

Comparison of Misspecified Calibrated Models: The Minimum Distance Approach

Viktoriya Hnatkovska*, Vadim Marmer*, and Yao Tang†

August 28, 2011

Abstract

This paper proposes several testing procedures for comparison of misspecified calibrated models. The proposed tests are of the Vuong-type (Vuong, 1989; Rivers and Vuong, 2002). In our framework, the econometrician selects values for model's parameters in order to match some characteristics of data with those implied by the theoretical model. We assume that all competing models are misspecified, and suggest a test for the null hypothesis that they provide equivalent fit to data characteristics, against the alternative that one of the models is a better approximation. We consider both nested and non-nested cases. We also relax the dependence of models' ranking on the choice of a weight matrix by suggesting averaged and sup-norm procedures. The methods are illustrated by comparing the cash-in-advance and portfolio adjustment cost models in their ability to match the impulse responses of output and inflation to money growth shocks.

JEL classification: C51; C52

Keywords: misspecified models; calibration; minimum distance estimation

*Department of Economics, University of British Columbia, 997 - 1873 East Mall, Vancouver, BC V6T 1Z1, Canada. E-mail addresses: hnatkovs@mail.ubc.ca (Hnatkovska), vadim.marmer@ubc.ca (Marmer).

†Department of Economics, Bowdoin College, 9700 College Station, Brunswick, Maine 04011-8497, USA. Email: ytang@bowdoin.edu.

1 Introduction

This paper proposes a method of formally comparing calibrated models. Calibrationists often acknowledge that their models are misspecified and will be rejected by data if formally tested (Canova, 1994). Their objective is not to assess whether a particular model is true, but rather which features of data it can be used to capture. Furthermore, calibrationists are often interested in learning which of several competing models provides a better fit to data.

In a typical calibration exercise, the calibrationist selects values for a model's parameters in order to match some characteristics of observed data with those implied by the theoretical model. For example, a model can be calibrated to match empirical moments, cross-correlations, or impulse responses. Such characteristics will be referred to as properties of a reduced-form model or the reduced-form parameters, since they can be consistently estimated from data without knowing the true data generating process (DGP).

After choosing parameter values, a structural model is evaluated by comparing model-implied reduced-form characteristics with those of actual data (Gregory and Smith, 1991, 1993; Cogley and Nason, 1995; Kim and Pagan, 1995). However according to the calibration approach, a model is only an approximation and therefore should not be regarded as a null hypothesis to be statistically tested (Prescott, 1991). Instead of testing whether a specification is correct, one can evaluate a model by comparing it against some benchmark model. This approach has been advocated in Diebold and Mariano (1995), West (1996), White (2000), and other papers.

In this paper, we view assigning parameters' values in calibration as an example of classical minimum distance (CMD) estimation. We then compare models by the means of an asymptotic test. Following the approach of Vuong (1989) and Rivers and Vuong (2002), RV hereafter, our null hypothesis states that two misspecified models provide an equivalent approximation to data. Vuong (1989) proposed such a test in the maximum likelihood framework, and RV discussed it in a more general setting allowing for a broad class of lack-of-fit criteria including that of GMM.

The contribution of our paper relative to RV is as follows. While RV focused solely on non-nested models, the CMD framework allows us to analyze both non-nested and nested cases. The nested case is particularly important because, if the null of model equivalence is not rejected, one can replace a bigger model with a more parsimo-

nious one. Furthermore, many hypotheses can be expressed in terms of parameter restrictions and thus fall into the nested category. The nested and non-nested cases also correspond to different asymptotic null-distributions. Focusing on the non-nested case, RV describes generic conditions sufficient for the difference between sample lack-of-fit criteria of two models to have an asymptotically normal null distribution. We derive a similar result for the non-nested case in our framework, however, we also show that our test statistic has a mixed χ^2 asymptotic null-distribution in the nested case.¹

We also address the issue of choosing weights for reduced-form characteristics when comparing models. When models are misspecified, pseudo-true values of their parameters and ranking of the models depend on the choice of the weighting scheme. In this paper, we relax the dependence of model ranking on the choice of the weighting scheme by suggesting averaged and sup-norm procedures. The averaged test corresponds to the null hypothesis that two models have equal lack-of-fit on average. On the other hand, according to the null hypothesis of the sup-norm test, one model cannot outperform another for any choice of the weight matrix.

Comparison of misspecified models has been studied recently from the Bayesian perspective by Schorfheide (2000), and using a Kolmogorov-type test by Corradi and Swanson (2007). Kan and Robotti (2009) use the Hansen-Jagannathan distance to design a Vuong-type test for comparison of misspecified asset pricing models. Misspecification has been also addressed by Dridi et al. (2007) using indirect inference.

We apply our methodology to compare two standard monetary business cycle models: the cash-in-advance (CIA) model, and the Lucas (1990) and Fuerst (1992) model with portfolio adjustment costs (PAC). The two models have the same underlying structure except for the information sets that agents possess when making their decisions. We judge the performance of the models based on their ability to replicate the response of inflation and output growth to an unanticipated monetary shock in the U.S. data. A structural vector autoregression (SVAR) is employed to obtain model-free estimates of such impulse responses. We find that the null hypothesis that CIA and PAC models have the same fit cannot be rejected. We therefore conclude that the frictions underlying the PAC model do not play a significant role in approximating the inflation-output dynamics.

The paper proceeds as follows. Section 2 introduces the framework. Section

¹This is consistent with the results of Vuong (1989) obtained in the MLE framework.

3 describes our basic tests for model comparison based on a quasi-likelihood ratio (QLR) statistic. We discuss the distribution of the suggested statistic in the cases of nested, strictly non-nested and overlapping models. Section 4 discusses the averaged and sup-norm tests, and confidence sets for weighting schemes. Section 5 illustrates the technique with an empirical application. The proofs and additional details are given in the supplement to this paper.² We use the following notation. For a matrix B , let $\|B\| = \sqrt{\text{tr}(B'B)}$, where tr denotes trace. When B is symmetric and positive definite, $B^{1/2}$ denotes the symmetric square-root matrix.

2 Definitions

Let $Y_n(\omega)$ be a data matrix of the sample size n defined on the probability space (Ω, \mathcal{F}, P) . All random quantities in this paper are some functions of the data Y_n . We use h to denote an m -vector of reduced-form characteristics (parameters). Its true value, $h_0 \in R^m$, depends on the true unknown structural model of the economy and its parameters. For example, h can be a vector of moments, cross-correlations, impulse responses, etc. While the true structural model is unknown, we will assume that the reduced-form parameter h can be estimated consistently from the data. Let \hat{h}_n denote an estimator of h . We assume that \hat{h}_n has the following properties.

