathcal{I}(\phi_1,\phi_2)=\begin{bmatrix}\dfrac{1}{1-\phi_1^2} & \dfrac{-1}{1-\phi_1\phi_2}\\ \dfrac{-1}{1-\phi_1\phi_2} & \dfrac{1}{1-\phi_2^2}\end{bmatrix} \tag{2.3}$$

From (2.4) we can compute the determinant of $\mathcal{I}(\phi_1,\phi_1)$

$$det\left(\mathcal{I}(\phi_1,\phi_2)\right)=\frac{(\phi_1-\phi_2)^2}{(1-\phi_1\phi_2)^2(1-\phi_1^2)(1-\phi_2^2)} \tag{2.4}$$

Since non-singularity is equivalent to the determinant of the matrix being different from zero, from (2.4) it is immediate that $\phi_1 \neq \phi_2$ is necessary and sufficient for identification in the ARMA(1,1) model.

The information matrix is also useful for detecting weak identification problems. A parameter is identified but poorly when the information in the sample is very little, or if it is barely possible to distinguish that parameter from the other parameters. In this case the information matrix has full rank, but is very close to being singular. The strength of identification may be measured using the result that the asymptotic covariance matrix of an efficient estimator is equal to the inverse of the information matrix divided by the sample size. Thus, the asymptotic variances of the estimators of the ARMA parameters ϕ_1 and ϕ_2 are:

$$\text{var}(\hat{\phi}_1)=\frac{(1-\phi_1\phi_2)^2(1-\phi_1^2)}{T(\phi_1-\phi_2)^2},\quad \text{var}(\hat{\phi}_2)=\frac{(1-\phi_1\phi_2)^2(1-\phi_2^2)}{T(\phi_1-\phi_2)^2} \tag{2.5}$$

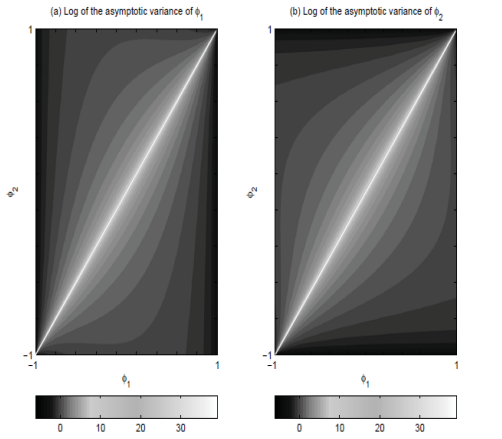
The formulas in (2.5) reveal that the asymptotic variances are large when $\phi_1 \approx \phi_2$. This suggests that the estimates of the autoregressive and moving average parameters will be very imprecise when their true values are similar. Therefore, ϕ_1 and ϕ_2 are weakly identified.

Note that both variances in (2.5) depend on the values of ϕ_1 and ϕ_2. Thus, for a given sample size T, the strength of identification of either parameter is determined by the true values of both parameters. This can be seen very clearly in Chart 3 which shows how the asymptotic variances vary across different regions in the parameter space.

To gain some intuition about the relationship between the parameter values and the strength of identification, consider the following decomposition of the information matrix (2.4)

Chart 3

ASYMPTOTIC VARIANCES OF THE PARAMETERS OF AN ARMA(1,1) PROCESS



Source: Author's calculations.

$$\mathcal{I}(\phi_1,\phi_2)=\begin{bmatrix}\frac{1}{\sqrt{1-\phi_1^2}} & 0\\ 0 & \frac{1}{\sqrt{1-\phi_2^2}}\end{bmatrix}\begin{bmatrix}1 & \frac{\sqrt{(1-\phi_1^2)(1-\phi_2^2)}}{\phi_1\phi_2-1}\\ \frac{\sqrt{(1-\phi_1^2)(1-\phi_2^2)}}{\phi_1\phi_2-1} & 1\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{1-\phi_1^2}} & 0\\ 0 & \frac{1}{\sqrt{1-\phi_2^2}}\end{bmatrix} \quad (2.6)$$

Note that the first and the last terms on the right hand side are the same diagonal matrix with elements equal to the square roots of the diagonal elements of $\mathcal{I}(\phi_1,\phi_2)$. This matrix tells us how much information there is in the sample about each parameter if the other parameter was known. For instance, $(1-\phi_1^2)/T$ is the asymptotic variance of an efficient estimator of ϕ_1 if ϕ_2 was known. Therefore, the closer is $|\phi_1|$ to 1, the more information there is about ϕ_1, for a given value of ϕ_2. Similarly, the closer is $|\phi_2|$ to 1, the more information there is about ϕ_2, for a given value of ϕ_1.

