Dynare Working Papers Series
http://www.dynare.org/wp/

Choosing the variables to estimate singular DSGE models: Comment

Iskrev Nikolay

Working Paper no. 41

October 2014

CEPREMAP
CENTRE POUR LA RECHERCHE ECONOMIQUE ET SES APPLICATIONS
142, rue du Chevaleret — 75013 Paris — France
http://www.cepremap.fr

Choosing the variables to estimate singular DSGE models: Comment

Nikolay Iskrev
Banco de Portugal

Abstract

In a recent article Canova et al. (2014) study the optimal choice of variables to use in the estimation of a simplified version of the Smets and Wouters (2007) model. In this comment I examine their conclusions by applying a different methodology to the same model. The results call into question most of Canova et al. (2014) findings.

Keywords: DSGE models, Observables, Identification, Information matrix, Cramér-Rao lower bound

JEL classification: C32, C51, C52, E32

1 Introduction

DSGE models are usually estimated using only a subset of the variables that are present in them. This is partly due to the fact that some variables, such as capital, are not observed. However, even variables for which data exist are often not utilized. One could explain this with the restriction that the number of included observables should not be greater than the number of shocks in the model. While there are ways to get around this restriction,[1] it is a fact that in the literature there are similar DSGE models estimated with different sets of observables. It is not clear what motivates these different choices, nor what consequences that has on the empirical findings.

In a recent article Canova, Ferroni, and Matthes (2014) (CFM henceforth) seek to provide some guidance on how to select the most informative among several available sets of observables. They propose the use of two criteria which rank different combinations of variables according to measures of identification and information content. The first criterion starts by selecting the sets of variables that satisfy a rank condition for identification of the free model parameters. To pick the best among the selected sets, measures of closeness to a convoluted singular system of all observables are computed in terms of sensitivity of the log-likelihood function to parameters of interest. The one yielding smallest discrepancy is chosen as the most informative. The second criterion is based on Bierens (2007) and uses convolutions of both the singular and non-singular systems with the same non-singular distribution. The combination of variables whose convoluted distribution is closest to the convoluted singular system of all available observables is selected as being the most informative.

CFM apply their selection criteria to a simplified version of the Smets and Wouters (2007) model. The model has 4 shocks and a total of 7 observables, namely output (y_t), consumption (c_t), investment (i_t), wages (w_t), hours (h_t), inflation (π_t), and nominal interest rate (r_t). Thus, 35 combinations of variables are available to use in estimation. Among these, as most informative overall the authors select y_t, c_t, i_t and either w_t or h_t. Furthermore, it is argued that the ranking of different sets of variables does not depend on the value of the parameters at which the model is evaluated, and is robust to increasing the number of shocks as in the original Smets and Wouters (2007) model.

The purpose of this comment is to evaluate these claims, applying a different approach to the same model. As in Iskrev (2010), where the choice of observables is studied with respect to the original Smets and Wouters (2007) model, here I use criteria based on the expected Fisher information matrix (FIM). Using the FIM has several advantages. First, as the name suggests, it is a measure of the amount of information about the parameters available in a sample (see Rothenberg (1971)). It takes the model as it is and does not require convoluting the true data density as the measures CFM use do.[2] Second, FIM depends on the set of observables and the sample size, but does not depend on actual data. Thus, the information one could expect to have in different sets of observables and in samples of different sizes can be measured and compared prior to estimation. Third, using the FIM one can compute measures of expected estimation uncertainty with respect to each model parameter. In general, there is a trade-off between the amount of information contained in different sets of observables with respect

[1] One is to introduce measurement errors in the observed series. Another is the approach in Bierens (2007).