Assumption 2.1. (a) $n^{1/2}(\hat{h}_n - h_0) \rightarrow_d N(0, \Lambda_0)$, where Λ_0 is positive definite $m \times m$ matrix. (b) There is $\hat{\Lambda}_n$ such that $\hat{\Lambda}_n \rightarrow_p \Lambda_0$.

According to Assumption 2.1(a), \hat{h}_n is a consistent and asymptotically normal estimator of h_0 . Part (b) of the assumption requires that Λ_0 can be estimated consistently from the data. The above assumptions are of high level and can be verified from more primitive conditions for mixing processes (White, 2001) or linear processes (Phillips and Solo, 1992).

Let $\theta \in \Theta \subset R^k$ be a vector of deep parameters corresponding to a structural model specified by the calibrationist. We assume that one can compute analytically the value of the reduced-form parameters h given the model and a value of θ . The mapping from the space of θ to the space of reduced-form parameters is given by the

²The supplement is available as Hnatkovska et al. (2011) from the UBC Working Papers series. In the supplement, we also consider several extensions of our basic test, including the case when a model is estimated using one set of reduced-form parameters and evaluated with respect to another.

function $f : \Theta \rightarrow R^m$, which we call the binding function using the terminology of indirect inference (Gouriéroux et al., 1993).³ In the remainder of the paper, structural models are referred to by their binding functions.

The estimator of θ is denoted by $\hat{\theta}_n$ and chosen to minimize the distance between the sample reduced-form characteristics of the data, \hat{h}_n , and those implied by the chosen structural model f . Let A_n be a possibly random $m \times m$ weight matrix. The weight matrix can be nonrandom and is chosen by the calibrationist to put more weight on more important reduced-form parameters. Alternatively, it can be data dependent and, therefore, random.

Assumption 2.2. $A_n' A_n \rightarrow_p A' A$, where $A' A$ is of full rank.

The calibrated θ , or the CMD estimator of θ , is given by the value that minimizes the weighted distance function:

$$\hat{\theta}_n(A_n) = \arg \min_{\theta \in \Theta} \left\| A_n \left(\hat{h}_n - f(\theta) \right) \right\|^2. \quad (2.1)$$

The following definition is similar to Definitions 1 and 2 of Hall and Inoue (2003), HI hereafter.

Definition 2.1. A structural model f is misspecified if $\inf_{\theta \in \Theta} \|h_0 - f(\theta)\| > 0$.

Naturally, a structural model chosen by the calibrationist is correctly specified in the sense of Definition 2.1 in the unlikely case that f is the true data generating process. Also, model f can be correctly specified according to Definition 2.1 in the case of exact identification, i.e. when $m = k$, even if the structural model and its binding function describe an incorrect DGP.⁴ Typically, the number of reduced-form parameters available for calibration exceeds k , and therefore, the calibrationist can always choose a binding function and reduced-form parameters so that a model is overidentified and thus misspecified.

We assume that the calibrationist considers two competing structural models. The second structural model is given by the binding function g and the vector of deep parameters $\gamma \in \Gamma \subset R^l$. Let $\hat{\gamma}_n$ be the calibrated value of γ , where $\hat{\gamma}_n$ is constructed similarly to $\hat{\theta}_n$ in (2.1): $\hat{\gamma}_n(A_n) = \arg \min_{\gamma \in \Gamma} \|A_n(\hat{h}_n - g(\gamma))\|^2$.

³If the binding function has to be computed by simulations, as it often happens in practice, one can assume that the number of simulations, say τ_n increases with the sample size n but at a faster rate so that $n/\tau_n \rightarrow 0$. Under this assumption, all the results in the paper continue to hold.

⁴See the discussion of overidentification and misspecification in HI.

Assumption 2.3. *Models f and g are misspecified according to Definition 2.1.*

Next, we define pseudo-true values of the structural parameters θ and γ . A pseudo-true value minimizes the distance between h_0 and the binding functions for a given weight matrix A :

Assumption 2.4. *There exists a unique $\theta_0(A) \in \Theta$ such that for all $\theta \in \Theta$, $\|A(h_0 - f(\theta_0(A)))\| \leq \|A(h_0 - f(\theta))\|$. There exists a unique $\gamma_0(A) \in \Gamma$ such that for all $\gamma \in \Gamma$, $\|A(h_0 - g(\gamma_0(A)))\| \leq \|A(h_0 - g(\gamma))\|$.*

In the above assumption, the pseudo-true values are written as functions of A to emphasize that different choices of the weight matrix may lead to different minimizers of $\|A(h_0 - f(\theta))\|$.⁵ For notational brevity, we may suppress the dependence of pseudo-true values on A if there is no ambiguity regarding the choice of A . The uniqueness of pseudo-true values is a standard assumption in the literature.⁶ When the pseudo-true value lies in the interior of Θ , it uniquely solves the following equation, provided that f is differentiable:

$$\frac{\partial f(\theta_0(A))'}{\partial \theta} A' A (h_0 - f(\theta_0(A))) = 0. \quad (2.2)$$

Due to Assumption 2.2 on A , for θ_0 to be unique it is sufficient that $\partial f(\theta_0(A)) / \partial \theta'$ has rank k .

Given two misspecified models, the calibrationist's objective is to choose the one that provides a better A -weighted fit to the reduced-form parameter h_0 . We suggest a testing procedure for the null hypothesis that the two models have equal fits:

$$H_0 : \|A(h_0 - f(\theta_0(A)))\| = \|A(h_0 - g(\gamma_0(A)))\|. \quad (2.3)$$

The null is tested against alternatives in which one of the models provides a better fit. The calibrationist prefers model f if the following alternative is true:

$$H_f : \|A(h_0 - f(\theta_0(A)))\| < \|A(h_0 - g(\gamma_0(A)))\|. \quad (2.4)$$

Similarly, the calibrationist prefers model g when $H_g : \|A(h_0 - f(\theta_0(A)))\| > \|A(h_0 -$

⁵This point has been discussed in Maasoumi and Phillips (1982) and HI.

⁶See Assumption 3 of RV and Assumption 3 of HI.

$g(\gamma_0(A))\|$ is true. Note that different weighting schemes may correspond to a different ranking of f and g .

In order to test the null hypothesis in (2.3), it is natural to consider a sample counterpart of the difference in fits between the two competing models:

$$QLR_n \left(\hat{\theta}_n(A_n), \hat{\gamma}_n(A_n) \right) = \left\| A_n \left(\hat{h}_n - g(\hat{\gamma}_n) \right) \right\|^2 - \left\| A_n \left(\hat{h}_n - f(\hat{\theta}_n) \right) \right\|^2. \quad (2.5)$$

We make the following assumption.