Next, consider the matrix in the middle. It is a correlation matrix which tells us how similar is the effect on the distribution of x_t of a small change in one parameter, say ϕ_1, to that of a small change in the other parameter ϕ_2. Note that $\mathcal{I}(\phi_1,\phi_2)$ is singular only when the correlation matrix in (2.6) is singular, which occurs if and only if the off-diagonal element, $\frac{\sqrt{(1-\phi_1^2)(1-\phi_2^2)}}{\phi_1\phi_2-1}$ is equal to -1. In this case a small change, say increase in ϕ_1, is exactly the same as a small decrease in the other parameter. When the correlation is close to, but different from 1 in absolute value, the effect of changing one parameter is almost the same as, though different from, that of changing the other one. Therefore, the middle term in (2.6) accounts for the loss of information about either parameter due to the uncertainty regarding the true value of the other parameter.

The information matrix approach to identification is possible only when the distribution of the data is known. What if we can not or do not want to assume that ε_t in (2.1) normally distributed? A reasonable approach in this case is to base the identification analysis on the ACF of x_t. As we already saw, it is straightforward to establish the non-identifiability of the autoregressive and moving average pa-

rameters at $\phi_1 = \phi_2$ using the theoretical ACF of the ARMA(1,1) process. More formally, we may proceed as follows: let $\gamma = [\gamma_0, \gamma_1, ..., \gamma_{k-1}]'$ be the vector of the first k -autocovariances of x_t. Then $\boldsymbol{\theta}$ is identified at $\boldsymbol{\theta}_0$ if the $(k \times 3)$-dimensional matrix $\partial\gamma / \partial \boldsymbol{\theta}$ has rank equal to 3 when evaluated at $\boldsymbol{\theta}_0$. The intuition behind this condition is very simple: the matrix has full column rank (equal to the dimension of $\boldsymbol{\theta}$) if and only if the vectors $\partial\gamma / \partial\phi_1, \partial\gamma / \partial\phi_2, \partial\gamma / \partial\sigma$ are linearly independent. For this to hold it must be impossible to match the effect on the moments of changing one parameter by changing the other two parameters. That is, each parameter plays a distinct role in determining the properties of the model, which is what identification requires.

Weak identification, on the other hand, means that the effect of changing one parameter on the moments of x_t can be approximated very closely by that of changing other parameters. This results in derivatives which are almost linearly dependent; for instance, having collinearity between $\partial\gamma / \partial\phi_1$ and $\partial\gamma / \partial\phi_2$ of nearly one (in absolute value) means that the effect of changing ϕ_1 on γ is very similar to that of changing ϕ_2 .

Table 1 illustrates the moments-based approach to identification in the ARMA(1,1) model. Columns 2 to 4 show the values of the derivatives of the first 10 autocovariances when the true values of the parameters are $\phi_1 = \phi_2 = 0, \sigma = 1$. As we can see, the derivatives with respect to ϕ_1 and ϕ_2 are perfectly negatively correlated. Thus the rank of $\partial\gamma / \partial\theta$ is only 2 and the ϕ_1 and ϕ_2 are not identified. Columns 5 to 7 similarly show the derivatives of γ evaluated at $\phi_1 = .7, \phi_2 = .8, \sigma = 1$. The degree of collinearity between $\partial\gamma / \partial\phi_1$ and $\partial\gamma / \partial\phi_2$ is $-.98$, which is high but less than -1. Thus, ϕ_1 and ϕ_2 are still identified though weakly.

Table 1

DERIVATIVE OF THE ACF OF A NARMA PROCESS

$\gamma(i)$	$\phi_1 = 0, \phi_2 = 0, \sigma = 1$			$\phi_1 = .7, \phi_2 = .8, \sigma = 1$		
	$\partial\gamma / \partial\phi_1$	$\partial\gamma / \partial\phi_2$	$\partial\gamma / \partial\sigma$	$\partial\gamma / \partial\phi_1$	$\partial\gamma / \partial\phi_2$	$\partial\gamma / \partial\sigma$
0	0.00	-0.00	4.00	-1.35	1.57	4.08
1	4.00	-4.00	0.00	3.13	-2.90	-0.35
2	1.60	-1.60	0.00	1.85	-2.03	-0.24
3	0.64	-0.64	0.00	1.05	-1.42	-0.17
4	0.26	-0.26	0.00	0.57	-1.00	-0.12
5	0.10	-0.10	0.00	0.28	-0.70	-0.08
6	0.04	-0.04	0.00	0.11	-0.49	-0.06
7	0.02	-0.02	0.00	0.02	-0.34	-0.04
8	0.01	-0.01	0.00	-0.03	-0.24	-0.03
9	0.00	-0.00	0.00	-0.05	-0.17	-0.02
10	0.00	-0.00	0.00	-0.05	-0.12	-0.01

Source: Author's calculations.

3. DSGE MODELS

In this section I discuss parameter identification in DSGE models. I will start with a brief outline of the general setup and then turn to analysis of two prototypical DSGE models.

