Corradi, Valentina; Swanson, Norman R.

Working Paper

Evaluation of Dynamic Stochastic General Equilibrium Models Based on Distributional Comparison of Simulated and Historical Data

Working Paper, No. 2003-20

Provided in Cooperation with:
Department of Economics, Rutgers University

Suggested Citation: Corradi, Valentina; Swanson, Norman R. (2003) : Evaluation of Dynamic Stochastic General Equilibrium Models Based on Distributional Comparison of Simulated and Historical Data, Working Paper, No. 2003-20, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:
https://hdl.handle.net/10419/23176

Evaluation of Dynamic Stochastic General Equilibrium Models Based on Distributional Comparison of Simulated and Historical Data*

Valentina Corradi[1] and Norman R. Swanson[2]

[1]Queen Mary, University of London and [2]Rutgers University

February 2003
this revision: September 2003

Abstract

We take as a starting point the existence of a joint distribution implied by different dynamic stochastic general equilibrium (DSGE) models, all of which are potentially misspecified. Our objective is to compare "true" joint distributions with ones generated by given DSGEs. This is accomplished via the construction of a new tool for comparing the empirical joint distribution of historical time series with the empirical distribution of simulated time series. The tool draws on recent advances in the theory of the bootstrap, Kolmogorov type testing, and other work on the evaluation of DSGEs, aimed at comparing the second order properties of historical and simulated time series. We begin by fixing a given model as the "benchmark" model, against which all "alternative" models are to be compared. Our comparison is done using a distributional generalization of White's (2000) reality check. In particular, we test whether at least one of the alternative models provides a more "accurate" approximation to the true cumulative distribution than does the benchmark model. Accuracy is measured in terms of distributional square error. As the data are simulated using estimated parameters (as well as previously calibrated parameters), the limiting distribution of the test statistic is a Gaussian process with a covariance kernel that reflects the contribution of parameter estimation error. Thus, the limiting distribution is not nuisance parameter free, and critical values cannot be tabulated. In order to address this issue, we show the validity of two versions of the block bootstrap in our context. An illustrative example is also given, in which the testing approach is applied to a real business cycle model. It is shown that alternative versions of the model in which calibrated parameters are allowed to vary slightly perform equally well. On the other hand, there are stark differences between models when the shocks driving the models are assigned non-plausible variances and/or distributional assumptions.

JEL classification: C12, C22.
Keywords: Real business cycles, output, empirical distribution, simulated models, model selection.

---

* Valentina Corradi, Department of Economics, Queen Mary, University of London, Mile End, London E14NS, U.K., v.corradi@qmul.ac.uk. Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu. The authors wish to thank the editors of this special issue, and in particular, we wish to thank Roko Aliprantis and an anonymous referee for very useful comments on an earlier version of this paper. In addition, we are grateful to Michael Bordo, Roberto Chang, Stephanie Schmitt-Grohe, John Landon-Lane, Michael Pakko, and Frank Schorfeide for helpful comments. Corradi gratefully acknowledges financial support from the ESRC, grant code R000230006, and Swanson has benefited from the support of Rutgers University in the form of a Research Council grant.

1 Introduction

In this paper, we merge recent econometric advances in bootstrapping time series and Kolmogorov type testing with recent developments in the evaluation of dynamic stochastic general equilibrium (DSGE) models. This is accomplished via the construction of a new tool for comparing the empirical joint distribution of historical time series with the empirical distribution of simulated time series.

Since the seminal papers by Kydland and Prescott (1982), Long and Plosser (1983) and King, Plosser and Rebelo (KPR: 1988a,b), there has been substantial attention given to the problem of reconciling the dynamic properties of data simulated from real business cycle (RBC) models with the historical record. A partial list of advances in this area includes: (i) the examination of how RBC simulated data reproduce the covariance and autocorrelation functions of actual time series (see e.g. Watson (1993)); (ii) the comparison of RBC and historical spectral densities (see e.g. Diebold, Ohanian and Berkowitz (1998)); (iii) the evaluation of the difference between the second order time series properties of vector autoregression (VAR) predictions and out-of-sample predictions from RBC models (see e.g. Schmitt-Grohe (2000)); (iv) the construction of Bayesian odds ratios for comparing RBC models with unrestricted VAR models (see e.g. Chang, Gomes and Schorfheide (2002), and Fernandez-Villaverde (2002)); (v) the comparison of historical and simulated data impulse response functions (see e.g. Cogley and Nason (1995)); and (vi) the formulation of "reality" bounds for measuring how close the density of an RBC model is to the density associated with an unrestricted VAR model (see e.g. Bierens and Swanson (2000)). The papers listed above are mainly concerned with the issue of model evaluation. Another strand of the literature is instead mainly concerned with providing alternatives to calibration (see e.g. De Jong, Ingram and Whiteman (2000) for a Bayesian perspective in which prior distributions are constructed around calibrated structural parameters). In most of the above papers, the issue of singularity (i.e. when the number of variables in the model is larger than the number of shocks) is circumvented by considering only a subset of variables, for which a nondegenerate distribution exists.[1] Our work is closest to the first strand of literature. In particular,

