Kleibergen, F., and Mavroeidis, S. (2008), "Inference on Subsets of Parameters in GMM Without Assuming Identification," working paper, Brown University.

Kleibergen, F., and Paap, R. (2006), "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 133, 97–126.

Krause, M., Lubik, T. A., and Lopez-Salido, D. (2008), "Inflation Dynamics With Search Frictions: A Structural Econometric Analysis," Working Paper 08-01, Richmont Fed.

Kuester, K., Mueller, G., and Stoelting, S. (2009), "Is the New Keynesian Phillips Curve Flat?" *Economics Letters*, 103, 39–41.

Lewbel, A. (1991), "The Rank of Demand Systems: Theory and Nonparametric Estimation," *Econometrica*, 59, 711–730.

Lubik, T. A., and Schorfheide, F. (2004), "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94 (1), 190–216.

Lucas, R. E. J. (1976), "Econometric Policy Evaluation: A Critique," in *The Philips Curve and Labor Markets. Carnegie–Rochester Conference Series on Public Policy*, eds. K. Brunner and A. Meltzer, Amsterdam: North-Holland.

Ma, A. (2002), "GMM Estimation of the New Keynesian Phillips Curve," *Economics Letters*, 76, 411–417.

Magnusson, L. M., and Mavroeidis, S. (2009), "Identifying Euler Equation Models via Stability Restrictions," working paper, Brown University.

Martins, L. F., and Gabriel, V. J. (2006), "Robust Estimates of the New Keynesian Phillips Curve," Discussion Paper 0206, University of Surrey, Dept. of Economics.

Mavroeidis, S. (2005), "Identification Issues in Forward-Looking Models Estimated by GMM With an Application to the Phillips Curve," *Journal of Money Credit and Banking*, 37 (3), 421–449.

——— (2006), "Testing the New Keynesian Phillips Curve Without Assuming Identification," Economics Working Paper 2006-13, Brown University, available at *http://ssrn.com/abstract=905261*.

Mavroeidis, S., Chevillon, G., and Massmann, M. (2008), "Inference in Models With Adaptive Learning, With an Application to the New Keynesian Phillips Curve," working paper, Brown University.

McConnell, M. M., and Perez-Quiros, G. (2000), "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?" *The American Economic Review*, 90 (5), 1464–1476.

Moreira, M. J., (2003), "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica*, 71, 1027–1048.

Nason, J. M., and Smith, G. W. (2008), "Identifying the New Keynesian Phillips Curve," *Journal of Applied Econometrics*, 23 (5), 525–551.

Newey, W. K., and McFadden, D. (1994), "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, Vol. 4, eds. R. Engle and D. McFadden, Amsterdam: North-Holland, Chapter 36, pp. 2113–2148.

Newey, W. K., and West, K. D. (1987), "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55 (3), 703–708.

Newey, W., and Windmeijer, F. (2009), "GMM With Many Weak Moment Conditions," *Econometrica*, 77 (3), 687–719.

Pesaran, M. H. (1987), *The Limits to Rational Expectations*, Oxford: Blackwell Publishers.

Phillips, P. C. B. (1983), "Exact Small Sample Theory in the Simultaneous Equations Model," in *Handbook of Econometrics*, Vol. 1, eds. Z. Griliches and M. Intrilligator, Amsterdam: North-Holland.

Robin, J.-M., and Smith, R. J. (2000), "Tests of Rank," *Econometric Theory*, 16, 151–175.

Robins, J. M. (2004), "Optimal Structural Nested Models for Optimal Sequential Decisions," in *Proceedings of the Second Seattle Symposium on Biostatistics*, eds. D. Y. Lin and P. Heagerty, New York: Springer.

Rothenberg, T. J. (1984), "Approximating the Distributions of Econometric Estimators and Test Statistics," in *Handbook of Econometrics*, Vol. 2, eds. Z. Griliches and M. D. Intrilligator, Amsterdam: North-Holland, Chapter 15, pp. 881–935.

Rudd, J., and Whelan, K. (2005), "New Tests of the New-Keynesian Phillips Curve," *Journal of Monetary Economics*, 52 (6), 1167–1181.