3.1. Generalities

A DSGE model is summarized by a system of non-linear equations. Currently, most studies involving either simulation or estimation of DSGE models use linear approximations of the original models. That is, the model is first expressed in terms of stationary variables, and then linearized around the steady-state values of these variables. Once linearized, most DSGE models can be written in the following form:

$$\Gamma_0(\theta)z_t = \Gamma_1(\theta)E_t z_{t+1} + \Gamma_2(\theta)z_{t-1} + \Gamma_3(\theta)u_t \tag{3.1}$$

where z_t is a $m-$dimensional vector of endogenous and exogenous state variables, and the structural shocks u_t are independent and identically distributed n -dimensional random vectors with $Eu_t = 0, Eu_t u_t^{'} = I_n$. The elements of the matrices $\Gamma_0, \Gamma_1, \Gamma_2$ and Γ_3 are functions of a $k-$dimensional vector of deep parameters $\boldsymbol{\theta}$, where $\Theta \subset \mathbb{R}^k$ is a point in $\subset^k$. The parameter space Θ is defined as the set of all theoretically admissible values of $\boldsymbol{\theta}$.

There are several algorithms for solving linear rational expectations models (see for instance Blanchard and Kahn (1980), Anderson and Moore (1985), Klein (2000), Christiano (2002), Sims (2002)). Depending on the value of $\boldsymbol{\theta}$, there may exist zero, one, or many stable solutions. Assuming that a unique solution exists, it can be cast in the following form

$$z_t = A(\theta)z_{t-1} + B(\theta)u_t \tag{3.2}$$

where the $(m \times m)$ matrix A and the $(m \times n)$ matrix B are unique for each value of θ.

The model in (3.2) cannot be taken to the data directly since some of the variables in z_t are not observed. Instead, the solution of the model is expressed in a state space form, with a transition equation given by (7), and a measurement equation

$$x_t = s(\theta) + C(\theta)z_t \tag{3.3}$$

where x_t is a l -dimensional vector of observed state variables, s is a l -dimensional vector, and C is a $l \times m$ matrix.

The log-likelihood function of the data $X = [x_1, ..., x_T]$ may be computed using the Kalman filter if the structural shocks u_t are (assumed to be) jointly normally distributed. In this case the expected information matrix may be derived analytically as discussed in Iskrev (2008).

3.2. Identification in the RBC model

The first model I consider is a version of the one-sector stochastic growth model of Hansen (1985) with investment-specific technology shock. Below I outline the main features of the model.

3.2.1 The model

The representative household preferences are characterized by the lifetime utility function:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln(c_t) - \phi n_t \right) \tag{3.4}$$

where c_t is consumption in period t and n_t is the total labor supplied by the household.

Aggregate output is produced using capital k_t and labor using the following production function:

$$y_t = \exp(z_t) k_t^{1-\alpha} n_t^{\alpha} \tag{3.5}$$

where z_t is total factor productivity and follows an AR(1) process:

$$z_t = \rho_z z_{t-1} + \varepsilon_t^z, \quad \varepsilon_t^z \sim (0, \sigma_z^2) \tag{3.6}$$

The law of motion for aggregate capital is:

$$k_{t+1} = (1-\delta) k_t + \exp(u_t) i_t \tag{3.7}$$

where u_t is investment-specific technology and follows an AR(1) process:

$$u_t = \rho_u u_{t-1} + \varepsilon_t^u, \quad \varepsilon_t^u \sim (0, \sigma_u^2) \tag{3.8}$$

The resource constraint of the economy is:

$$c_t + i_t = y_t \tag{3.9}$$

3.2.2 Identification analysis

The model is log-linearized around the deterministic steady state of the variables, and the system is expressed as in (3.2). There are four potentially observable variables: output, consumption, hours worked and investment. Since there are only two structural shocks, we can use at most two variables to estimate the model with maximum likelihood; those may be any two of the for variables, or some linear combinations of them. The model has 8 deep parameters, which are collected in the vector $\boldsymbol{\theta} = [\alpha, \beta, \delta, \phi, \rho_z, \rho_u, \sigma_z, \sigma_u]$.

Let us first consider the case of using only one variable. This is an useful exercise as it tells us which variable is most informative for which of the (identifiable) parameters In this case in the measurement equation (3.3) x_t and s are scalars, and C is a row vector with 1 in the position of the observed

variable, and zeros elsewhere.

The identifiability of θ may be established using either the information matrix or the moment-based approach. Both show that 2 of the 8 parameters are not identified; these are β and δ, which, when there is only one observable variable, and irrespectively which one it is, cannot be identified separately. This is easy to see from the fact that the derivatives of the moments with respect to β and δ are collinear. However, if either β or δ is known, the remaining 7 parameters are identified. Table 2 shows the relative asymptotic standard deviations, defined as $\frac{std(\hat{\theta}_i)}{|\theta_i|}$, with each observable assuming that either β or δ are known. Note that there are substantial differences in the precision with which the parameters may be estimated depending on which variable is used and also on whether β or δ is known. For instance, output (y) is most informative for α if β is known and δ is estimated, but hours worked (n) is most informative when β is estimated and δ is known.

The reason why the relative standard deviations are reported is that they provide a measure of the identification strength which is independent of the value of the parameter. This permits us to determine which parameters are relatively better and which are relatively worse identified.