[2] This approach follows Bierens (2007) where the theoretical model is assumed to be misspecified and the singular distribution it implies is convoluted and compared to a convoluted distribution of an a-theoretical econometric model which is assumed to represent the true data generating process.

to different parameters. Quantifying the amount of information for each parameter provides a clearer understanding of the trade-offs involved in selecting one set of observables over another. The measures CFM use do not provide such information. And fourth, the FIM can be evaluated analytically for linearized Gaussian model such as the one in Smets and Wouters (2007). This is very useful in practice since it allows many possible combinations of variables to be compared quickly for a large number of *a priori* plausible parameter values. Furthermore, the use of analytical derivatives minimizes the risk of reaching wrong conclusions as a result of numerical errors.

2 Analysis

In this section I apply the FIM approach to the model analyzed in CFM. I address three main questions: (1) is the rank condition useful for selecting the set of observables, (2) which is the most informative set of four variables out of the seven available, (3) are the results sensitive to changes in the parameter values and the number of shocks.

2.1 Is the rank condition useful?

I start by checking whether the parameters of the simplified SW model are identified if only four of the seven variables are observed. It is well known that four parameters - ξ_w, ξ_p, ϵ_w and ϵ_p, are not separately identifiable in the sense that in the linearized model ξ_w cannot be distinguished from ϵ_w, and ξ_p cannot be distinguished from ϵ_p. As in the original paper, I will assume that ϵ_w and ϵ_p are both known. This leaves 27 free parameters.

A necessary and sufficient condition for local identification is that the FIM has full rank. When evaluated at the parameter values from Table 2 in CFM, the FIM has full rank of 27 for all 35 combinations of four variables. Thus, the rank condition alone provides no useful information regarding the best set of variables to use in estimating the model.

2.2 Which are the best four observables?

Selecting the best combination of variables requires a criterion according to which to compare and rank the alternatives. Which criterion should be used depends on the purpose for which the model is estimated. In any case, the criterion would be a function of the estimated parameters and would rank higher sets of observables that are more informative about the relevant function of the parameters of interest $\boldsymbol{\theta}$.

When the objective is to minimize the estimation uncertainty about $\boldsymbol{\theta}$ as a whole, a popular criterion to use is the natural logarithm of the determinant of the inverse of the FIM, i.e. $\ln(\det(\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})))$. This is known in the optimal design literature as D-optimality criterion. The well-known Cramér-Rao (CR) theorem tells us that, depending on whether the asymptotic FIM is used or the finite sample one, its inverse gives either a lower bound on the asymptotic covariance matrix of any consistent estimator of $\boldsymbol{\theta}$, or a lower bound on the covariance matrix of any unbiased estimator $\boldsymbol{\theta}$. Furthermore, the diagonal elements of $\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})$ are lower bounds on the variances of estimators of individual parameters. This can be used to construct a criterion which assigns different weights to the parameters, to reflect their relative importance for the researcher. An example of such a criterion is the weighted geometric average of the diagonal elements of $\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})$,

$$\left(\prod_{i=1}^{k} \mathrm{b}_i^{w_i}\right)^{1/\sum_{i=1}^{k} w_i} \tag{2.1}$$

where b_i is the i-th diagonal element of $\boldsymbol{\mathcal{I}}^{-1}(\boldsymbol{\theta})$, k is the number of free parameters, and w_i is the weight assigned to θ_i. The geometric average is more appropriate to use than the arithmetic average since parameters typically have different range.

In what follows I use the finite sample FIM in order to take a proper account of the size of the sample, which is set to T=150, as in CFM.[3] I report three versions of the weighted geometric average criterion with: (1) equal weights on all free parameters; (2) equal weights on the free structural parameters and zero weights on the shock parameters; (3) equal weights on the six parameters emphasized in CFM, namely λ, ι_p, ξ_p, σ_l, r_π, and r_y, and zero weights on all other parameters. To be comparable with CFM, I assume that δ, λ_w and c_g are known. This leaves 24 free parameters, 17 of which are structural and the other 7 are shock parameters.