Assumption 2.5. *(a) Θ and Γ are compact. (b) $\theta_0(A)$ and $\gamma_0(A)$ lie in the interior of Θ and Γ respectively. (c) f and g are continuous on Θ and Γ respectively. (d) The binding functions f and g are twice continuously differentiable in the neighborhoods of $\theta_0(A)$ and $\gamma_0(A)$ respectively. (e) F_0 and G_0 are non-singular, where*

$$F_0 = \frac{\partial f(\theta_0(A))'}{\partial \theta} A' A \frac{\partial f(\theta_0(A))}{\partial \theta'} - M_{f,0}, \quad (2.6)$$

$$M_{f,0} = (I_k \otimes (h_0 - f(\theta_0(A)))' A' A) \frac{\partial}{\partial \theta'} \text{vec} \left(\frac{\partial f(\theta_0(A))}{\partial \theta'} \right),$$

$$G_0 = \frac{\partial g(\gamma_0(A))'}{\partial \gamma} A' A \frac{\partial g(\gamma_0(A))}{\partial \gamma'} - M_{g,0}, \quad (2.7)$$

$$M_{g,0} = (I_l \otimes (h_0 - g(\gamma_0(A)))' A' A) \frac{\partial}{\partial \gamma'} \text{vec} \left(\frac{\partial g(\gamma_0(A))}{\partial \gamma'} \right).$$

One can show that under Assumptions 2.1, 2.2, 2.4, and 2.5(a)-(c), $\hat{\theta}_n$ and $\hat{\gamma}_n$ consistently estimate the corresponding pseudo-true values. As a result, QLR_n consistently estimates $\|A(h_0 - g(\gamma_0(A)))\| - \|A(h_0 - f(\theta_0(A)))\|$. Assumptions 2.5(c)-(e) are used to derive the asymptotic null distribution of QLR_n .⁷

The CMD estimator of θ solves $(\partial f(\hat{\theta}_n)' / \partial \theta) A_n' A_n (\hat{h}_n - f(\hat{\theta}_n)) = 0$. Using the mean value theorem twice to expand $f(\hat{\theta}_n)$ and $\partial f(\hat{\theta}_n) / \partial \theta'$, and taking into account the population first-order condition (2.2), we obtain the following expansion for the CMD estimators in the misspecified case:^{8,9}

⁷ Assumption 2.5(e) is similar to Assumption 5 of HI.

⁸ We also assume that $n^{1/2}(A_n' A_n - A' A) = O_p(1)$.

⁹ The result in (2.8) is similar to that of equation (9) in HI. However, in the case of CMD it involves one term less than in the GMM case. This is due to the fact that, in the case of CMD, the data enters the criterion function through h , and the parameters through the binding function. The result in (2.8) is also similar to Assumptions 10 and 13 of RV.

$$\begin{aligned}
n^{1/2} \left(\hat{\theta}_n - \theta_0 \right) &= F_0^{-1} \frac{\partial f(\theta_0)'}{\partial \theta} A' A n^{1/2} \left(\hat{h}_n - h_0 \right) \\
&\quad + F_0^{-1} \frac{\partial f(\theta_0)'}{\partial \theta} n^{1/2} (A'_n A_n - A' A) (h_0 - f(\theta_0)) + o_p(1). \quad (2.8)
\end{aligned}$$

The expansion (2.8) shows that in the misspecified case, the asymptotic distribution of the CMD estimators and QLR_n depends on that of the reduced-form parameters and the weight matrices. In this paper, we will focus on the case $A_n = A$ (fixed, known weight matrices). In the supplement to this paper, we also consider the case where $n^{1/2} \text{vec}(A'_n A_n - A' A)$ is asymptotically normal. Such a situation arises when $A'_n A_n$ is a function of sample moments as in the case of Hansen-Jagannathan distance (Kan and Robotti, 2009).

3 Model comparison

The distribution of the QLR_n statistic in (2.5) depends on the relationship between the two models. Following Vuong (1989), we consider (i) nested, (ii) strictly non-nested, and (iii) overlapping models f and g . Define

$$\mathcal{F} = \{h \in R^m : h = f(\theta), \theta \in \Theta\}, \quad \mathcal{G} = \{h \in R^m : h = g(\gamma), \gamma \in \Gamma\}.$$

The relationship between the two structural models can be defined in terms of \mathcal{F} and \mathcal{G} which represent the spaces spanned by f and g respectively for the reduced-form parameter h .

Definition 3.1. *The two structural models f and g are said to be (a) nested if $\mathcal{F} \subset \mathcal{G}$ or $\mathcal{G} \subset \mathcal{F}$, (b) strictly non-nested if $\mathcal{F} \cap \mathcal{G} = \emptyset$, and (c) overlapping if $\mathcal{F} \cap \mathcal{G} \neq \emptyset$, $\mathcal{F} \not\subset \mathcal{G}$, and $\mathcal{G} \not\subset \mathcal{F}$.*

Remarks: (i) Note that the nested case is not necessarily a result of some restrictions on the structural parameters. Models f and g can be very different in terms of their underlying structures so that θ and γ are not directly comparable, and still be nested with respect to the spaces they span for h . (ii) In the strictly non-nested case, for no combination of structural parameters' values can model f give the same reduced-form value h as model g . Thus, in the calibration context, the non-nested case appears to be much more restrictive than the nested and overlapping cases.

3.1 Nested models

Suppose that $\mathcal{G} \subset \mathcal{F}$. In this case, model g cannot provide a better fit than model f . Therefore, the calibrationist is interested in testing H_0 against H_f . Since models are nested, and under Assumption 2.4 of unique pseudo-true values, the null hypothesis can be equivalently stated as $f(\theta_0) = g(\gamma_0)$. Theorem 3.1 then describes the asymptotic properties of QLR_n in the nested case.