The results in Table 2 suggest that although it is possible to estimate most parameters with only one observable, the estimates are likely to be very imprecise. With two observed variables there is much more information about the parameters, and thus the estimation uncertainty, captured by the asymptotic standard deviation, is greatly reduced. This can be seen in Table 3, which reports the relative asymptotic standard deviations with each pair of observables. From the table we can see that all parameters are identified; generally, the best identified parameters are β, ρ_z and ρ_u, while the worst identified are ϕ, σ_z and σ_u.

To determine the causes for why some parameters are better and other worse identified, we can use a decomposition of the information matrix analogous to that in equation (2.6). Using it, we can express the relative standard deviation for a given parameter as a product of two terms: a sensitivity component, which is large for parameters which do not play an important role in the model, and a collinearity component, which is large for parameters whose role in the model is easy to approximate with other parameters. This decomposition is shown in Table 4. We can see that the reason why β

Table 2

IDENTIFCATION STRENGTH IN THE RBC MODEL WITH ONE OBSERVABLE

Par.	true	c	y	i	n	c	y	i	n
α	0.670	187	35	135	62	197	97	11797	30
ϕ	0.025	981	3103	656	268	*fixed*	*fixed*	*fixed*	*fixed*
δ	0.980	*fixed*	*fixed*	*fixed*	*fixed*	26	98	1667	6.1
β	2.000	287	652	918	8.4	289	1246	58	47
ρ_z	0.950	5.6	20	37	29	5.6	20	37	29
ρ_u	0.970	8.6	17	32	20	8.6	17	32	20
σ_z	1.000	241	103	545	952	250	257	7497	1097
σ_u	1.000	306	1843	2401	1051	289	2102	927	929

Source: Author's calculations.

Note: Each column of the table shows the relative asymptotic standard deviations of θ when there is only one observed variable (shown in the first row)and either β or δ is assumed known. The results are obtained using the expected information matrix and T=100.

Table 3

IDENTIFCATION STRENGTH IN THE RBC MODEL WITH TWO OBSERVABLES

Par.	true	(c,y)	(c,i)	(c,n)	(y,i)	(y,n)	(i,n)
α	0.670	0.325	0.274	0.321	0.285	0.307	0.288
δ	0.025	1.027	0.454	1.125	0.454	1.119	0.453
β	0.980	0.024	0.025	0.025	0.018	0.025	0.021
ϕ	2.000	3.005	1.958	1.178	1.420	1.164	0.726
ρ_z	0.950	0.033	0.059	0.033	0.050	0.033	0.043
ρ_u	0.970	0.052	0.051	0.051	0.051	0.050	0.052
σ_z	1.000	0.299	0.381	0.295	0.264	0.283	0.346
σ_u	1.000	0.604	0.606	0.590	0.679	0.582	0.717

Source: Author's calculations.

Note: Each column of the table shows the relative asymptotic standard deviations of θ when there are only two observed variables (shown in the first row) The results are obtained using the expected information matrix and T=100 .

is so well identified is that its sensitivity component is very low; this implies that β is a very important determinant of the empirical properties of the model variables. On the other extreme is ϕ , which has very large sensitivity component, and because of that is the worst identified parameter. Strong collinearity explains the different strength of identification of σ_z and σ_u which have the same sensitivity components. Other parameters with strong collinearity are α , δ and ρ_u . As was already discussed in Section 2, strong collinearity implies that two or more parameters play similar role in the model. It is interesting to know what these parameters are. A simple way to find out is to compute coefficients of pairwise collinearity, which measure how similar the effects of two parameters are. This is done in Table 5 and we can see that there is a strong negative collinearity between σ_u and ρ_u on one hand and between β and δ , on the other. Thus, having higher volatility of the investment specific shock is similar to having lower persistence of the same shock, and having more patient consumers is similar to having lower depreciation rate. Furthermore, we can also see that when the included observables

Table 4

SENSITIVITY AND COLLINEARITY IN THE RBC MODEL WITH TWO OBSERVABLES

Par.	(c,y)		(c,i)		(c,n)		(y,i)		(y,n)		(i,n)	
	sens.	col.	sens.	col.	sens.	col.	sens.	col.	sens.	col.	sens.	col.
α	0.056	5.8	0.028	9.7	0.057	5.7	0.014	19.8	0.057	5.4	0.003	93.2
δ	0.191	5.4	0.087	5.2	0.191	5.9	0.045	10.0	0.192	5.8	0.015	30.0
β	0.005	4.8	0.005	5.0	0.005	5.0	0.001	13.6	0.005	5.0	0.000	73.4
ϕ	2.374	1.3	1.799	1.1	1.127	1.0	1.195	1.2	1.100	1.1	0.647	1.1
ρ_z	0.014	2.4	0.014	4.2	0.014	2.4	0.014	3.6	0.014	2.4	0.014	3.1
ρ_u	0.008	6.5	0.008	6.5	0.008	6.4	0.008	6.5	0.008	6.4	0.008	6.5
σ_z	0.071	4.2	0.071	5.4	0.071	4.2	0.071	3.7	0.071	4.0	0.071	4.9
σ_u	0.071	8.5	0.071	8.5	0.071	8.3	0.071	9.6	0.071	8.2	0.071	10.1

Source: Author's calculations.