[1] A novel alternative to calibration is proposed by Bierens (2003). He solves the singularity problem via convolution of the singular distribution of an RBC model with a non singular distribution. The same convolution is also applied to the associated non singular distribution of the econometric model, which is an

our paper attempts to add to the model evaluation literature by introducing a measure of the "goodness of fit" of RBC models that is based on applying standard notions of Kolmogorov distance and drawing on advances in the theory of the bootstrap.[2]

The papers cited above primarily address the case in which the objective is to test for the correct specification of some aspects of a given candidate model. In the case of RBC models, however, it is usually crucial to account for the fact that all models are approximations, and so are misspecified. This is the reason why the testing procedure that we develop allows us to evaluate the relative degree of misspecification of a given group of competing models based on the comparison of empirical distribution functions of historical data with those of RBC simulated data. The RBC models of interest in our context are simulated using both calibrated parameters (with calibrated values suggested by KPR, for example), and parameters estimated by using actual data, along the lines of Christiano (1988) and Christiano and Eichenbaum (1992). One important feature of our approach is that we begin by fixing a given RBC model as the "benchmark" model, against which all "alternative" models are compared. The comparison is done using a distributional generalization of White's (2000) reality check, which assesses whether at least one of the alternative models provides a more "accurate" approximation to the true cumulative distribution than does the benchmark model. One key element of our approach is that we measure "accuracy" in terms of square error, as in Corradi and Swanson (2002). We also outline the relationship between our measure of accuracy and the Kullback-Leibler Information Criterion (KLIC). DSGE model evaluation based on KLIC measures of accuracy is considered by Fernandez-Villaverde and Rubio-Ramirez (2001), and Chan, Gomes and Schorfheide (2002).

As mentioned above, our statistic is based on comparison of historical and simulated distributions. The limiting distribution of the statistic is a functional over a Gaussian process with covariance kernel that reflects the contribution of parameter estimation error. This limiting distribution is thus not nuisance parameter free, and critical values cannot be tab-

unrestricted VAR. Parameters are then estimated via the maximization of an information criterion which gives the probability that the distribution of the convoluted VAR model is generated by the distribution of the convoluted RBC model, conditional on the data.

[2]In recent years, Kolmogorov type distance measures for testing distributional equality have been extended to the case of dependent observations and/or parameter estimation error (see e.g. Andrews (1997), Bai (2003)). Tests of this sort generally have limiting distributions that are not nuisance parameter free, and critical values cannot be tabulated. Papers addressing this issue in the bootstrap literature include Andrews (2002), Goncalves and White (2002), Hall and Horowitz (1996), Horowitz (2002), Inoue and Shintani (2003), Naik-Nimbalkar and Rajarshi (1994).

ulated. In order to obtain valid asymptotic critical values, we suggest two block bootstrap procedures, each of which depends on the relative rate of growth of the actual and simulated sample size. In addition, we circumvent the issue of singularity by considering a subset of variables (and their lagged values) for which a non singular distribution exists.

Our testing framework can be used to address questions of the following sort: (i) For a given RBC model, what is the relative usefulness of different sets of calibrated parameters for mimicking different dynamic features of output growth? (ii) Given a fixed set of calibrated parameters, what is the relative performance of RBC models driven by shocks with a different marginal distribution?

In order to illustrate how the proposed testing framework, we consider the RBC model of Christiano (1988)), characterized by flexible labor supply, capital depreciation, and two shocks - a permanent shock affecting technology and a transitory shock affecting preferences. Data are then simulated and various versions of the model are compared, in terms of their ability to reproduce the joint distribution of current output, lagged output, current hours worked, and lagged hours worked. The illustrations suggest that the methodology outlined in this paper provides a useful additional tool for examining the relevance of different RBC models vis a vis how well competing models capture the dynamic structural characteristics of the historical record.