Lemma 3.1. *Suppose that Assumption 2.4 holds, and models f and g are nested according to Definition 3.1. Then, under H_0 in (2.3), $f(\theta_0) = g(\gamma_0)$.*

Theorem 3.1. *Suppose that $A_n = A$ for all $n \geq 1$, A is of full rank, Assumptions 2.1, 2.3-2.5 hold, and $\mathcal{G} \subset \mathcal{F}$. (a) Under H_0 , $nQLR_n(\hat{\theta}_n, \hat{\gamma}_n)$ converges in distribution to $Z' \Lambda_0^{1/2} A' A (W_{g,0} - W_{f,0}) A' A \Lambda_0^{1/2} Z$, where $Z \sim N(0, I_m)$, $W_{f,0} = W_{f,0}(1) - W_{f,0}(2) - W_{f,0}(3)$, and*

$$\begin{aligned} W_{f,0}(1) &= \frac{\partial f(\theta_0)}{\partial \theta'} F_0'^{-1} \frac{\partial f(\theta_0)'}{\partial \theta} A' A \frac{\partial f(\theta_0)}{\partial \theta'} F_0^{-1} \frac{\partial f(\theta_0)'}{\partial \theta}, \\ W_{f,0}(2) &= \frac{\partial f(\theta_0)}{\partial \theta'} \left(F_0'^{-1} + F_0^{-1} \right) \frac{\partial f(\theta_0)'}{\partial \theta}, \\ W_{f,0}(3) &= \frac{\partial f(\theta_0)}{\partial \theta'} F_0'^{-1} (M_{f,0}' + M_{f,0}) F_0^{-1} \frac{\partial f(\theta_0)'}{\partial \theta}. \end{aligned}$$

The matrix $W_{g,0}$ is defined similarly to $W_{f,0}$ with θ_0 , $\partial f/\partial \theta$, F_0 , and $M_{f,0}$ replaced by γ_0 , $\partial g/\partial \gamma$, G_0 , and $M_{g,0}$ respectively. (b) Under H_f , $P(nQLR_n(\hat{\theta}_n, \hat{\gamma}_n) > c) \rightarrow 1$ for all constants $c > 0$.

According to Theorem 3.1, $nQLR_n$ has a mixed χ^2 asymptotic distribution under the null. This result is similar to the one established by Vuong (1989) for the MLE in the case of nested models. The asymptotic distribution in part (a) is nonstandard and depends on the unknown parameters h_0 , θ_0 , γ_0 , and Λ_0 . However, this distribution and its critical values can be approximated by simulations using consistent estimators of the unknown parameters. First, let $\hat{W}_{f,n}$ and $\hat{W}_{g,n}$ be the consistent plug-in estimators of $W_{f,0}$ and $W_{g,0}$ defined in part (a) of Theorem 3.1. To construct $\hat{W}_{f,n}$ and $\hat{W}_{g,n}$, one replaces h_0 , θ_0 , γ_0 , and Λ_0 with \hat{h}_n , $\hat{\theta}_n$, $\hat{\gamma}_n$, and $\hat{\Lambda}_n$ respectively. Next, simulate $Z_r \sim N(0, I_m)$ and calculate $QLR_{nr} = Z_r' \hat{\Lambda}_n^{1/2} A' A (\hat{W}_{g,n} - \hat{W}_{f,n}) A' A \hat{\Lambda}_n^{1/2} Z_r$. The asymptotic distribution of QLR_{nr} (as $n \rightarrow \infty$,) is as described in part (a) of Theorem 3.1. Repeating this for $r = 1, \dots, R$ with Z_r being drawn independently across r 's, the

simulated critical value $c_{1-\alpha,n,R}$ is the $1-\alpha$ sample quantile of $\{QLR_{nr} : r = 1, \dots, R\}$. Thus, in the nested case, one rejects the null when $nQLR_n(\hat{\theta}_n, \hat{\gamma}_n) > c_{1-\alpha,n,R}$.

3.2 Strictly non-nested models

In the case of strictly non-nested models, the space of reduced-form parameters spanned by f and that spanned by g have no common points. In this case, the asymptotic null distribution of $n^{1/2}QLR_n$ is normal.¹⁰ The asymptotic normality is due to the presence of $(f(\theta_0) - g(\gamma_0))'A'A(\hat{h}_n - h_0) = O_p(n^{-1/2})$ in the asymptotic expansion for QLR_n . When the models are nested, this term is absent because $f(\theta_0) = g(\gamma_0)$ under H_0 in the nested case.

Theorem 3.2. *Suppose that $A_n = A$ for all $n \geq 1$, and A has full rank. Suppose that Assumptions 2.1, 2.3-2.5 hold, and $\mathcal{F} \cap \mathcal{G} = \emptyset$. (a) Under H_0 , $n^{1/2}QLR_n(\hat{\theta}_n, \hat{\gamma}_n) \rightarrow_d N(0, \omega_0^2)$ with $\omega_0 = 2\|\Lambda_0^{1/2}A'A(f(\theta_0) - g(\gamma_0))\|$. (b) Under H_f , $P(n^{1/2}QLR_n(\hat{\theta}_n, \hat{\gamma}_n) > c) \rightarrow 1$, and $P(n^{1/2}QLR_n(\hat{\theta}_n, \hat{\gamma}_n) < -c) \rightarrow 1$ under H_g for all constants $c > 0$.*

Note that the asymptotic variance ω_0^2 is strictly positive since $f(\theta_0) \neq g(\gamma_0)$ when the models are non-nested.¹¹ Let $\hat{\omega}_n$ be the plug-in estimator of ω_0 . One rejects the null in favor of H_f when $n^{1/2}QLR_n(\hat{\theta}_n, \hat{\gamma}_n)/\hat{\omega}_n > z_{1-\alpha/2}$, where z_α is the α quantile of the standard normal distribution. If $n^{1/2}QLR_n(\hat{\theta}_n, \hat{\gamma}_n)/\hat{\omega}_n < -z_{1-\alpha/2}$, the null should be rejected in favor of H_g .

3.3 Overlapping models

When f and g are overlapping, one can adopt a sequential testing procedure proposed in Vuong (1989). In the first step, one tests whether $f(\theta_0) = g(\gamma_0)$. This hypothesis is rejected when $nQLR_n$ exceeds a critical value from the mixed χ^2 distribution, say $c_{1-\alpha_1}$, where α_1 denotes the significance level used in this step. If the hypothesis $f(\theta_0) = g(\gamma_0)$ is not rejected, one concludes that the two models have the same fit and the procedure stops. If the restriction $f(\theta_0) = g(\gamma_0)$ is rejected in the first step, one continues to the second step.

Asymptotically, the restriction $f(\theta_0) = g(\gamma_0)$ can be rejected either because $f(\theta_0) \neq g(\gamma_0)$ while the models have the same population lack of fit (H_0 is true),

¹⁰This is consistent with Vuong (1989) and RV.