Table 1: Ranking of observables, baseline parameterization

	geometric average			D-optimality
rank	24 parameters	17 parameters	6 parameters	
1	c, i, π, r	c, i, π, r	c, i, π, r	c, i, π, r
2	y, c, w, r	y, w, π, h	c, i, w, r	y, w, π, h
3	y, c, w, h	y, c, π, h	c, w, π, h	c, i, w, r
33	c, i, w, π	y, i, π, r	y, c, i, π	y, i, π, r
34	y, c, i, π	y, c, i, π	y, i, r, h	y, i, r, h
35	y, c, i, h	$y, c, i, h,$	y, c, i, h	y, c, i, h

Note: The table shows the best 3 and the worst 3 sets of observables according to the geometric average and D-optimality criteria. The geometric average criterion is computed for 3 groups of parameters: 24 (all free) parameters; 17 (all except shock) parameters; 6 (only λ,ι_p,ξ_p,σ_l,r_π,r_y) parameters.

Table 1 shows the best three and the worst three sets of variables according to each criterion. The criteria are in full agreement regarding the most informative and the least informative sets, all selecting (c, i, π, r) as best and (y, c, i, h) as worst. There is also a broad agreement in the rest of the ranking, and in particular in the absence of (y, c, i) among the most informative sets. The differences between most and least informative combinations can be seen in Figure 1, which shows the sorted values for of the different criteria.

Table 2 presents the values of the individual CR bounds for the three most informative and the three least informative sets of variables as selected by the geometric average criterion with weights only on the six deep parameters. In addition, the set (y, c, i, w) is also included as it was selected by CFM as being one of the two most informative combinations.

[3]The asymptotic FIM is defined as the limit of the average finite sample FIM, which in turn is the negative expected Hessian of the log-likelihood function. With the asymptotic FIM information accumulates at a constant rate T and therefore the ranking of observables does not depend on the sample size. With the final sample FIM information may accumulate at different and changing rates for different sets of observables. Thus, the ranking may change with the sample size.

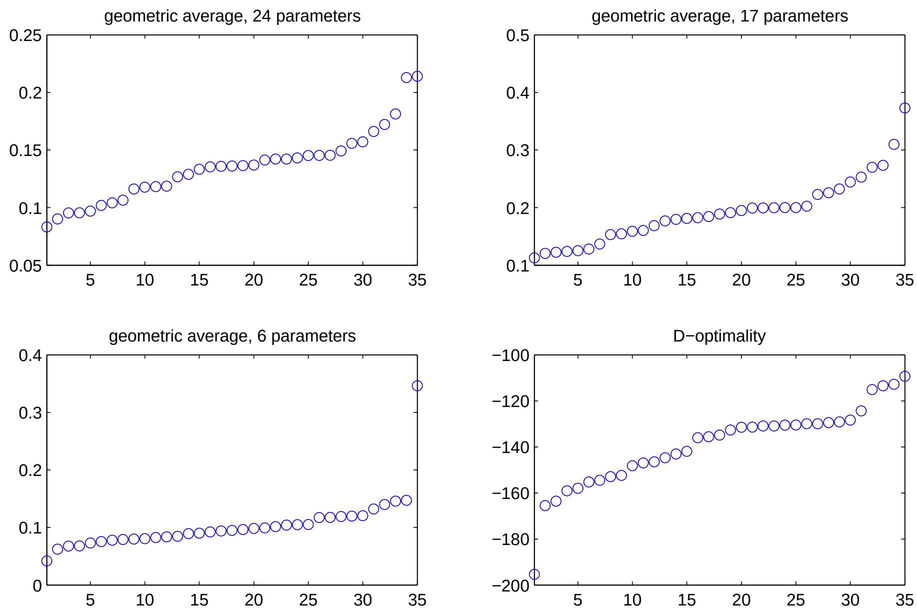

Figure 1: Sorted values of the variables selection criteria.