The rest of the paper is organized as follows. Section 2 outlines the testing framework, describes the test statistic, and shows that the limiting distribution of the statistic is a zero mean Gaussian process with a covariance kernel that reflects both the contribution of parameter estimation error and the time series structure of the data. This is all done under the assumption that all models may be misspecified. In Section 3, the construction of bootstrap critical values is outlined, and the first order validity of the block bootstrap is established under two different assumptions with respect to the limit of the ratio of the actual sample size and the simulated sample period. An empirical illustration is given in Section 4, and concluding remarks are gathered in Section 5. All proofs are collected in the appendix.

2 Testing Framework

Our objective is to compare the joint distribution of historical variables with the joint distribution of the simulated variables. Hereafter, for sake of simplicity, but without loss of generality, we limit our attention to the joint empirical distribution of (actual and model-based) current and previous period output. In principle, if we have a model driven by k shocks, then we can consider the joint CDF of k variables plus an arbitrary (but finite) number of lags. Consider m RBC models, and set model 1 as the benchmark model. Let $\Delta \log X_t$, $t = 1, ..., T$ denote actual historical output (growth rates) and let $\Delta \log X_{j,n}$, $j = 1, ..., m$ and $n = 1, ..., S$, denote the output series simulated under model j, where S denotes the simulated sample length. As mentioned above, some parameters are assumed to be kept fixed (at their calibrated values), while other parameters are estimated. Along these lines, denote $\Delta \log X_{j,n}(\widehat{\theta}_{j,T})$, $n = 1, ..., S$, $j = 1, ..., m$ to be a sample of length S drawn (simulated) from model j and evaluated at the parameters estimated under model j, where parameter estimation is done using the T available historical observations.[3] Now, let $Y_t = (\Delta \log X_t, \Delta \log X_{t-1})$ and $Y_{j,n}(\widehat{\theta}_{j,T}) = (\Delta \log X_{j,n}(\widehat{\theta}_{j,T}), \Delta \log X_{j,n-1}(\widehat{\theta}_{j,T}))$. Also, let $F_0(u; \theta_0)$ denote the distribution of Y_t evaluated at u and $F_j(u; \theta_j^\dagger)$ denote the distribution of $Y_{j,n}(\theta_j^\dagger)$, where $\theta_j^\dagger$ is the probability limit of $\widehat{\theta}_{j,T}$, taken as $T \to \infty$, and where $u \in U \subset \Re^2$, possibly unbounded. Accuracy is measured in terms of the squared (approximation) error associated with model j, $j = 1, ..., m$, defined as a (weighted) average over U of $\left(\left(F_j(u; \theta_j^\dagger) - F_0(u; \theta_0)\right)^2\right)$. Thus, the rule is to choose Model 1 over Model 2, say, if

$$\int_U \left(\left(F_1(u; \theta_1^\dagger) - F_0(u; \theta_0)\right)^2\right) \phi(u) du < \int_U \left(\left(F_2(u; \theta_2^\dagger) - F_0(u; \theta_0)\right)^2\right) \phi(u) du,$$

where $\int_U \phi(u) du = 1$ and $\phi(u) \geq 0$ for all $u \in U \subset \Re^2$. For any evaluation point, this measure defines a norm and is a typical goodness of fit measure.

Another measure of distributional accuracy available in the literature (see e.g. Vuong (1989)), is the KLIC, according to which we should choose Model 1 over Model 2 if $E(\log f_1(Y_t; \theta_2^\dagger) - \log f_2(Y_t; \theta_2^\dagger)) > 0$.[4] The KLIC thus chooses the model which on average gives higher proba-

[3] The reason why we use differences in the above is that the RBC model in our empirical illustration is characterized by permanent shocks to technology. In general, we require both the actual and the simulated series to be (strictly) stationary and mixing, and hence differencing may or may not be appropriate.