——— (2006), "Can Rational Expectations Sticky-Price Models Explain Inflation Dynamics?" *American Economic Review*, 96 (1), 303–320.

——— (2007), "Modelling Inflation Dynamics: A Critical Survey of Recent Research," *Journal of Money Credit and Banking*, 39, 155–170.

Sbordone, A. M. (2002), "Prices and Unit Labor Costs: A New Test of Price Stickiness," *Journal of Monetary Economics*, 49, 265–292.

Smets, F., and Wouters, R. (2007), "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *AER*, 97 (3), 586–606.

Stock, J., and Watson, M. (1999), "Forecasting Inflation," *Journal of Monetary Economics*, 44 (2), 293–335.

——— (2008), "Phillips Curve Inflation Forecasts," Working Paper 14322, National Bureau of Economic Research.

Stock, J. H., and Wright, J. H. (2000), "GMM With Weak Identification," *Econometrica*, 68, 1055–1096.

Stock, J. H., and Yogo, M. (2005), "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, eds. D. W. K. Andrews and J. H. Stock, Cambridge: Cambridge University Press, pp. 80–108.

Stock, J. H., Wright, J. H., and Yogo, M. (2002), "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20, 518–530.

West, K. D. (1997), "Another Heteroscedasticity- and Autocorrelation-Consistent Covariance Matrix Estimator," *Journal of Econometrics*, 76, 171–191.

White, H. (1984), *Asymptotic Theory for Econometricians*, Orlando, FL: Academic Press.

Woodford, M. (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton, NJ: Princeton University Press.

Zivot, E., Startz, R., and Nelson, C. R. (1998), "Valid Confidence Intervals and Inference in the Presence of Weak Instruments," *International Economic Review*, 39, 1119–1144.

Comment

Fabio Canova

ICREA-UPF, AMeN, and CEPR, Department of Economics, Trias Fargas 25-27, 08005 Barcelona, Spain (*fabio.canova@upf.edu*)

I discuss the identifiability of a structural New Keynesian Phillips curve when it is embedded in a small-scale dynamic stochastic general equilibrium model. Identification problems emerge because not all the structural parameters are recoverable from the semistructural ones and the objective functions I consider are poorly behaved. The solution and the moment mappings are responsible for the problems.

1. INTRODUCTION

Kleinbergen and Mavroeidis (KM) have written an excellent article, compactly reviewing what we know about the identification of the parameters of a New Keynesian Phillips curve when estimated by the generalized method of moments (GMM) and including interesting Monte Carlo evidence to shed light on the properties of various identification robust methods proposed in the literature. This comment takes on two issues of interest for applied macroeconomists that the paper has left on the back burner: Nowadays structural Phillips curves are typically considered, as opposed to the semistructural Phillips curves that KM use; for policy exercises, a Phillips curve is typically

Journal of Business & Economic Statistics
July 2009, Vol. 27, No. 3
DOI: 10.1198/jbes.2009.08262

embedded into a small- or medium-scale general equilibrium (DSGE) model. Therefore, the identification of its parameters requires a system-wide rather than a single equation perspective.

To discuss these issues I will first write a canonical small-scale structural model that constitutes the backbone of those medium-scale models currently used in policy institutions for forecasting and policy evaluation. I will then discuss the difference between the structural and the semistructural versions of such a model and examine the identification of the parameters when impulse responses or likelihood-based methods are used to construct the objective function.

I want to stress that this comment is concerned with population identification problems. That is, the problems I highlight are intrinsic to the theory rather than specific to a dataset or a sample. Their solutions therefore require alterations of the theory rather than the acquisition of better or longer datasets and/or a careful selection of objective functions to be optimized.