Note: Each column of the table shows the sensitivity and collinearity components of the relative asymptotic standard deviations of θ when there are two observed variables (shown in the first row).The results are obtained using the expected information matrix and T=100 .

Table 5

STRONGEST PAIRWISE COLLINEARITY IN THE RBC MODEL WITH TWO OBSERVABLES

Par.	(c,y)		(c,i)		(c,n)		(y,i)		(y,n)		(i,n)	
	pcol	par.	pcol	par.	pcol	par.	pcol	par.	pcol	par.	pcol	par.
α	0.70	σ_z	-0.95	β	0.71	σ_z	-0.98	δ	0.71	σ_z	-0.999	β
δ	-0.97	β	-0.74	α	-0.98	β	-0.98	α	-0.98	β	-0.997	α
β	-0.97	δ	-0.95	α	-0.98	δ	-0.96	α	-0.98	δ	-0.999	α
ϕ	0.12	α	0.04	α	-0.04	α	0.04	α	-0.06	α	-0.004	α
ρ_z	-0.56	β	-0.65	α	-0.56	β	-0.49	α	-0.56	β	-0.12	β
ρ_u	-0.97	σ_u	-0.97	σ_u	-0.97	σ_u	-0.97	σ_u	-0.97	σ_u	-0.97	σ_u
σ_z	0.70	α	0.72	α	0.71	α	0.45	α	0.71	α	-0.71	δ
σ_u	-0.97	ρ_u	-0.97	ρ_u	-0.97	ρ_u	-0.97	ρ_u	-0.97	ρ_u	-0.97	ρ_u

Source: Author's calculations.

Note: The table shows which parameters are most strongly related to each deep parameter as well as the value of the pairwise collinearity ($pcol$) coefficients. The results are obtained using the expected information matrix and T=100.

are output and investment, α is strongly collinear with both β and δ. This means that the effect of these parameters on the moments and cross moments of output and investment are difficult to distinguish.

3.3. Identification in the New Keynesian model

In this section I consider a small-scale New Keynesian model studied in An and Schorfheide (2007). A brief description of the model follows.

3.3.1 The model

The representative household maximizes lifetime utility function

$$E_t[\sum_{s=0}^{\infty}\beta^s(\frac{(C_{t+s}\ /\ A_{t+s}-1)^{1-\tau}}{1-\tau}-N_{t+s})], \tag{3.10}$$

subject to a budget constraint:

$$P_tC_t+B_t+T_t=P_tW_tN_t+R_{t-1}B_{t-1}+P_tD_t+P_tSC_t, \tag{3.11}$$

where C_{t+s} is consumption, $N_{t+s}(j)$ is hours worked, P_t is the price of the final good, W_t is the real wage, R_t is the interest on the government bonds B_t, D_t is the residual real profit, T_t is lump-sum taxes and SC_t is net cash flow from trading state-contingent securities. A_t is stock of habit given by the level of technology in the intermediate good sector, and evolves according to

$$\Delta\ln A_t=\ln\gamma+\ln z_t,\qquad \ln z_t=\rho_z\ln z_{t-1}+\varepsilon_t^z,\quad \varepsilon_t^z\sim\mathbb{N}(0,\sigma_z^2)$$

There is a perfectly competitive sector producing a single final good from intermediate inputs $Y_t(j)$ using the technology

$$Y_t = (\int_0^1 Y_t(j)^{1-\nu} dj)^{\frac{1}{1-\nu}} \tag{3.12}$$

The final goods firm maximize profits given by

$$P_t Y_t - \int_0^1 P_t(i) Y_t(i) di, \tag{3.13}$$

where $P_t(i)$ is the price of intermediate good $Y_t(i)$.

Intermediate goods are produced in a monopolistically competitive sector. Each variety i is produced by a single firm using the following production technology:

$$Y_t(i) = A_t N_t(i) \tag{3.14}$$

The intermediate goods firm j maximizes the present value of its future profits

$$E_t[\sum_{s=0}^{\infty} \beta^s Q_{t+s|t} (\frac{P_{t+s}(j)}{P_{t+s}} Y_{t+s}(j) - W_{t+s} N_{t+s}(j) - AC_{t+s}(j))], \tag{3.15}$$

where $Q_{t+s|t}$ is the time t value to the consumers of a unit of the final good in period $t+s$; $AC_t(j) = \frac{\phi}{2}(\frac{P_t(j)}{P_{t-1}(j)} - \pi)^2 Y_t(j)$ is the cost of adjusting prices and π is the steady state rate of inflation.

The central bank sets the nominal interest rate according to the following rule

$$\frac{R_t}{r^* \pi^*} = \exp(\varepsilon_t^r)(\frac{R_{t-1}}{r^* \pi^*})^{\rho_r} [(\frac{\pi_t}{\pi^*})^{\psi_1} (\frac{Y_t}{Y_{t-1}\gamma})^{\psi_2}]^{1-\rho}, \tag{3.16}$$

where r^* is the steady state real interest rate, π_t is the gross inflation rate, π^* is the inflation target rate, and $\varepsilon_t^r \sim \mathbb{N}(0, \sigma_r^2)$ is a monetary policy shock.