¹¹This agrees with the conclusions in Section 6 of RV.

or because one of the models fits the data better (H_f or H_g is true). In the second step, one rejects H_0 of equivalence of the two models when $n^{1/2}QLR_n/\hat{\omega}_n > z_{1-\alpha_2/2}$, in which case f is the preferred model, or $n^{1/2}QLR_n/\hat{\omega}_n < -z_{1-\alpha_2/2}$, in which case g is preferred. Here α_2 denotes the significance level in the second step. If H_0 is not rejected in the second step, one concludes that the two models are equivalent. Vuong (1989) shows that the asymptotic significance level of the sequential procedure is $\max(\alpha_1, \alpha_2)$.

4 Averaged and sup-norm tests for model comparison

We argued earlier that the choice of the weight matrix A plays a crucial role when models are misspecified. In this section, we propose testing procedures that take into account models' relative performance for various weighting schemes.

Let \mathbb{A} be a set of $m \times m$ weight matrices, $\mathcal{B}(\mathbb{A})$ be a σ -field generated by open subsets of \mathbb{A} , and π be a probability measure on $\mathcal{B}(\mathbb{A})$. Let $\lambda_{(s)}(A'A)$ denote the s -th largest eigenvalue of $A'A$. We make the following assumption.¹²

Assumption 4.1. (a) \mathbb{A} is compact. (b) For $s = \max\{k, l\} + 1$ and some $\delta > 0$, $\inf_{A \in \mathbb{A}} \lambda_{(s)}(A'A) > \delta$. (c) Assumptions 2.4 and 2.5 hold for all $A \in \mathbb{A}$.

Remark. Part (b) of the above assumption states that uniformly over \mathbb{A} , the weight matrices must be at least of rank $\max\{k, l\} + 1$, so that f and g are both overidentified for any choice of $A \in \mathbb{A}$. To satisfy this condition, one can choose \mathbb{A} such that all its elements have full rank m , however, this is not necessary as we illustrate next. Suppose that h consists of two groups of reduced-form parameters, $h = (h'_1, h'_2)'$, where h_1 is an m_1 -vector, h_2 is an m_2 -vector, and $m_1 + m_2 = m$. Suppose further that f provides a better fit to h_1 , and g provides a better fit to h_2 . In such a situation, the econometrician might be interested in varying the relative weights of h_1 and h_2 in the overall fit of a model. For this purpose, \mathbb{A} can be defined as follows. Let $[\underline{c}, \bar{c}] \subset [0, 1]$, A_1 be a full rank $m_1 \times m_1$ matrix, and A_2 be a full rank $m_2 \times m_2$ matrix

¹²We thank the referee for his comments and suggestions for Assumption 4.1 and the primitive conditions below.

such that $\|A_1\| = \|A_2\|$. Now, consider

$$\begin{aligned} \mathbb{A} &= \{A(c) : c \in [\underline{c}, \bar{c}]\}, \text{ where} \\ A(c) &= \begin{pmatrix} \sqrt{c}A_1 & 0 \\ 0 & \sqrt{1-c}A_2 \end{pmatrix}. \end{aligned}$$

In this example, c determines the relative weight given to the first group of reduced-form parameters, $cA_1' A_1$, while the weight given to the second group is $(1-c)A_2' A_2$. When $m_1 > \max\{k, l\}$ and $m_2 > \max\{k, l\}$, the two models can be estimated and compared using h_1 or h_2 alone. In this case, one can use $[\underline{c}, \bar{c}] = [0, 1]$, and while for $c = 0, 1$ the matrix $A(c)$ does not have a full rank, the model comparison problem is well defined.

The null hypothesis of the averaged procedure is stated as follows:

$$H_0^a : \int_{\mathbb{A}} (\|A(h_0 - g(\gamma_0(A)))\|^2 - \|A(h_0 - f(\theta_0(A)))\|^2) \pi(dA) = 0.$$

According to H_0^a , the two models f and g provide equivalent approximations to h_0 on average, where the average is taken in the class \mathbb{A} with respect to the probability measure π . For example, \mathbb{A} may consist of a finite number of matrices A with π assigning equal weights to all A 's. Note that the pseudo-true values $\theta_0(A)$ and $\gamma_0(A)$ continue to depend on A . The null hypothesis H_0^a will be tested against the alternatives

$$\begin{aligned} H_f^a &: \int_{\mathbb{A}} (\|A(h_0 - g(\gamma_0(A)))\|^2 - \|A(h_0 - f(\theta_0(A)))\|^2) \pi(dA) > 0, \text{ or} \\ H_g^a &: \int_{\mathbb{A}} (\|A(h_0 - g(\gamma_0(A)))\|^2 - \|A(h_0 - f(\theta_0(A)))\|^2) \pi(dA) < 0. \end{aligned}$$

The null hypothesis of the sup-norm procedure is given by

$$H_0^s : \sup_{A \in \mathbb{A}} (\|A(h_0 - g(\gamma_0(A)))\|^2 - \|A(h_0 - f(\theta_0(A)))\|^2) \leq 0.$$

According to H_0^s , model f cannot outperform model g for any considered weight matrix $A \in \mathbb{A}$. Thus, H_0^s imposes a much stronger restriction than H_0^a . The null H_0^s

will be tested against the following alternative:

$$H_f^s : \sup_{A \in \mathbb{A}} (\|A(h_0 - g(\gamma_0(A)))\|^2 - \|A(h_0 - f(\theta_0(A)))\|^2) > 0.$$

According to H_f^s , there is a weight matrix A such that model f outperforms model g .

Again, we consider the QLR statistic defined in (2.5), however, it is now explicitly indexed by A :

$$QLR_n(\hat{\theta}_n(A), \hat{\gamma}_n(A), A) = \left\| A(\hat{h}_n - g(\hat{\gamma}_n(A))) \right\|^2 - \left\| A(\hat{h}_n - f(\hat{\theta}_n(A))) \right\|^2.$$

The averaged and sup-norm statistics are given by:

$$\begin{aligned} AQLR_n &= \int_{\mathbb{A}} QLR_n(\hat{\theta}_n(A), \hat{\gamma}_n(A), A) \pi(dA), \\ SQLR_n &= \sup_{A \in \mathbb{A}} QLR_n(\hat{\theta}_n(A), \hat{\gamma}_n(A), A). \end{aligned}$$

The asymptotic null distributions and ranking of models according to $AQLR_n$ or $SQLR_n$ depend on the choice of the measure π . The asymptotic null distributions of the averaged and sup-norm statistics also depend on whether f and g are nested or non-nested. When the models are nested, model g cannot outperform model f , and the inequality in H_0^s holds as an equality. We have the following result.