2. A PROTOTYPE SMALL–SCALE NEW KEYNESIAN MODEL

The baseline model I consider has log-linearized optimality conditions of the form:

$$y_t = \frac{h}{1+h} y_{t-1} + \frac{1}{1+h} E_t y_{t+1} + \frac{1}{\phi}(i_t - E_t \pi_{t+1}) + v_{1t}, \tag{1}$$

$$\pi_t = \frac{\omega}{1+\omega\beta}\pi_{t-1} + \frac{\beta}{1+\omega\beta}E_t\pi_{t+1} + \frac{(\phi+\nu)(1-\zeta\beta)(1-\zeta)}{(1+\omega\beta)\zeta} y_t + v_{2t}, \tag{2}$$

$$i_t = \lambda_r i_{t-1} + (1-\lambda_r)(\lambda_\pi \pi_{t-1} + \lambda_y y_{t-1}) + v_{3t}, \tag{3}$$

where h is the degree of habit persistence, ϕ the relative risk aversion coefficient, β the discount factor, ω the degree of price indexation, ζ the degree of price stickiness, ν the elasticity of labor supply, while $\lambda_r, \lambda_\pi, \lambda_y$ are monetary policy parameters. v_{1t} and v_{2t} are AR(1) processes with parameters ρ_1, ρ_2, while v_{3t} is iid. The variances of the shocks are denoted by $\sigma_i^2, i = 1, 2, 3$. Equation (1) is a log-linearized Euler condition; the second is a version of a New Keynesian Phillips curve obtained by log-linearizing the optimal pricing decision around a zero steady state inflation; and the third is a policy rule. The model has 14 structural parameters: $\theta_1 = (h, \phi, \beta, \omega, \nu, \zeta, \lambda_r, \lambda_\pi, \lambda_y)$ are economic parameters and $\theta_2 = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \rho_1, \rho_2)$ are auxiliary parameters. While the specification is rather standard, two features of (1)–(3) are worth discussing. First, the policy rule is backward looking—this allows us to name v_{3t} as a monetary policy innovation. Second, there is habit in consumption, a feature typically absent from basic versions of the theory, but always included in the larger-scale structures.

The semistructural version of the model eschews the cross-equation restrictions that the theory imposes on the coefficients and is of the form:

$$y_t = a_1 y_{t-1} + a_2 E_t y_{t+1} + a_3(i_t - E_t\pi_{t+1}) + v_{1t}, \tag{4}$$

$$\pi_t = a_4\pi_{t-1} + a_5 E_t\pi_{t+1} + a_6 y_t + v_{2t}, \tag{5}$$

$$i_t = a_7 i_{t-1} + a_8\pi_{t-1} + a_9 y_{t-1} + v_{3t}. \tag{6}$$

Note that (5) corresponds to the specification used by KM. This version of the model also has 14 parameters, $\alpha = (a_1, \ldots, a_9)$ and $\theta_2 = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \rho_1, \rho_2)$ but, following the logic of rank and order conditions, one can see that even when all the parameters of (4)–(6) are identifiable it is impossible to recover all the θ_1 from estimates of the a's—ζ and ν enter multiplicatively and only in the slope parameter a_6, while a_1 and a_2 contain information only about h. Hence, conditioning on a model where variables are expressed in deviation from the steady state and absent external information, it is impossible to examine the structural determinants of the slope of the Phillips curve and, consequently, back out estimates of the frequency of price adjustments, ζ. To solve this problem it is necessary to specify additional equations that allow the elasticity of labor supply ν to be identifiable. For example, one could solve the model around a flexible price equilibrium rather than the steady state and add the definition of flexible output to the system of equations.

3. MAPPING THE SEMISTRUCTURAL MODEL INTO A POPULATION OBJECTIVE FUNCTION

Local identification of the parameters of the model (4)–(6) requires that the objective function have a unique extremum in correspondence with the true parameter vector; that the Hessian of the objective function be of full rank in the neighborhood of the true parameter vector; and that the curvature of the objective function in the neighborhood of the true parameter vector be sufficient to translate the objective function information into parameters information.

Absent the first condition, models with different theoretical features may be observationally equivalent given a particular objective function (see Sargent 1978; Kennan 1988; Neely, Roy, and Whiteman 2001; Kim 2003; Beyer and Farmer 2004; and Canova and Sala 2006; among others). Clearly, observational equivalence crucially depends on the selected objective function. The second condition ensures that under-identification of pathologies where the objective function is insensitive to variations in one or more parameters (see Choi and Phillips 1992 and Canova and Sala 2009) will be absent. For example, the belief that the discount factor β is hard to estimate with cyclical data in a real business cycle model can be formalized by showing that the rank of the Hessian of the objective function is deficient for any true $\beta \in [0.96, 0.9999]$.