The government collects lump-sum taxes in order to finance its consumption so as to respect the following budget constraint

$$P_t G_t + B_{t-1} R_{t-1} = T_t + B_t, \tag{3.17}$$

where $G_t = \zeta_t Y_t$ is government consumption in terms of final good, and $\zeta_t = 1 - 1/g_t$ where g_t is random variable evolving according to

$$\ln g_t = (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \varepsilon_t^g, \quad \varepsilon_t^g \sim \mathbb{N}(0, \sigma_g^2)$$

3.3.2 Identification analysis

Again, the model is log-linearized around the deterministic steady state of the variables, and the system may be expressed as in (3.2). There are four potentially observable variables: output, consumption, inflation and the nominal interest rate. Since there are only three structural shocks, we can use at most three variables to estimate the model with maximum likelihood. The model has 14 deep

parameters, which are collected in the vector $\boldsymbol{\theta} = [\tau, \nu, \phi, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, r^*, \pi^*, \gamma, \sigma_r, \sigma_g, \sigma_z]$.

Let us first consider identification with only two observed variables. Two of the 14 parameters, ϕ and ν, are not identified with any pair of observables. Examining the derivatives of the moments shows that this is due to the perfect collinearity of the derivatives with respect to these two parameters. Therefore, if either one of the two parameters is fixed, the other one would be identified along with the other 12 parameters. An exception to this conclusion is the case when only output and consumption are observed. Then we have to fix three more parameters, in addition to ν or ϕ. For example, if we fix ν, ψ_2, π^* and σ_r, we could identify the remaining 10 parameters.

The reason why the (output,consumption) pair is less informative is that the behavior of the two variables in the model is very similar. Therefore, consumption adds very little information to that already contained in output. This can be seen in Table 6, which shows the asymptotic standard deviations for each pair of observables assuming that some of the elements of $\boldsymbol{\theta}$ are known. The estimation uncertainty of most parameters is much larger, compared to the other pais of observables, even though more parameters are assumed known. Note that, as in the RBC model, there is a substantial difference in the information content of different variables. Also, which pair of variables is best to use for estimation depends on the parameters one is most interested in. For instance, the policy response to inflation parameter ψ_1 is best identified with (π, r) while the policy response to output growth ψ_2 is best identified with (y, r).

Next, consider using three out of the four observables to estimate $\boldsymbol{\theta}$. Table 7 reports the asymptotic

Table 6

PARAMETER IDENTIFCATION IN THE NKM MODEL WITH TWO OBSERVABLES

Par.	true	(y,π)	(y,r)	(y,c)	(π,r)	(π,c)	(r,c)	(y,π)	(y,r)	(y,c)	(π,r)	(π,c)	(r,c)
τ	2.00	3.9	3.3	554	689	2.3	12	3.9	3.3	554	689	2.3	12
ν	0.10	*fixed*	*fixed*	*fixed*	*fixed*	*fixed*	*fixed*	7.5	786	614	32	2.7	3160
ϕ	7.50	8.3	873	682	36	3.0	3511	*fixed*	*fixed*	*fixed*	*fixed*	*fixed*	*fixed*
ψ_1	1.50	20	11	2719	8.9	206	90	20	11	2719	8.9	206	90
ψ_2	1.00	18	14	*fixed*	230	193	72.7	18.4	14.5	*fixed*	230	193	72
ρ_r	0.96	0.7	0.5	118	21	7.5	4.5	0.7	0.5	118	21	7.5	4.5
ρ_g	0.95	0.95	0.7	0.1	*fixed*	*fixed*	21	0.2	0.7	0.1	*fixed*	21.7	0.2
ρ_z	0.65	0.8	0.9	111	34	0.3	0.3	0.8	0.9	111	34	0.3	0.3
r^*	0.40	3927	4432	293802	13845	562	17825	3927	4432	293802	13845	562	17825
π^*	4.00	0.3	443	*fixed*	0.3	0.3	1782	0.3	443	*fixed*	0.3	0.3	1782
γ	0.50	0.3	0.3	0.3	2769	0.3	0.3	0.3	0.3	0.3	2769	0.3	0.3
σ_r	0.20	26	1.9	*fixed*	493	18	11	26	1.9	*fixed*	493	18	11
σ_g	0.80	2.7	2.0	0.1	596	276	80	2.7	2.0	0.1	596	276	80
σ_z	0.45	1.5	1.7	207	139	0.5	0.5	1.5	1.7	207	139	0.5	0.5

Source: Author's calculations.

Note: Each column of the table shows the relative asymptotic standard deviations of θ when there are two observed variables (shown in the first row)and some deep parameters are assumed known. The results are obtained using the expected information matrix and T=100.