[4] Recently, Giacomini (2002) proposes an extension which uses a weighted (over Y_t) version of the KLIC, and Kitamura (2002) suggests a generalization for choosing among models that satisfy some conditional moment restrictions.

bility to events which have actually occurred. Also, it leads to simple Likelihood Ratio tests. Our approach is an alternative to the KLIC that should be viewed as complementary in some cases, and preferred in others. For example, if we are interested in measuring accuracy for a given confidence interval, this cannot be done in an obvious manner using the KLIC, while it can easily done using our measure. Furthermore, we often do not have an explicit form for the density implied by the various models we are comparing. Of course, model comparison could then be done using kernel density estimators, at the cost of nonparametric convergence rates.

Turning again to our main discussion, note that within our context, the hypotheses of interest are:

$$H_0 : \max_{j=2,\ldots,m} \int_U \left(\left(F_0(u;\theta_0) - F_1(u;\theta_1^\dagger) \right)^2 - \left(F_0(u) - F_j(u;\theta_j^\dagger) \right)^2 \right) \phi(u) du \leq 0$$

$$H_A : \max_{j=2,\ldots,m} \int_U \left(\left(F_0(u) - F_1(u;\theta_1^\dagger) \right)^2 - \left(F_0(u) - F_j(u;\theta_j^\dagger) \right)^2 \right) \phi(u) du > 0.$$

Thus, under H_0, no model can provide a better approximation than model 1. If interest focuses on confidence intervals or on testing the null of equal accuracy of two distribution models (analogous to the pairwise conditional mean comparison setup of Diebold and Mariano (1995)), then analogous hypotheses can be constructed by replacing terms of the type $\int_U \left(\left(F_0(u;\theta_0) - F_1(u;\theta_1^\dagger) \right)^2 \right) \phi(u) du$ with terms such a $\left(\left(F_1(\overline{u};\theta_1^\dagger) - F_1(\underline{u};\theta_1^\dagger) \right) - \left(F_0(\overline{u};\theta_0) - F_0(\underline{u};\theta_0) \right) \right)^2$ or by simply removing $\max_{j=2,\ldots,m}$ from the statement of the hypotheses, respectively. In order to test H_0 versus H_A, the relevant test statistic is $\sqrt{T} Z_{T,S}$, where:

$$Z_{T,S} = \max_{j=2,\ldots,m} \int_U Z_{j,T,S}(u) \phi(u) du, \tag{1}$$

and

$$\begin{aligned} Z_{j,T,S}(u) \;=\; & \frac{1}{T} \sum_{t=1}^{T} \left(1\{Y_t \leq u\} - \frac{1}{S} \sum_{n=1}^{S} 1\{Y_{1,n}(\widehat{\theta}_{1,T}) \leq u\} \right)^2 \\ & - \frac{1}{T} \sum_{t=1}^{T} \left(1\{Y_t \leq u\} - \frac{1}{S} \sum_{n=1}^{S} 1\{Y_{j,n}(\widehat{\theta}_{j,T}) \leq u\} \right)^2, \end{aligned}$$

with $\widehat{\theta}_{j,T}$ an estimator of $\theta_j^\dagger$ that satisfies Assumption 2 below. From equation (1), it is immediate to see that the computational burden of our procedure increases with the di-

mensionality of U (i.e. as the number of variables and their lags increases).[5] Unfortunately, Monte Carlo integration techniques, such as importance sampling or one of its accelerated versions (see e.g. Danielsson and Richard (1993)) are not applicable in our context. This is because $Z_{j,T,S}(u)\phi(u)$ is not a joint density and can be either negative or positive. An alternative possibility would be to compute the statistic $\widetilde{Z}_{T,S} = \max_{j=2,...,m} \frac{1}{T}\sum_{i=k}^{T} Z_{j,T,S}(Y_i)$, where k denotes the highest lag order. If Y_t were an *iid* vector-valued process, then $\sqrt{T}Z_{T,S}$ and $\sqrt{T}\widetilde{Z}_{T,S}$ would be asymptotically equivalent, as shown in Andrews (1997). However, in our case, Y_t is a dependent process. Thus, the argument used in Andrews' proof (based on his Lemma A6) does not apply. Nevertheless, as T gets large, $(Y_k, Y_{k+1}, ...Y_T)$ will become a dense subset in U, and so we conjecture that $\sqrt{T}Z_{T,S}$ and $\sqrt{T}\widetilde{Z}_{T,S}$ are asymptotically equivalent even in the dependent case. In the sequel, we require the following assumptions:

Assumption A1: Y_t is stationary-ergodic $\beta-$mixing processes with size -4, for $j = 1, ..., m$.[6]