As can be seen from the results, choosing one combination of variables over another usually involves a trade-off in terms of information about different parameters. Even the least informative according to all criteria set, (y, c, i, h), is the most informative one, amongst those in the table, for two parameters, ρ_{ga} and σ_a. The overall best set (c, i, π, r) yields the lowest (among these seven sets) CR bounds for 14 of the 24 free parameters, including four of the six deep parameters CFM focus on. The set (y, c, i, w), which is one of the two preferred combinations of CFM, is most informative (among the seven sets in the table) for five parameters, including two of the six they focus on, ξ_p and σ_l. However, choosing that combination involves large information loss with respect to the other four parameters, especially the Taylor rule ones, which is not surprising since interest rate and inflation are naturally very informative variables for those parameters. Notice, in panel B of the table, that the value of the loss function for (y, c, i, w) is about twice as large as for any of the top three combinations. Hence, the only way to pick (y, c, i, w) as best would be to assign a much larger weights on σ_l and ξ_p than on the other four parameters.

One of the criteria used by CFM ranks the sets of variables on the basis of the sensitivity of the likelihood to a selected group of parameters. The measures they use compare the scores of the non-singular and convoluted singular systems, and require simulated data to compute. A simpler and more direct measure of sensitivity to a single parameter θ_i is the expected curvature of log-likelihood function, given by $\mathrm{E}\left(\frac{\partial \ell_T(\boldsymbol{\theta})}{\partial \theta_i}\right)^2$. Note that this is just the i-th diagonal element of the FIM and does not require data to compute. It is interesting to compare the sets of variables which maximize the likelihood sensitivity with the ones which minimize the CR bounds. This is done in Table 3. For

Table 2: Most and least informative sets of observables

param.	(c,i,π,r)	(c,i,w,r)	(c,w,π,h)	(y,c,i,w)	(y,c,i,π)	(y,i,r,h)	(y,c,i,h)
ρ_{ga}	1.680	0.794	0.502	0.294	0.372	0.198	0.173
α	0.045	0.040	0.079	0.051	0.076	0.069	0.069
ψ	0.080	0.097	0.229	0.126	0.179	0.145	0.170
β	0.009	0.009	0.018	0.011	0.015	0.018	0.013
φ	2.489	2.813	2.823	6.089	6.817	2.207	2.470
σ_c	0.123	0.228	0.260	0.224	0.506	0.358	0.281
λ	0.022	0.027	0.039	0.038	0.100	0.090	0.128
Φ	0.341	0.638	0.757	0.152	1.521	0.256	0.262
ι_w	0.139	0.376	0.076	0.225	1.037	1.218	1.238
ξ_w	0.134	0.026	0.021	0.021	0.183	0.160	0.409
ι_p	0.039	0.325	0.057	0.198	0.099	1.318	1.137
ξ_p	0.147	0.133	0.112	0.048	0.419	0.309	0.519
σ_l	0.676	0.947	0.190	0.188	2.021	1.324	0.965
r_π	0.250	1.007	0.772	2.959	1.431	1.844	2.505
$r_{\triangle y}$	0.031	0.036	0.159	0.134	0.142	0.116	0.591
r_y	0.048	0.125	0.071	0.336	0.195	0.163	0.777
ρ	0.023	0.055	0.035	0.080	0.089	0.089	0.254
ρ_a	0.015	0.017	0.026	0.025	0.030	0.022	0.023
ρ_g	0.005	0.009	0.015	0.014	0.015	0.021	0.014
ρ_I	0.041	0.055	0.046	0.093	0.098	0.067	0.065
σ_a	0.399	0.091	0.153	0.218	0.395	0.049	0.046
σ_g	0.150	0.136	0.317	0.046	0.088	0.074	0.057
σ_I	0.059	0.078	0.325	0.094	0.097	0.078	0.085
σ_r	0.015	0.017	0.100	0.122	0.231	0.027	0.341
			B. Overall, geometric average				
24 parameters	0.083	0.106	0.119	0.118	0.213	0.157	0.214
17 parameters	0.113	0.160	0.137	0.153	0.310	0.270	0.373
6 parameters	0.042	0.062	0.068	0.121	0.146	0.147	0.346
			C. Overall, D-optimality				
24 parameters	-195	-164	-147	-142	-115	-113	-109