Table 7

PARAMETER IDENTIFCATION IN THE NKM MODEL WITH THREE OBSERVABLES

Par.	true	(y,π,r)	(y,π,c)	(y,r,c)	(π,r,c)	(y,π,r)	(y,π,c)	(y,r,c)	(π,r,c)
τ	2.00	0.36	0.56	0.26	0.32	0.36	0.56	0.26	0.32
ν	0.10	*fixed*	*fixed*	*fixed*	*fixed*	0.71	1.36	188.38	0.72
ϕ	7.50	0.79	1.52	209.31	0.80	*fixed*	*fixed*	*fixed*	*fixed*
ψ_1	1.50	1.24	1.37	1.06	3.58	1.24	1.37	1.06	3.58
ψ_2	1.00	1.41	1.36	1.18	3.28	1.41	1.36	1.18	3.28
ρ_r	0.96	0.03	0.10	0.03	0.14	0.03	0.10	0.03	0.14
ρ_g	0.95	0.03	0.04	0.04	0.03	0.03	0.04	0.04	0.03
ρ_z	0.65	0.27	0.29	0.27	0.21	0.27	0.29	0.27	0.21
r^*	0.40	3.40	323.22	1062.13	3.39	3.40	323.22	1062.13	3.39
π^*	4.00	0.31	0.31	106.21	0.31	0.31	0.31	106.21	0.31
γ	0.50	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
σ_r	0.20	0.39	2.32	0.28	0.35	0.39	2.32	0.28	0.35
σ_g	0.80	0.15	0.07	0.07	2.22	0.15	0.07	0.07	2.22
σ_z	0.45	0.30	0.35	0.33	0.17	0.30	0.35	0.33	0.17

Source: Author's calculations.

Note: Each column of the table shows the relative asymptotic standard deviations of θ when there are two observed variables(shown in the first row) and and either ν or ϕ is assumed known. The results are obtained using the expected information matrix and T=100.

standard deviations for each triplet of observables assuming that either ν or ϕ is known. As in Table 6, which one of the two parameters is fixed has no effect on the standard deviation of the other parameters. Worst identified with all combinations of observables are the response coefficients of the Taylor rule (ψ_1 and ψ_2), the price stickiness and inverse elasticity of demand parameters (ϕ and ν), and the steady state interest rate (r^*); best identified are the interest rate smoothing parameter (ρ_r) and the government consumption shock parameter ρ_g.

Table 8 shows the decompositions of the relative standard deviations into sensitivity and collinearity components. Note that most of the worst identified parameters are also the ones with the largest collinearity components. Thus, these parameters are poorly identified because their effects on the empirical properties of the observables are easy to mimic with other parameters. An exception is r^*, which is poorly identified because of the very large sensitivity component. This implies that the value of r^* is of little consequence empirically. Note that both r^* and π^* have huge collinearity components when π is not among the observables. For example, the value for π^* translates into a multiple collinearity coefficient of $.999999875$.[1] This means that π^* is almost impossible too distinguish from other model parameters unless its effect on inflation is accounted for. Computing the pairwise collinearity coefficients, reported in Table 9, shows that when inflation is not among the observables, the collinearity between π^* and r^* is $.966$. There we also see that the policy response to inflation ψ_1 is highly collinear with either the price stickiness parameter ϕ or the interest rate smoothing parameter

(1) The multiple collinearity coefficient measures the degree of collinearity between a given parameter and all other model parameters.

Table 8

SENSITIVITY AND COLLINEARITY IN THE NKM MODEL WITH THREE OBSERVABLES

Par.	(y,π,r)		(y,π,c)		(y,r,c)		(π,r,c)	
	sens.	**col.**	**sens.**	**col.**	**sens.**	**col.**	**sens.**	**col.**
τ	0.045	8.1	0.043	13.1	0.027	9.7	0.030	10.7
ϕ	0.022	36.8	0.067	22.6	0.402	520.5	0.038	21.3
ψ_1	0.011	116.4	0.073	18.9	0.029	36.5	0.016	220.6
ψ_2	0.021	66.7	0.295	4.6	0.058	20.4	0.035	93.6
ρ_r	0.001	44.5	0.003	34.3	0.002	15.7	0.001	152.7
ρ_g	0.010	2.9	0.041	1.0	0.026	1.5	0.013	2.4
ρ_z	0.109	2.4	0.100	2.9	0.078	3.5	0.088	2.4
r^*	0.431	7.9	21.597	15.0	0.550	1932.4	0.432	7.9
π^*	0.010	29.7	0.033	9.3	0.053	2003.3	0.017	18.2
γ	0.082	3.1	0.247	1.0	0.107	2.4	0.082	3.1
σ_r	0.071	5.5	0.070	33.0	0.071	4.0	0.070	5.0
σ_g	0.071	2.1	0.071	1.0	0.071	1.0	0.071	31.4
σ_z	0.071	4.3	0.071	4.9	0.071	4.6	0.071	2.4

Source: Author's calculations.

Note: Each column of the table shows the sensitivity and collinearity components of the relative asymptotic standard deviations of when there are three observed variables (shown in the first row).The results are obtained assuming $\nu = .10$ is known, and using the expected information matrix with T=100.