The first two conditions rule out somewhat extreme kinds of identification pathologies. The third safeguards against more subtle weak and partial identification problems. Deficiencies in the curvature of the objective function in the neighborhood of the true parameter vector imply that parameter changes only marginally affect the objective function—it is either nearly flat in some dimensions (weak identification) or displays ridges (partial identification).

The mapping from the parameters of the model (4)–(6) to a given objective function may fail to meet these three necessary criteria for identification because three types of transformations are needed to go from the former to the latter. First, the model needs to be solved—this involves a nonlinear and typically numerical transformation. Second, some sufficient statistic (unconditional moments or impulse responses) is computed to summarize the informational content of the solution—this is another nonlinear transformation. Third, an objective function expressing the distance between model-based and actual summary statistics is constructed—this can be a highly nonlinear transformation if, for example, one compares turning points of economic activity. When likelihood-based methods are used, the last two steps are combined and the VAR(1) solution is used directly to construct the likelihood or the kernel of the posterior. When some variables appearing in the solution are omitted because, for example, they are unobservable, the solution for the observables is an ARMA(∞, ∞) (see Canova 2007) so one extra nonlinear transformation is needed.

It is difficult to study in theory how these nonlinear transformations repackage the information contained in the parameters. However, one can use graphical and exploratory analyses to detect problems. To compare my conclusions with those of KM, I will focus attention solely on the identification of a_4, a_5, and a_6, which give us information about the structural parameters β, ω and, given estimates of a_3, about the conglomerate of ζ and ν. To make the discussion concrete, I choose the true parameter vector θ to be $\beta = 0.985$, $\phi = 2.0$, $\nu = 1.0$, $\zeta = 0.68$, $\omega = 0.70$, $h = 0.85$, $\rho_r = 0.2$, $\rho_y = 1.1$, $\rho_\pi = 1.5$, $\rho_1 = 0.65$, $\rho_2 = 0.65$, $\sigma_1 = 0.003$, $\sigma_2 = 0.002$, $\sigma_3 = 0.001$, in line with the estimates of Rabanal and Rubio-Ramirez (2005). These values imply that $a_4 = 0.4143$, $a_5 = 0.5830$, $a_6 = 0.2759$ are the true values of the parameters of interest.

I consider three objective functions: one measures the distance of responses to monetary policy shocks—twenty equally weighted responses of the three variables are used; the second is the likelihood function, constructed under normality of the disturbances; the third is the kernel of the posterior, obtained using informative priors for the structural parameters entering a_4, a_5, and a_6 and centered at the true values with small spreads.

4. ARE THE PARAMETERS OF THE PHILLIPS CURVE THEORETICALLY IDENTIFIABLE?

For this class of models and for my choice of "true" θ, all the objective functions have a unique local extremum. Five of the eigenvalues of the Hessian of the distance function are exactly zero—those corresponding to ρ_1 and ρ_2, which are under-identified from monetary policy shocks, and those corresponding to σ_i, $i = 1, 2, 3$, which are under-identified from any scaled impulse responses. The other two objective functions have no eigenvalue with this feature. Nevertheless, six eigenvalues of both the Hessian of the distance function and of the likelihood function are small relative to the average eigenvalue—weak and partial identification problems could be present. To examine whether these eigenvalues are associated with a_4, a_5, and a_6, I graphically explore how the objective functions change when these parameters vary in the neighborhood of the true parameter vector (see the range presented in the x-axis in Figure 1), keeping all other parameters fixed at their true values.