Note: Panel A shows the values of the Cramér-Rao lower bounds for sample size $T = 150$. Panel B shows the geometric averages of the bounds for three groups of parameters: 24 (all free) parameters; 17 (all except shock) parameters; 6 (only λ,ι_p,ξ_p,σ_l,r_π,r_y). Panel C shows the values of $\ln(\det(\boldsymbol{\mathcal{I}}^{-1}))$. Lower values always indicate more information.

eighteen parameters, including four of the six CFM focus on, the likelihood is most sensitive when (c, i, π, r) is observed. For ten of these (c, i, π, r) is also the most informative combination. Overall, for 13 of the 24 parameters the most sensitive and most informative sets coincide. This discrepancy can be explained with the fact that sensitivity is only one of the factors that determines the amount of information. The other factor has to do with how distinct the effect of each parameter on the likelihood is from the effects of the other free parameters. To see why this might matter, consider a situation where there are two parameters which affect very strongly the distribution of a set of variables, but the effects are exactly identical. Even though the likelihood sensitivity is very strong, the two parameters would be unidentifiable from that set of variables. Yet, the parameters may be identified with another set of variables, even thought the sensitivity is much weaker. For more details on this, see Iskrev (2010).

Table 3: Most sensitive and most informative sets of variables

param.	most sensitive	most informative
ρ_{ga}	y, c, i, h	y, w, π, h
α	c, i, π, r	y, i, w, r
ψ	c, i, π, r	c, i, π, r
β	c, i, π, r	c, i, r, h
φ	c, i, π, r	c, i, r, h
σ_c	c, i, π, r	c, i, π, r
λ	c, i, π, r	c, i, π, r
Φ	c, i, π, r	y, w, π, h
ι_w	y, w, π, h	y, c, w, π
ξ_w	c, i, π, r	y, c, w, h
ι_p	y, w, π, h	y, w, π, h
ξ_p	y, w, π, h	y, w, π, h
σ_l	c, i, π, r	y, c, w, h
r_π	c, i, π, r	c, i, π, r
$r_{\triangle y}$	c, i, π, r	c, i, π, r
r_y	c, i, π, r	c, i, π, r
ρ	c, i, π, r	c, π, r, h
ρ_a	c, i, π, r	y, w, π, h
ρ_g	c, i, π, r	c, i, π, r
ρ_I	c, i, π, r	c, i, π, r
σ_a	y, w, π, h	y, w, π, h
σ_g	y, i, w, r	y, c, w, r
σ_I	c, i, π, r	c, i, π, r
σ_r	c, i, π, r	c, i, π, r

Note: The most sensitive set of variables for θ_i is the one minimizing the inverse of the i-th diagonal element of $\boldsymbol{\mathcal{I}}$. The most informative is the one minimizing the i-th diagonal element of $\boldsymbol{\mathcal{I}}^{-1}$.

2.3 Are the results robust to changes in the parameter values and the number of shocks?

The results presented in the last section are conditional on the particular parameter values and the assumptions CFM make regarding the number of shocks and the stationarity of the observables. Here I check whether the optimal selection of observables is robust to changes in the parameter values and the model specification.

I consider three alternatives. First, I change the parameter values keeping the rest of the model as before. Instead of the baseline parametrization, which is similar to the posterior mean in Smets and Wouters (2007), I use the prior mean from that paper. The prior mean is a natural choice since the analysis is supposed to happen prior to estimation. In the other two cases the model is as specified in Smets and Wouters (2007), i.e. featuring seven shocks and deterministic trend, with the trending variables observed in terms of growth rates. I refer to this as the SW specification. That model is evaluated at the means of the prior and the posterior distributions. In order for the results to be comparable with those in the previous section, I maintain the same group of free parameters. Therefore, I assume that the parameters of the three additional shocks as well as the trend parameter are known.

Table 4 shows a summary of the results using the same criteria as before. Clearly, while there is considerable consistency in the ranking across different criteria, the optimal combination of variables is not invariant to the parametrization. Also, the two sets, (y, c, i, h) and (y, c, i, w), recommended by CFM, are consistently ranked among the least informative, especially when the focus is on the six deep parameters.