Assumption A2: For $j = 1, ..., m$: $\sqrt{T}\left(\widehat{\theta}_{j,T} - \theta_j^{\dagger}\right) = A_j(\theta_j^{\dagger})\frac{1}{\sqrt{T}}\sum_{t=2}^{T} q_j(Y_t, \theta_j^{\dagger}) + o_P(1)$, where $\frac{1}{\sqrt{T}}\sum_{t=2}^{T} q_j(Y_t, \theta_j^{\dagger})$ satisfies a central limit theorem and $A_j(\theta_j^{\dagger})$ is positive definite.[7]

Assumption A3: For $j = 1, ..., m$: (i) $\forall \theta_j \in \Theta_j$, with Θ_j a compact set in $\Re^{p_j}$ and $Y_{j,n}(\theta_j)$ is a strictly stationary ergodic $\beta-$mixing process with size -4, where p_j is the number of estimated parameters in model j; (ii) $Y_{j,n}(\theta_j)$ is continuously differentiable in the interior of Θ_j, for $n = 1, ..., S$; (iii) $\nabla_{\theta_j} Y_{j,n}(\theta_j)$ is $2r-$dominated in Θ_j, uniformly in n for $r > 2$;[8] (iv) $F_j(u; \theta_j^{\dagger})$ is twice continuously differentiable in u; and (v) for at least one j, $F_j(u; \theta_j^{\dagger}) \neq F_1(u; \theta_1^{\dagger})$ for $u \in \widetilde{U}$, where $\widetilde{U}$ is a subset of U of non-zero Lebesgue measure.

A2 requires that $\sqrt{T}\left(\widehat{\theta}_{j,T} - \theta_j^{\dagger}\right)$ is asymptotically normal with a positive definite covariance matrix. Thus, given the size condition in **A1**, **A2** is satisfied by OLS, NLS, and QMLE, under mild conditions, such as finite $(4+\delta)$th moments and unique identifiability. It is satis-

[5] For example, if U is a two-dimensional subset of $\Re^2$, and ϕ is uniform on U, then

$$Z_{T,S} = \frac{1}{N_1 N_2} \max_{j=2,...,m} \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} Z_{j,T,S}(u_{i,j}).$$

[6] $\beta-$mixing is a memory requirement stronger than $\alpha-$mixing, but weaker than (uniform) $\phi-$mixing.

[7] Given the size condition in A1, A2 is satisfied by the LS, NLS, QMLE estimator, under mild conditions, such as finite $(4+\delta)$th moments and unique identifiability.

[8] This means that $|\nabla_{\theta_j} Y_{j,n}(\theta_j)| \leq D_{j,n}$, with $\sup_n E(D_{j,n}^{2r}) < \infty$ (see e.g. Gallant and White (1988), p.33).

fied for the GMM-type estimator of Christiano and Eichenbaum (1992) and the estimator of Bierens (2003). With regard to **A3**(i), whenever the production function is a Cobb-Douglas type, and the shock to technology follows a unit root process in logs, then output follows a unit root process in logs, and the growth rate is stationary. This is not necessarily true in the case of more general CES production functions. **A3**(ii) need only hold for estimated parameters. When solving RBC models, we often obtain a (linear) ARMA representations for the variables of interest, in terms of final (or reduced form) parameters. Therefore, because of linearity, **A3**(ii) holds straightforwardly for the final parameters. Hence, if the structural (deep) parameters are smooth functions of the final parameters, as is often the case **A3**(ii) is satisfied. **A3**(iii) is a standard assumption (see e.g. Duffie and Singleton (1993)), and **A3**(iv) is always satisfied for linearized solutions of RBC models. Finally, **A3**(v) ensures that at least one competing model is nonnested with the benchmark model. This in turn ensures that the covariance matrix of the statistic is positive semi-definite. Hereafter, for notational simplicity, let $F_j(u) = F_j(u;\theta_j^\dagger)$.