Theorem 4.1. *Suppose that Assumptions 2.1, 2.3, 4.1 hold, and $\mathcal{G} \subset \mathcal{F}$. Let $Z \sim N(0, I_m)$. For a given $A \in \mathbb{A}$, define $W_{f,0}(A)$ and $W_{g,0}(A)$ as $W_{f,0}$ and $W_{g,0}$ respectively in Theorem 3.1. (a) Under H_0^a , $nAQLR_n \rightarrow_d Z' \Lambda_0^{1/2} (\int A'A(W_{g,0}(A) - W_{f,0}(A))A'A\pi(dA)) \Lambda_0^{1/2} Z$. (b) Under H_0^s , $nSQLR_n \rightarrow_d \sup_{A \in \mathbb{A}} (Z' \Lambda_0^{1/2} A'A(W_{g,0}(A) - W_{f,0}(A))A'A \Lambda_0^{1/2} Z)$.*

According to Theorem 4.1, when the models are nested, the asymptotic distribution of the averaged statistic is mixed χ^2 . However, the weights are now given by the average of matrices $W_{f,0}$ and $W_{g,0}$. Note that $W_{f,0}$, $W_{g,0}$, F_0 , $M_{f,0}$ depend on A . Since $W_{f,0}(A)$ and $W_{g,0}(A)$ can be estimated consistently by the plug-in method, the critical values of the mixed χ^2 distribution can be computed by simulations as described in Section 3.1. The asymptotic null distribution of the sup-norm statistic depends on the sup-norm transformation of the mixed χ^2 distribution. Its critical values can be obtained similarly by simulations.

In the non-nested case, the asymptotic null distribution is a functional of a Gaussian process. Note that when the models are non-nested, H_0^s does not determine the null distribution uniquely. It is a composite hypothesis, and the null distribution depends on whether the restriction is binding or not, while the least favorable alternative, as usual, corresponds to the case where the restriction in H_0 is binding.

Theorem 4.2. *Suppose that Assumptions 2.1, 2.3, 4.1 hold, and $\mathcal{F} \cap \mathcal{G} = \emptyset$. Let $\{X(A) \in R : A \in \mathbb{A}\}$ be a mean zero Gaussian process such that the covariance of $X(A_1)$ and $X(A_2)$, $A_1, A_2 \in \mathbb{A}$, is $\omega_0(A_1, A_2)$, where $\omega_0(A_1, A_2) = 4(f(\theta_0(A_1)) - g(\gamma_0(A_1)))' A_1' A_1 \Lambda_0 A_2' A_2 (f(\theta_0(A_2)) - g(\gamma_0(A_2)))$. **(a)** Under H_0^a , $n^{1/2} AQLR_n \rightarrow_d N(0, \int_{\mathbb{A}} \int_{\mathbb{A}} \omega_0(A_1, A_2) \pi(dA_1) \pi(dA_2))$. **(b)** Under H_0^s , $\lim_{n \rightarrow \infty} P(n^{1/2} SQLR_n > c) \leq P(\sup_{A \in \mathbb{A}} X(A) > c)$.*

According to Theorem 4.2, the averaged statistic has a normal distribution. The variance is given by the weighted average of variances and covariances of the QLR statistics for different A 's. For the sup-norm statistic, the asymptotic distribution is that of the sup-norm of the Gaussian process, and the critical values for a test based on $SQLR_n$ can be obtained by simulations. In the case of overlapping models, one can follow a sequential procedure similar to the one discussed in Section 3.3.¹³

5 Application

In this section we apply our method to two monetary macroeconomic models, the cash-in-advance (CIA) model and the portfolio adjustment cost (PAC) model.¹⁴ We compare the performance of the two models based on their ability to match the impulse responses of output growth and inflation to a monetary growth shock, which comprise h in our application. We obtain a consistent estimate of h_0 from an SVAR model for GDP growth and inflation. The identification scheme employed for the SVAR follows Blanchard and Quah (1989) with the identifying restriction that money is neutral in the long run. This restriction is satisfied by both CIA and PAC.

¹³In Section S.3 of the supplement, we extend the results of Theorem 4.1 and 4.2 to allow for estimated weight matrices.

¹⁴Detailed discussions of these models can be found in Christiano (1991) and Christiano and Eichenbaum (1992).

5.1 CIA Model

The model economy is populated by a representative household, a firm, and a financial intermediary. At the beginning of period t the household owns the economy's entire money stock, M_t , and decides how to allocate it between purchases of consumption goods and deposits in the financial intermediary, $M_t - Q_t$, where Q_t is money allocated to purchases of consumption goods. Consumption purchases must be financed with Q_t and wage earnings. Thus, the objective of the household is to choose real consumption, C_t , working hours, H_t , and nominal deposit to solve the following problem:

$$\max_{\{C_t, H_t, M_{t+1}, Q_t\}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t [(1 - \phi) \ln C_t + \phi \ln (1 - H_t)] \right],$$

subject to $P_t C_t \leq Q_t + W_t H_t$, $Q_t \leq M_t$, and $M_{t+1} = (Q_t + W_t H_t - P_t C_t) + R_{H,t}(M_t - Q_t) + F_t + B_t$. Here \mathbb{E}_0 denotes expectation at date 0, β is the subjective discount factor, and ϕ is the share of leisure in per period utility. P_t denotes economy's price level, while W_t and $R_{H,t}$ denote nominal wage rate and return on deposit. The household also receives nominal profits paid by the firm, F_t , and the financial intermediary, B_t .

The production technology in the economy is $Y_t = K_t^\alpha (\mathcal{A}_t N_t)^{1-\alpha}$, where K_t , N_t , and \mathcal{A}_t are capital stock, labor input, and labor-augmenting technology, respectively. Firms must pay the total wage bill up-front to workers, so they borrow $W_t N_t$ from the financial intermediary. Loans must be repaid at the end of period t . The representative firm's problem is

$$\max_{F_t, K_{t+1}, N_t, L_t} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^{t+1} \frac{F_t}{C_{t+1} P_{t+1}} \right],$$

subject to $F_t \leq L_t + P_t [Y_t - K_{t+1} + (1 - \delta)K_t] - W_t N_t - L_t R_{F,t}$, and $W_t N_t \leq L_t$.

The objective of the financial intermediary in this economy is simple. At the beginning of each period, it loans out the household's deposit $M_t - Q_t$ and the money injection X_t received from the central bank to the firm. At the end of period, it collects the loan plus interest $L_t R_{F,t}$ and pays the amount to the household. The household, firm and financial intermediary all take prices as given.