The distance function is rather flat in all dimensions (the elasticity is always smaller than 0.1) and somewhat asymmetric in a_4. When plotted in two dimensions it is still very flat, particularly for a_5, which is the forward-looking parameter of the Phillips curve. The log-likelihood function, which contains all

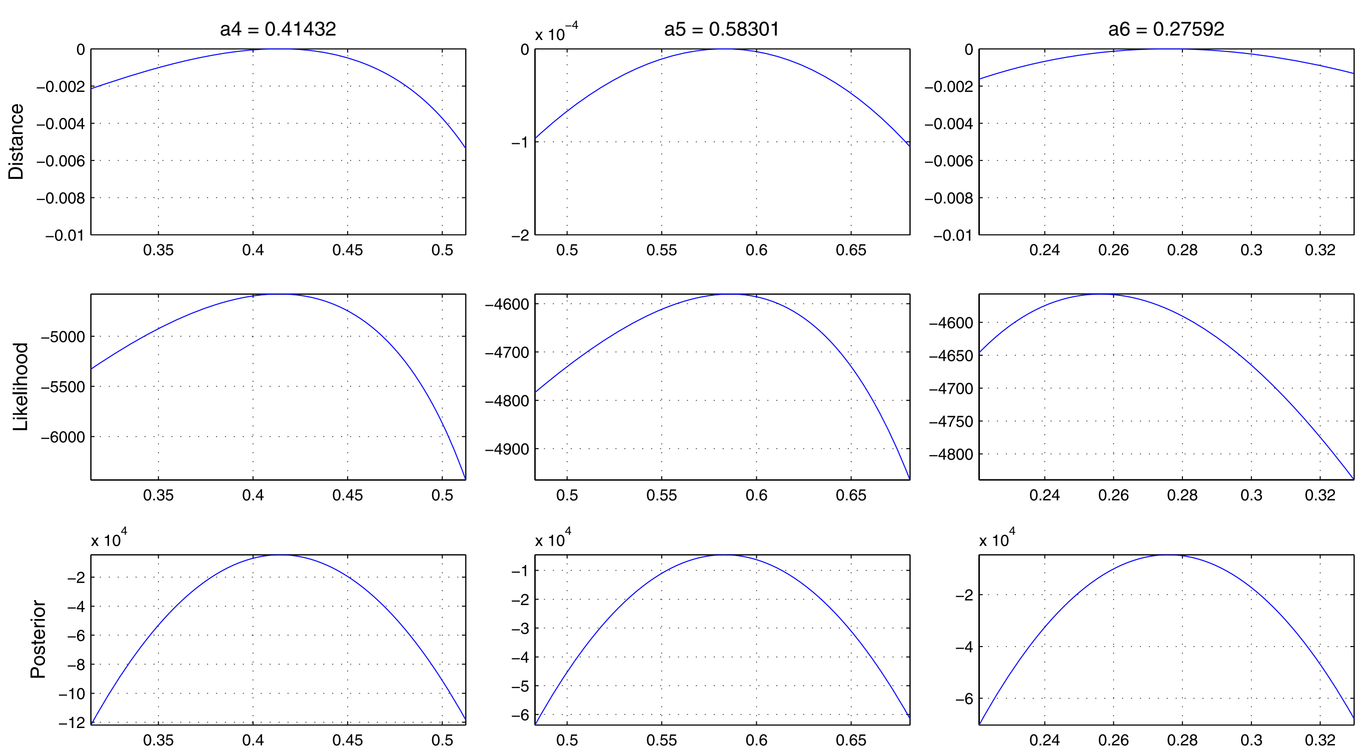

Figure 1. Shape of different objective functions.

the information of the model, is better behaved except for the marked asymmetry it displays in all dimensions. When plotted in two dimensions it has sufficient curvature in both a_4, a_5 and a_5, a_6 but displays diagonal ridges—a_4, a_5, and a_6 are not separately identifiable. The log posterior kernel, instead, is nicely peaked in all dimensions. Since the priors used in structural estimation are conventionally centered at calibrated values and with tight spreads—as we have done here—it is not difficult to see that the prior may determine the shape of the posterior.

These visual impressions are confirmed using the relative size of the eigenvalues of the Hessian of the objective function at the true parameters. For example, the eigenvalues of the Hessian of the distance associated with a_4 and a_6 are of the order of 10% of the average eigenvalue of both matrices and the one associated with a_5 is smaller than 0.001% of the average eigenvalue.