Proposition 1: Let Assumptions A1-A3 hold. (i) Assume that as $T, S \to \infty : T/S \to \delta$, $0 < \delta < \infty$, then:

$$\max_{j=2,\ldots,m} \sqrt{T} \int_U \left(Z_{j,T,S}(u) - \left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2 \right) \right) \phi(u) du \tag{2}$$
$$\xrightarrow{d} \max_{j=2,\ldots,m} \int_U Z_j(u)\phi(u)du,$$

where $Z_j(u)$ is a zero mean Gaussian process with covariance kernel:

$$\begin{aligned}
&K_j(u,u') = 4\left(F_1(u) - F_j(u)\right) C_{yy}(u,u') \left(F_1(u') - F_j(u')\right) + 4\mu_{F_1}(u)\delta^2 C_{y_1y_1}(u,u')\mu_{F_1}(u') \\
&+4\mu_{F_j}(u)\delta^2 C_{y_jy_j}(u,u')\mu_{F_j}(u') + 4\mu_{F_1}(u)\mu_{f_{1,\theta_1^\dagger}}(u)A(\theta_1^\dagger)C_{11}A(\theta_1^\dagger)\mu_{F_1}(u')\mu_{f_{1,\theta_1^\dagger}}(u') \\
&+4\mu_{F_j}(u)\mu_{f_{j,\theta_j^\dagger}}(u)A(\theta_j^\dagger)C_{jj}A(\theta_j^\dagger)\mu_{F_j}(u')\mu_{f_{1,\theta_j^\dagger}}(u') + 8\mu_{F_1}(u)\delta C_{y,y_1}(u,u')\left(F_1(u') - F_j(u')\right) \\
&-8\mu_{F_j}(u)\delta C_{y,y_j}(u,u')\left(F_1(u') - F_j(u')\right) + 8\mu_{F_1}(u)\mu_{f_{1,\theta_1^\dagger}}(u)A(\theta_1^\dagger)C_{y,1}(u')\left(F_1(u') - F_j(u')\right) \\
&-8\mu_{F_j}(u)\mu_{f_{j,\theta_j^\dagger}}(u)A(\theta_j^\dagger)C_{y,j}(u')\left(F_1(u') - F_j(u')\right) - 8\mu_{F_j}(u)\delta^2 C_{y_jy_1}(u,u')\mu_{F_1}(u') \\
&+8\mu_{F_1}(u)\delta C_{y_1,1}(u)A(\theta_1^\dagger)\mu_{F_1}(u')\mu_{f_{1,\theta_1^\dagger}}(u') - 8\mu_{F_1}(u)\delta C_{y_1,j}(u)A(\theta_j^\dagger)\mu_{F_j}(u')\mu_{f_{j,\theta_j^\dagger}}(u') \\
&+8\mu_{F_j}(u)\delta C_{y_j,j}(u)A(\theta_j^\dagger)\mu_{F_j}(u')\mu_{f_{j,\theta_j^\dagger}}(u') - 8\mu_{F_j}(u)\delta C_{y_j,1}(u)A(\theta_1^\dagger)\mu_{F_1}(u')\mu_{f_{1,\theta_1^\dagger}}(u') \\
&-8\mu_{F_1}(u)\mu_{f_{1,\theta_1^\dagger}}(u)A(\theta_1^\dagger)C_{1j}A(\theta_j^\dagger)\mu_{F_j}(u')\mu_{f_{1,\theta_j^\dagger}}(u')
\end{aligned}$$

with $C_{yy}(u,u') = E\left(\sum_{s=-\infty}^{\infty}(1\{Y_2 \leq u\} - F_0(u))(1\{Y_{2+s} \leq u'\} - F_0(u'))\right)$, and where F_0 denotes the "true" joint distribution of Y_t. Also,
$C_{y_j y_j}(u,u') = E\left(\sum_{s=-\infty}^{\infty}(1\{Y_{j,2} \leq u\} - F_j(u))(1\{Y_{j,2+s} \leq u'\} - F_0(u'))\right)$,
$C_{yy_j}(u,u') = E\left(\sum_{s=-\infty}^{\infty}(1\{Y_2 \leq u\} - F_0(u))(1\{Y_{j,2+s} \leq u'\} - F_j(u'))\right)$, and
$C_{jj} = E\left(\sum_{s=-\infty}^{\infty} q_j(Y_2,\theta_j^\dagger)q_j(Y_{2+s},\theta_j^\dagger)\right)$, $C_{y_j,j}(u) = E\left(\sum_{s=-\infty}^{\infty}(1\{Y_{j,2} \leq u\} - F_j(u))q_j(Y_{2+s},\theta_j^\dagger)\right)$.
Finally, $\mu_{F_j}(u) = E(1\{Y_t \leq u\} - F_j(u))$ and $\mu_{f_j,\theta_j^\dagger}(u) = E\left(f_j(u)\nabla_{\theta_j}Y_{j,n}(\theta_j^\dagger)\right)$, where f_j is the density of F_j.