Technology \mathcal{A}_t and money growth rate $m_t = M_{t+1}/M_t$ follow stochastic processes $\ln \mathcal{A}_t = \psi + \ln \mathcal{A}_{t-1} + \epsilon_{\mathcal{A},t}$, with $\epsilon_{\mathcal{A},t} \sim N(0, \sigma_{\mathcal{A}}^2)$, and $\ln m_t = (1 - \rho) \ln m_{ss} + \rho \ln m_{t-1} + \epsilon_{M,t}$, with $\epsilon_{M,t} \sim N(0, \sigma_M^2)$. Here m_{ss} is the steady state inflation rate, and ψ, ρ are parameters.

To solve the model, we first re-scale all real variables by technology level \mathcal{A}_t , prices by M_t/\mathcal{A}_t , and nominal variables by M_t . Then we log-linearize the equilibrium conditions around the deterministic steady state and solve the resulting system of linear difference equations. We use the solution to obtain the theoretical impulse responses of the model. These impulse responses are conditional on the structural model parameters $[m_{ss}, \alpha, \beta, \delta, \psi, \phi, \rho, \sigma_{\mathcal{A}}^2, \sigma_M^2]'$.

5.2 PAC Model

The production function and stochastic processes governing technology and money growth in the PAC model are the same as in the CIA model. The key difference between the two models is in the information sets that the household faces. In particular, in the PAC model the household's contingency plan for deposit holdings is not a function of period- t realizations of shocks. This rigidity of Q_t implies that any positive money shock must be absorbed by firms. For firms to be willing to do so voluntarily, the interest rate must fall. To make this liquidity effect persistent, Christiano (1991) introduces the second distinct feature of the PAC model - the adjustment cost $\tilde{p}_t = \alpha_1 [\exp(\alpha_2(Q_t/Q_{t-1} - m_{ss})) + \exp(-\alpha_2(Q_t/Q_{t-1} - m_{ss})) - 2]$. The household's problem in the PAC model is

$$\max_{\{C_t, H_t, M_{t+1}, Q_{t+1}\}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t [(1 - \phi) \ln C_t + \phi \ln(1 - H_t - \tilde{p}_t)] \right],$$

subject to the same sequence of constraints as in the CIA problem. The firm's problem and the financial intermediary's problem are identical to those in the CIA model. We solve this model and calculate its theoretical impulse responses using the same procedure as for CIA model. These impulse responses are conditional on the set of structural model parameters $[m_{ss}, \alpha, \beta, \delta, \psi, \phi, \rho, \sigma_{\mathcal{A}}^2, \sigma_M^2, \alpha_1, \alpha_2]'$.

5.3 Model Estimation and Comparison Results

As mentioned before, our vector h_0 consists of output growth and inflation impulse responses (seven periods each) to a money growth shock. We use the SVAR model with four lags for the GDP per capita growth and inflation series to obtain \hat{h}_n . The dataset used in the empirical analysis contains U.S. GDP per capita growth rate and inflation

rate obtained from the Basic Economics database produced by DRI/McGraw-Hill for the period of 1947:Q2-2009:Q2.¹⁵

In our comparison testing, we treat the two models as overlapping. This is because the PAC model, on the one hand, is more restrictive than CIA due to our assumption on the timing of households' decisions. On the other hand, PAC can generate richer dynamics than the CIA model, due to adjustment costs.

To reduce the computation time, in our calibration exercise we fix the values of some parameters. Following Christiano and Eichenbaum (1992), we set $\alpha = 0.36$, $\beta = 1.03^{-1/4}$, $\phi = 0.797$, $\delta = 0.012$, $\psi = 0.004$, and $m_{ss} = 0.012$. We borrow the value of $\sigma_{\mathcal{A}} = 0.014$ from Christiano (1991). We calibrate the rest of the parameters. Note that the presence of preset parameters does not impose an additional difficulty in our framework since the binding functions are assumed to be misspecified. Because our procedure requires that the parameter vectors are defined on compact sets, we restrict $\rho \in [0, 1]$ and $\sigma_M \in [0.0001, 0.01]$. In the log-linearized PAC model, α_1 and α_2 only enter through the combination $\alpha_1\alpha_2^2$. We draw on Christiano and Eichenbaum (1992) to restrict $\alpha_1\alpha_2^2 \in [10, 100]$. To implement our test, first we estimate $\theta = [\rho, \sigma_M^2]'$ by minimizing the distance between the theoretical impulse responses in the CIA model, $f(\theta)$, and the empirical impulse responses from the SVAR model, \hat{h}_n . Similarly, we estimate $\gamma = [\rho, \sigma_M^2, \alpha_1\alpha_2^2]'$ by minimizing the distance between the PAC-implied impulse responses $g(\gamma)$ and \hat{h}_n . Both distances are computed using the identity weight matrix: $A_n = I_{14}$.¹⁶ Table 1 summarizes our parameter estimates.

Figure 1 plots the models' implied impulse responses of inflation and output growth. The top panel shows that both CIA and PAC models are somewhat successful in replicating inflation dynamics, while the bottom panel indicates that both models lack in their ability to match the real-side dynamics.¹⁷ At the same time, the PAC model does marginally better in fitting output impulse responses than the CIA model.

¹⁵The GDP per capita series are obtained as a ratio of GDP series (*GDP215* in the DRI database) and population series (*POP* in the DRI database). For the price level, we choose the GDP deflator series (*GDPD15* in the DRI database).

¹⁶We also verify that each of the models is misspecified in the sense of Definition 2.1. For this purpose, we perform the usual overidentifying restrictions test. For model f , the test statistic is given by $\min_{\theta \in \Theta} n(\hat{h}_n - f(\theta))' \hat{\Lambda}_n^{-1} (\hat{h}_n - f(\theta))$ and has the χ_{m-k}^2 asymptotic null distribution. We find that the test statistic is equal to 20.98 for the CIA model, and to 144.7 for the PAC model. The corresponding p -values are 0.05 and 0.00, and we conclude that both models are misspecified.

¹⁷These results agree with the findings of Nason and Cogley (1994).