In conclusion, both the distance and the likelihood functions will be unable to appropriately identify the forward-looking parameter of the Phillips curve, but for different reasons. The distance function will not work because identification of a_5 is weak. The likelihood function will not work because a_5 is linearly related to the other parameters of the Phillips curve.

Which mapping is responsible for these information deficiencies? The solution and moment mappings both contribute. In the solution mapping, four of the nine eigenvalues are smaller than 0.20 of the average eigenvalue, while with the moment mapping two additional eigenvalues are smaller than 0.20 of the average. Since the smallest eigenvalue of the solution mapping is the one associated with a_5, identification of this parameter is difficult unless the model or the way it is solved is changed. Note that the use of higher-order approximations does not guarantee better identification properties in the population (see, e.g., Canova and Sala 2006).

5. ESTIMATION

Since it is unlikely that applied investigators will spend time altering the theory or refining their numerical solution techniques, estimation methods that work when identification problems exist are needed. While KM have made it clear that identification robust methods exist in the single equation GMM literature, no procedure has been devised for likelihood-based methods. Furthermore, while impulse response matching estimators share similarities with GMM, failure to use the continuously updating weighting matrix in the estimation precludes a direct extension of the GMM results.

In this section, I first show what identification problems imply when nonidentification robust methods are used to estimate the parameters of the Phillips curve and then I use ideas of the literature that KM review to construct estimates of the parameters of interest. The punchline is that when weak and partial identification problems are present, standard methods produce erratic estimates and meaningless standard errors, even in extremely large samples. However, estimation intervals obtained by inverting the objective function are practically identical to those obtained with standard methods because the distance function is extremely flat in many dimensions (compare with Nason and Smith 2008). This is perhaps unsurprising since the distance function I use is not a robust objective function in the sense of KM.

The exercise is as follows. Given the correct model and 500 initial conditions in the neighborhood of the true parameter vector, I estimate a_4, a_5, and a_6 using a distance function that measures how far output gap, inflation, and the nominal rate responses to monetary shocks in the model are from the true ones. Figures 2 and 3 present the histograms of initial and final estimates for two different choices of the true parameter vector; the