Table 9

STRONGEST PAIRWISE COLLINEARITY IN THE NKM MODEL WITH THREE OBSERVABLES

Par.	(y,π,r)		(y,π,c)		(y,r,c)		(π,r,c)	
	pcol.	**par.**	*pcol.*	**par.**	*pcol.*	**par.**	*pcol.*	**par.**
τ	-0.76	ρ_r	-0.90	ϕ	-0.95	ρ_r	-0.68	ρ_r
ϕ	0.96	π^*	0.99	π^*	0.76	σ_r	0.90	π^*
ψ_1	0.92	ϕ	-0.97	ρ_r	-0.91	ρ_r	-0.89	ρ_r
ψ_2	-0.91	ψ_1	-0.87	r^*	-0.89	ψ_1	-0.89	ψ_1
ρ_r	-0.96	ϕ	0.99	σ_r	-0.95	τ	-0.89	ψ_1
ρ_g	-0.70	σ_g	-0.04	σ_g	-0.21	ψ_2	-0.90	σ_g
ρ_z	-0.29	ψ_2	0.87	σ_z	0.95	σ_z	0.80	σ_z
r^*	0.94	γ	0.98	ϕ	0.97	π^*	0.94	γ
π^*	0.96	ϕ	0.99	ϕ	0.97	r^*	0.90	ϕ
γ	0.94	r^*	0.03	π^*	0.91	r^*	0.94	r^*
σ_r	0.11	ρ_r	0.99	ρ_r	0.76	ϕ	0.35	ρ_r
σ_g	-0.70	ρ_g	-0.04	ρ_g	-0.01	ρ_g	-0.90	ρ_g
σ_z	0.42	ψ_2	0.87	ρ_z	0.95	ρ_z	0.80	ρ_z

Source: Author's calculations.

Note: The table shows which parameters are most strongly related to each deep parameter as well as the value of the pairwise collinearity coefficients ($pcol$). The results are obtained assuming $\nu = .10$ is known, and using the expected information matrix with T=100.

ρ_r , while the response to output ψ_2 is highly collinear with either ψ_1 or r^* .

4. CONCLUDING REMARKS

In the recent years DSGE models are increasingly becoming an important tool for quantitative policy analysis. This has lead to a considerable research effort aimed to increasing the models' complexity and realism. As the number of number of features represented in the models increases, it becomes very difficult to understand by reasoning alone their separate contribution to the model performance vis-a-vis the reality they are supposed to explain. In this note I have tried to show that studying parameter identification may provide useful insights regarding the model parameters and the structural features they represent. The strength of parameter identification reflects their importance in determining the empirical implications of the model. Weak identification arises when some model features have little empirical relevance; this may occur either because they are unimportant on their own, or because they are redundant given the other features represented in the model. Since DSGE models provide a complete characterization of the dynamics of the model variables, parameter identification may be treated as a property of the underlying model and studied without a reference to a particular data set. I have illustrated this approach to parameter identification using two canonical macroeconomic model - a real business cycle model and a new Keynesian model. One limitation of this analysis is that only a single parameter value was considered. To obtain a complete picture of identification as a property of the model, one has to study it across different theoretically plausible parameter values. For a more detailed discussion of this and other important aspects of the a priori analysis of identification, the reader may consult the papers cited in the introduction.

REFERENCES

An, S., e F. Schorfheide (2007): "Bayesian Analysis of DSGE Models", *Econometric Reviews*, 26(2-4), 113-172.

Anderson, G., e G. Moore (1985): "A linear algebraic procedure for solving linear perfect foresight models", *Economics Letters*, 17(3), 247-252,http://ideas.repec.org/a/eee/ecolet/v17y1985i3p247-252.html.

Blanchard, O. J., e C. M. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations", *Econometrica*, 48(5), 1305-11, http://ideas.repec.org/a/ecm/emetrp/v48y1980i5p1305-11.html.

Chari, V. V., P. J. Kehoe, e E. R. McGrattan (2009): "New Keynesian Models: Not Yet Useful for Policy Analysis", *American Economic Journal: Macroeconomics*, 1(1), 242-66.

Christiano, L. J. (2002): "Solving dynamic equilibrium models by a method of undetermined coefficients", *Computational Economics*, 20(1-2).

Hansen, G. D. (1985): "Indivisible labor and the business cycle", *Journal of Monetary Economics*, 16(3), 309--327, http://ideas.repec.org/a/eee/moneco/v16y1985i3p309-327.html.

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

_____(2010a): "Local identification in DSGE models", *Journal of Monetary Economics*, 57(2), 189-202.

_____(2010b): *On the choice of observables in DSGE models*, *mimeo*.

Klein, P. (2000): "Using the generalized Schur form to solve a multivariate linear rational expectations model", *Journal of Economic Dynamics and Control*, 24(10), 1405-1423, http://ideas.repec.org/a/eee/dyncon/v24y2000i10p1405-1423.html.

Rothenberg, T. J. (1971): "Identification in Parametric Models", *Econometrica*, 39(3), 577-91, http://ideas.repec.org/a/ecm/emetrp/v39y1971i3p577-91.html.

Sims, C. A. (2002): "Solving Linear Rational Expectations Models", *Computational Economics*, 20(1-2), 1-20, http://ideas.repec.org/a/kap/compec/v20y2002i1-2p1-20.html.