To determine whether this result is statistically significant, we follow the two-step procedure described in Section 3.3. In the first-step, we find $nQLR_n = 0.00098$. The first-step asymptotic null-distribution is mixed χ_{14}^2 , and the simulated 5% and 10% critical values are 0.00063 and 0.00044, respectively. The first-step null-hypothesis $f(\theta_0) = g(\gamma_0)$ is rejected with the p -value 0.014. In the second step, one rejects the null of the models' equivalence by comparing the t -statistic $n^{1/2}QLR_n/(2\|\hat{\Lambda}_n^{1/2}f(\hat{\theta}_n) - \hat{g}(\hat{\gamma}_n)\|)$ with the standard normal critical values. The resulting value for the t -statistic is 0.59, and the null of the CIA and PAC equivalence is not rejected. Our findings therefore indicate that the rigidities underlying the persistent liquidity effect in the PAC model do not play a significant role in approximating the inflation and output impulse response dynamics.¹⁸

To account for potential differences in models' performance for various weighting schemes, we implement our averaged and sup-norm procedures of Section 4. In addition to the equal weights imposed by $A = I_{14}$, we consider three additional weighting schemes: (i) with all weight assigned to inflation impulse responses; (ii) with all weight assigned to the first period (on impact) impulse responses for output growth and inflation; (iii) with weight assigned to the first and second periods of output growth and inflation impulse responses. The choice of π is such that all alternative weighting schemes are given equal weights.

Based on the averaged test, we find that the CIA and PAC models provide equivalent approximations to the true h_0 on average ($nAQLR_n = 0.00069$ with the p -value of 0.96 in the first step). Based on the sup-norm test, we find that the CIA model cannot outperform the PAC model for any of the considered weight matrices in our list ($nSQLR_n = 0.0016$ the p -value of 0.97 in the first step).

Acknowledgements

We thank two anonymous referees, the guest co-editors, Brant Abbott, Don Andrews, Paul Beaudry, Mick Devereux, Atsushi Inoue, Tae-Seok Jang, Oscar Jorda, Amartya Lahiri, James MacKinnon, William McCausland, Xiaodan Gao, Eric Renault, Shinichi Sakata, Henry Siu, and Gregor Smith for very valuable comments

¹⁸Schorfheide (2000) finds that the PAC model provides a better approximation to the posterior mean impulse response function than CIA. According to our comparison criterion however, the two models are equivalent.

and suggestions. The second author gratefully acknowledges the research support of the Social Sciences and Humanities Research Council of Canada under grant number 410-2007-1998 and Hampton Research Endowment Fund under Small HSS Research Grant.

References

- Blanchard, O. J., Quah, D., 1989. The dynamic effects of aggregate demand and supply disturbances. *American Economic Review* 79 (4), 655–73.
- Canova, F., 1994. Statistical inference in calibrated models. *Journal of Applied Econometrics* 9 (S1), S123–S144.
- Christiano, L. J., 1991. Modeling the liquidity effect of a money shock. *Federal Reserve Bank of Minneapolis Quarterly Review* 15 (1), 3–34.
- Christiano, L. J., Eichenbaum, M., 1992. Liquidity effects and the monetary transmission mechanism. *American Economic Review* 82 (2), 346–53.
- Cogley, T., Nason, J. M., 1995. Output dynamics in real-business-cycle models. *American Economic Review* 85 (3), 492–511.
- Corradi, V., Swanson, N. R., 2007. Evaluation of dynamic stochastic general equilibrium models based on distributional comparison of simulated and historic data. *Journal of Econometrics* 136 (2), 699–723.
- Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13 (3), 253–263.
- Dridi, R., Guay, A., Renault, E., 2007. Indirect inference and calibration of dynamic stochastic general equilibrium models. *Journal of Econometrics* 136 (2), 397–430.
- Fuerst, T. S., 1992. Liquidity, loanable funds, and real activity. *Journal of Monetary Economics* 29 (1), 3–24.
- Gouriéroux, C., Monfort, A., Renault, E., 1993. Indirect inference. *Journal of Applied Econometrics* 8 (S1), S85–S118.

- Gregory, A. W., Smith, G. W., 1991. Calibration as testing: Inference in simulated macroeconomic models. *Journal of Business & Economic Statistics* 9 (3), 297–303.
- Gregory, A. W., Smith, G. W., 1993. Statistical aspects of calibration in macroeconomics. In: Maddala, G. S., Rao, C. R., Vinod, H. D. (Eds.), *Handbook of Statistics*. Vol. 11. North-Holland, Amsterdam, Ch. 25, pp. 703–719.
- Hall, A. R., Inoue, A., 2003. The large sample behavior of the generalized method of moments estimator in misspecified models. *Journal of Econometrics* 114 (2), 361–394.
- Hnatkovska, V., Marmer, V., Tang, Y., 2011. Supplement to comparison of misspecified calibrated models, Working Paper, University of British Columbia.
- Kan, R., Robotti, C., 2009. Model comparison using the Hansen-Jagannathan distance. *Review of Financial Studies* 22 (9), 3449–3490.
- Kim, K., Pagan, A., 1995. The econometric analysis of calibrated macroeconomic models. In: Pesaran, M. H., Wickens, M. R. (Eds.), *Handbook of Applied Econometrics: Macroeconomics*. Blackwell, Cambridge, Massachusetts, Ch. 7, pp. 356–390.
- Lucas, R. J., 1990. Liquidity and interest rates. *Journal of Economic Theory* 50 (2), 237–264.
- Maasoumi, E., Phillips, P. C. B., 1982. On the behavior of inconsistent instrumental variables estimators. *Journal of Econometrics* 19 (2–3), 183–201.
- Nason, J. M., Cogley, T., 1994. Testing the implications of long-run neutrality for monetary business cycle models. *Journal of Applied Econometrics* 9 (S1), S37–S70.
- Phillips, P. C. B., Solo, V., 1992. Asymptotics for linear processes. *Annals of Statistics* 20 (2), 971–1001.
- Prescott, E. C., 1991. Real business cycle theory: What have we learned? *Revista de Análisis Económico* 6 (2), 3–19.
- Rivers, D., Vuong, Q., 2002. Model selection tests for nonlinear dynamic models. *Econometrics Journal* 5 (1), 1–39.

- Schorfheide, F., 2000. Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics* 15 (6), 645–670.
- Vuong, Q. H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57 (2), 307–333.
- West, K. D., 1996. Asymptotic inference about predictive ability. *Econometrica* 64 (5), 1067–1084.
- White, H., 2000. A reality check for data snooping. *Econometrica* 68 (5), 1097–1126.
- White, H., 2001. *Asymptotic Theory For Econometricians*. Academic Press, San Diego.

Table 1: CIA and PAC parameters' estimates and their standard errors

Parameter		Range	CIA estimates	PAC estimates
ρ	money shock persistence	[0,1]	0.8997 (0.00001)	0.9555 (0.0192)
σ_M	std.dev. of money growth innovations	[0.0001,0.004]	0.0022 (0.0003)	0.00229 (0.0003)
$\alpha_1\alpha_2^2$	adjustment cost parameter	[10, 100]	-	99.99 (0.9838)

Figure 1: Data and model-implied impulse responses