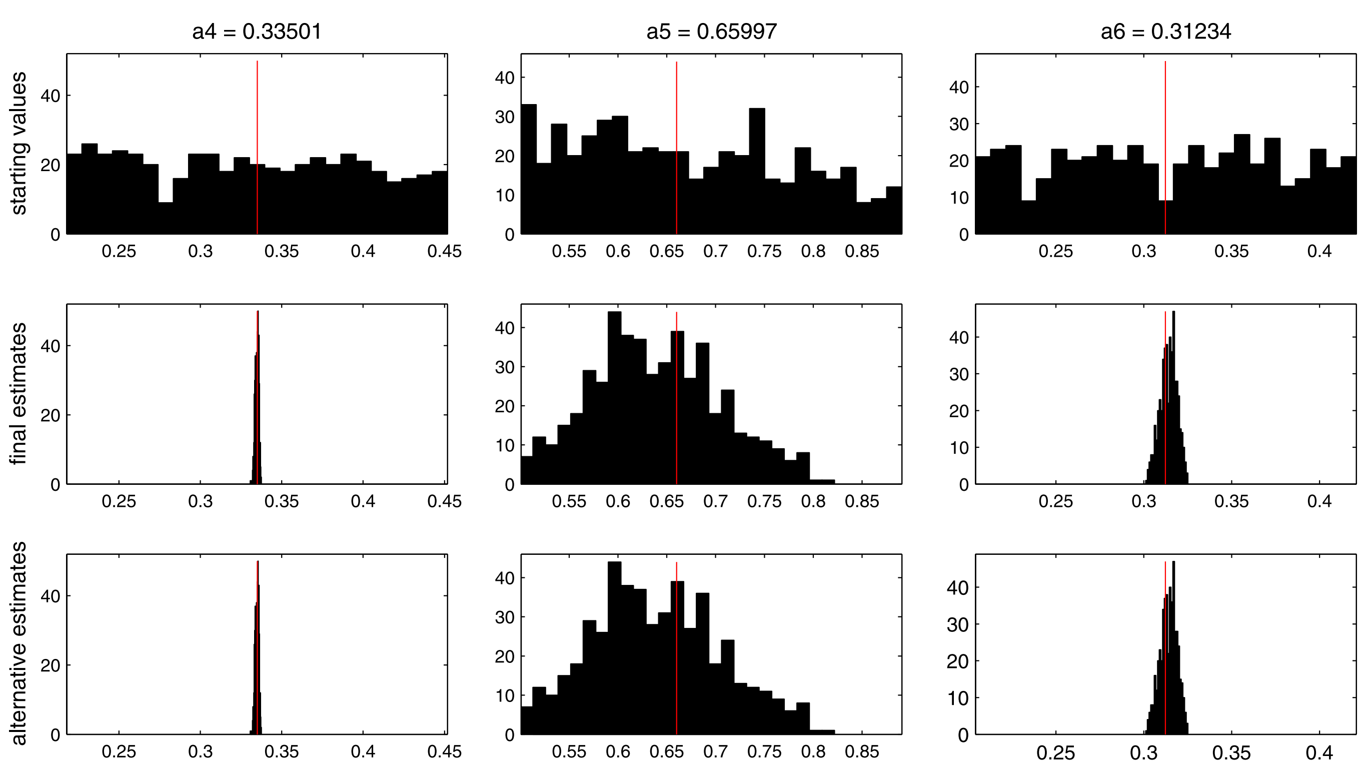

Figure 2. Histogram of initial conditions and estimates.