The optimal combinations of variables for each parameter are shown in Table 5. Panel A reports the results for the CFM specification of the model evaluated at the prior and posterior mean values. In panel B are shown the results for the SW specification. In the case of the CFM model, the optimal combinations of variables is the same for only four parameters. For the SW specification the number of coincidences is 11. There is only one parameter for which the most informative combination is the same across both parameterizations and model specifications. In very few cases the optimal combination of variables includes y, c, and i.

Table 4: Ranking of observables, alternative parameterizations

	geometric average			D-optimality
rank	24 parameters	17 parameters	6 parameters	
	A. CFM model, prior mean			
1	c, w, r, h	c, w, r, h	c, w, r, h	c, i, r, h
2	c, i, r, h	c, i, r, h	c, π, r, h	c, w, r, h
3	y, w, π, r	y, c, w, π	c, i, r, h	y, c, i, w
33	y, c, i, r	y, i, π, r	y, i, w, h	y, c, i, r
34	y, c, i, h	y, c, i, r	c, i, w, h	y, i, r, h
35	y, i, π, r	y, c, i, h	y, c, i, h	y, c, i, h
	B. SW model, prior mean			
1	y, i, h, r	y, i, h, r	c, w, h, π	y, i, h, r
2	y, c, h, r	y, c, h, r	c, h, π, r	y, c, h, r
3	y, i, h, π	i, h, π, r	y, c, h, π	y, c, i, r
33	y, c, i, w	y, c, i, h	y, i, w, h	c, w, h, π
34	c, i, w, h	y, c, i, w	y, c, i, w	c, i, w, π
35	c, i, w, π	c, i, w, π	y, c, i, h	c, w, π, r
	C. SW model, posterior mean			
1	y, i, h, r	y, i, π, r	c, h, π, r	y, i, h, r
2	y, i, π, r	i, w, h, r	c, i, π, r	y, h, π, r
3	y, c, h, r	y, i, h, r	w, h, π, r	y, c, h, r
33	y, w, h, π	c, i, w, π	y, c, i, r	y, c, i, w
34	y, c, i, h	y, c, i, π	y, c, i, w	y, w, h, π
35	c, i, w, π	y, c, i, h	y, c, i, h	y, c, i, h

Note: see note to Table 1.

2.4 The role of interest rate and inflation

One of the main conclusions of CFM is that neither interest rate nor inflation data should be selected, if one has to choose only four of the seven variables. This seems surprising given that four of the six parameters CFM focus on are the price stickiness and price indexation parameters and the inflation and output coefficients in the monetary policy rule. Intuitively, one may expect inflation and interest rate to be very important for the identification of these parameters.