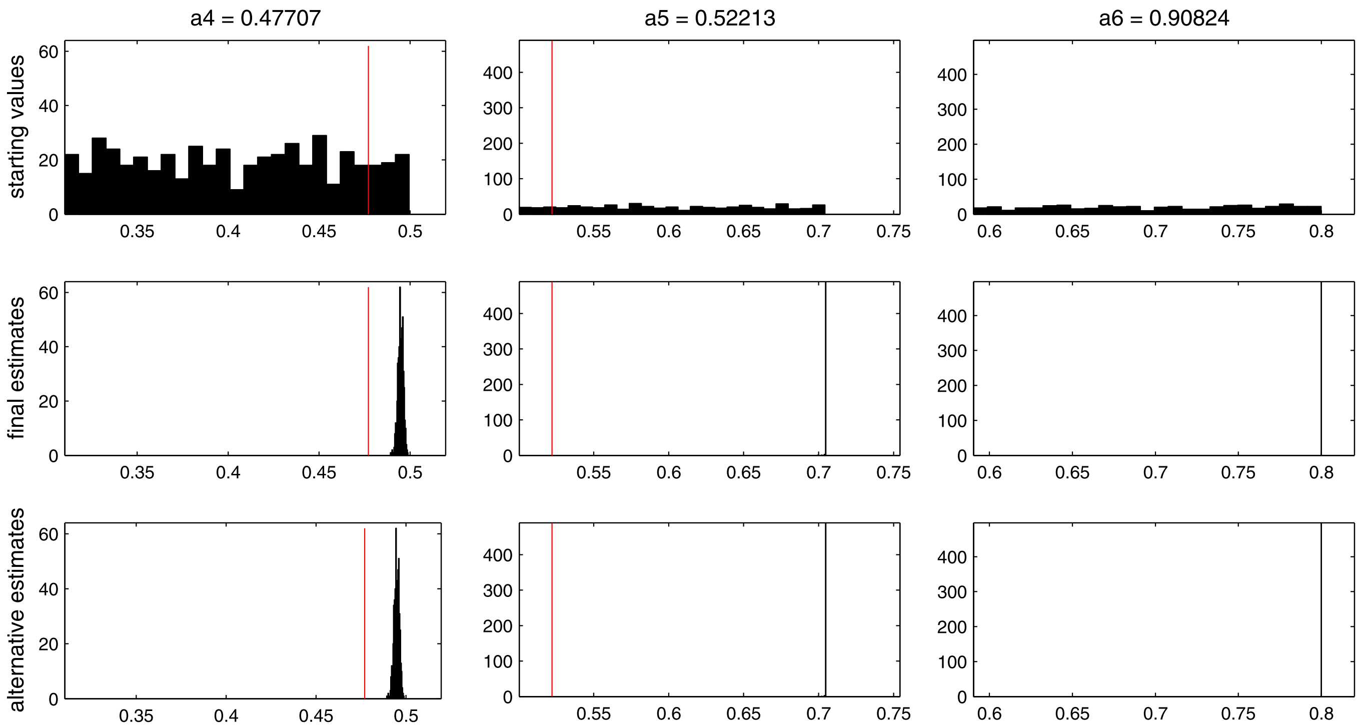

Figure 3. Histogram of initial conditions and estimates.

x-axis shows the range for the chosen initial conditions. When the forward-looking component of the Phillips curve is strong and the slope is economically different from zero, problems are concentrated in a_5—the range of final estimates of a_5 is only marginally smaller than the range of initial conditions and the correlation between initial conditions and final estimates is high (around 0.7).

When the forward- and the backward-looking components are roughly similar and marginal costs are important for inflation, estimates of all semistructural parameters are always away from the true parameters, the sum of estimates of a_4 and a_5 always exceeds one, the slope of the Phillips curve is systematically underestimated, and the bias is large (order of 10%–25%). When sample rather than population objective functions are available, all these problems could be greatly magnified.

Figures 2 and 3 show that the range of estimates of a_5 obtained by inverting the objective function is practically identical to the one obtained with standard minimum distance estimators—out of the 500 cases only 5 are eliminated. This occurs because, in all the simulations I have run, the value of the objective function at the estimates is close to the median value of the $\chi^2(51)$ distribution. This could have been expected. In Figure 1 the objective function is so flat in a_5 that estimates in the range [0.45, 0.80] only very marginally change its value.

To conclude, the problems that KM highlighted in their excellent review get compounded when the New Keynesian Phillips curve is embedded into a small scale DSGE model and multivariate estimation techniques are considered (see also Cochrane 2007); there are additional headaches for applied investigators when structural rather than semistructural estimation is attempted. The solution mapping seems to be responsible for the identification difficulties. Poorly behaved solution mappings are especially problematic because they leave applied investigators with no choice other than to respecify the structure they wish to estimate or refine their solution procedure.

ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support of the Spanish Ministry of Education through the grant SEJ-2004-21682-E and of the Barcelona Economic Program (CREA).

ADDITIONAL REFERENCES

Beyer, A., and Farmer, R. (2004), "On the Indeterminacy of New Keynesian Economics," Working Paper 323, ECB.

Canova, F. (2007), *Methods for Applied Macroeconomic Research*, Princeton, NJ: Princeton University Press.

Canova, F., and Sala, L. (2006), "Back to Square One: Identification Issues in DSGE Models," Working Paper 583, ECB.

Choi, I., and Phillips, P. C. (1992), "Asymptotic and Finite Sample Distribution Theory for IV Estimators and Tests in Partially Identified Structural Equations," *Journal of Econometrics*, 51, 113–150.

Cochrane, J. (2007), "Identification and Price Determination With Taylor Rules: A Critical Review," manuscript, University of Chicago.

Kennan, J. (1988), "An Econometric Analysis of the Fluctuations in Aggregate Labor Supply and Demand," *Econometrica*, 45, 969–990.

Kim, J. (2003), "Functional Equivalence Between Intertemporal and Multisectoral Investment Adjustment Costs," *Journal of Economic Dynamics and Control*, 27, 533–549.

Neely, C., Roy, A., and Whiteman, C. (2001), "Risk Aversion versus Intertemporal Substitution: A Case Study of Identification Failures in the Intertemporal Capital Asset Pricing Model," *Journal of Business & Economic Statistics*, 19, 295–403.

Rabanal, P., and Rubio-Ramirez, J. (2005), "Comparing New-Keynesian Models of the Business Cycle: A Bayesian Approach," *Journal of Monetary Economics*, 52, 1150–1162.

Sargent, T. (1978), "Estimation of Dynamic Labor Demand Schedules Under Rational Expectations," *Journal of Political Economy*, 86, 1009–1044.