Table 5: Optimal sets, alternative parameterizations

	CFM model		SW model	
param.	prior mean	posterior mean	prior mean	posterior mean
ρ_{ga}	y, i, w, h	y, w, π, h	y, c, i, h	y, c, i, h
α	y, c, i, w	y, i, w, r	c, i, h, r	y, c, i, r
ψ	c, i, r, h	c, i, π, r	y, i, w, h	y, c, i, w
β	c, i, r, h	c, i, r, h	c, h, π, r	c, h, π, r
φ	c, i, π, h	c, i, r, h	c, i, h, r	c, i, h, r
σ_c	c, i, r, h	c, i, π, r	c, i, h, r	c, i, h, r
λ	c, i, r, h	c, i, π, r	c, i, h, r	c, i, h, r
Φ	y, w, π, h	y, w, π, h	c, i, w, h	y, i, h, r
ι_w	c, i, w, π	y, c, w, π	c, w, h, π	c, w, h, π
ξ_w	y, w, π, r	y, c, w, h	c, w, h, π	c, w, h, r
ι_p	c, w, π, h	y, w, π, h	y, w, h, π	w, h, π, r
ξ_p	y, w, π, h	y, w, π, h	y, w, h, π	y, w, h, π
σ_l	c, w, r, h	y, c, w, h	c, w, h, π	c, w, h, r
r_π	c, i, r, h	c, i, π, r	y, h, π, r	c, i, π, r
$r_{\triangle y}$	c, w, r, h	c, i, π, r	y, i, π, r	c, i, h, r
r_y	c, i, r, h	c, i, π, r	c, h, π, r	c, i, π, r
ρ	c, π, r, h	c, π, r, h	y, h, π, r	c, h, π, r
ρ_a	c, i, r, h	y, w, π, h	y, i, w, h	y, i, h, r
ρ_g	y, c, i, w	c, i, π, r	y, c, i, h	y, c, h, r
ρ_I	c, i, r, h	c, i, π, r	c, i, h, r	c, i, h, r
σ_a	y, i, w, h	y, w, π, h	y, i, w, h	y, i, w, h
σ_g	y, w, π, r	y, c, w, r	y, c, i, h	y, c, i, h
σ_I	c, i, r, h	c, i, π, r	c, i, h, r	c, i, w, r
σ_r	c, π, r, h	c, i, π, r	c, h, π, r	c, h, π, r

Note: The table shows the most informative sets of variables for each parameter. The posterior mean parametrization of the CFM model is identical to the baseline case analyzed in section 2.2

One way to assess the importance of a variable is to measure the loss of information when the variable is excluded relative to when it is included as an observable. Table 6 shows, for the six deep parameters CFM focus on, the ratios of the Cramér-Rao lower bounds (CRLB) with one variable excluded relative to bounds with all seven variables observed. The largest number in each row shows which variable is most informative for the parameter in that row. According to that measure, inflation is the most informative variable for three parameters - ι_p, r_π and r_y. Interest rate and consumption are the most informative variables for λ, while consumption and hours are the most informative variables for σ_l. Lastly, wages is the most informative variable for ξ_p. Overall, the large numbers in the first two column of the table suggest that there is a significant loss of information from excluding π and r.

Table 6: Loss of information, posterior mean of SW model

	excluded variable						
	r	π	h	w	i	c	y
λ	1.8	1.1	1.4	1.2	1.4	1.8	1.2
ι_p	1.0	3.7	1.1	3.0	1.0	1.0	1.1
ξ_p	1.1	2.1	2.3	2.5	1.3	1.2	1.3
σ_l	1.4	1.2	1.7	1.2	1.2	1.7	1.1
r_π	1.8	2.2	1.2	1.1	1.1	1.2	1.0
r_y	1.8	2.1	1.2	1.1	1.1	1.1	1.0

Note: The numbers shown are the CRLBs with one excluded variable relative to the CRLBs with all 7 variables.

3 Concluding remarks

The results can be summarized as follows: (1) The rank condition for identification is not informative about the optimal choice of observables for the model analyzed. In general, it could be useful if only one or a very few of the many possible sets of variables satisfy the identification condition. This seems to rarely be the case in practice. (2) At the baseline parametrization of the model the most informative set of variables includes consumption, investment, interest rate, and inflation. (3) The most informative set of variables is not invariant to the parametrization of the model or the number of shocks.

All of these findings call into questions the results in CFM. Perhaps the most important one is (3), which implies that to select the best set of variables to use in estimation it is not sufficient to determine what they are at a single point in the parameter space. A reasonable solution in a Bayesian context would be to base the selection of variables on the expected value of the preferred criterion, with the expectation taken over the prior distribution of the parameters.

References

Bierens, H. J. (2007): "Econometric analysis of linearized singular dynamic stochastic general equilibrium models," *Journal of Econometrics*, 136, 595–627.

Canova, F., F. Ferroni, and C. Matthes (2014): "Choosing the variables to estimate singular DSGE models," *Journal of Applied Econometrics, forthcoming.*

Iskrev, N. (2010): "Evaluating the strength of identification in DSGE models. An a priori approach," Working paper series, Banco de Portugal.

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

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