(ii) Assume that as $T,S \to \infty : S/T^2 \to 0$ and $T/S \to 0$, then:

$$\max_{j=2,\ldots,m}\sqrt{T}\int_U\left(Z_{j,T,S}(u) - \left((F_0(u)-F_1(u))^2 - (F_0(u)-F_j(u))^2\right)\right)\phi(u)du \xrightarrow{d} \max_{j=2,\ldots,m}\int_U \widetilde{Z}_j(u)\phi(u)du,$$

where $\widetilde{Z}_j(u)$ is a zero mean Gaussian process with covariance kernel:

$$\begin{aligned}
&\widetilde{K}_j(u,u') = 4\left(F_1(u)-F_j(u)\right)C_{yy}(u,u')\left(F_1(u')-F_j(u')\right)\\
&+4\mu_{F_1}(u)\mu_{f_1,\theta_1^\dagger}(u)A(\theta_1^\dagger)C_{11}A(\theta_1^\dagger)\mu_{F_1}(u')\mu_{f_1,\theta_1^\dagger}(u') + 4\mu_{F_j}(u)\mu_{f_j,\theta_j^\dagger}(u)A(\theta_j^\dagger)C_{jj}A(\theta_j^\dagger)\mu_{F_j}(u')\mu_{f_1,\theta_j^\dagger}(u')\\
&+8\mu_{F_1}(u)\mu_{f_1,\theta_1^\dagger}(u)A(\theta_1^\dagger)C_{y,1}(u')\left(F_1(u')-F_j(u')\right) - 8\mu_{F_j}(u)\mu_{f_j,\theta_j^\dagger}(u)A(\theta_j^\dagger)C_{y,j}(u')\left(F_1(u')-F_j(u')\right)\\
&-8\mu_{F_1}(u)\mu_{f_1,\theta_1^\dagger}(u)A(\theta_1^\dagger)C_{1j}A(\theta_j^\dagger)\mu_{F_j}(u')\mu_{f_1,\theta_j^\dagger}(u').
\end{aligned}$$

Notice that when $T/S \to 0$, then $\frac{\sqrt{T}}{S}\sum_{n=1}^{S}\left(1\{Y_{j,n}(\theta_1^\dagger) \leq u\} - F_j(u)\right) \xrightarrow{pr.} 0$, uniformly in u, so that the covariance kernel of the limiting distribution does not reflect the contribution of the error term due to the fact we replace the "true" distribution of the simulated series with its empirical counterpart (i.e. simulation error vanishes). From Proposition 1, we see that when all competing models provide an approximation to the true joint distribution that is as accurate as the benchmark, then the limiting distribution is a zero mean Gaussian process with a covariance kernel that reflects the contribution of parameter estimation error, the time series structure of the data and, for $\delta > 0$, the contribution of simulation error. This is the case when

$$\int_U\left((F_0(u)-F_1(u))^2 - (F_0(u)-F_j(u))^2\right)\phi(u)du = 0, \text{ for all } j.$$

It follows that in this case, the limiting distribution of

$$\max_{j=2,\ldots,m}\sqrt{T}\int_U\left(Z_{j,T,S}(u) - \left((F_0(u)-F_1(u))^2 - (F_0(u)-F_j(u))^2\right)\right)\phi(u)du$$

is the same as that of $\sqrt{T}Z_{T,S}$, and so the critical values of the limiting distribution on the RHS of equation (2) provide valid asymptotic critical values for $\sqrt{T}Z_{T,S}$. On the other hand, when $\int_U \left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2 \right) \phi(u)du < 0$ some j, so that at least one alternative model is less accurate than the benchmark, then these critical values can be viewed as upper bounds for critical values from the distribution of $\sqrt{T}Z_{T,S}$. Note also that when all competing models are less accurate than the benchmark model, the statistic diverges to minus infinity. Finally, under the alternative, $\int_U \left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2 \right) \phi(u)du > 0$ for some j, so that $\sqrt{T}Z_{T,S}$ diverges to infinity. Therefore, the test has correct asymptotic size if all models are equally good, is conservative when some model is strictly dominated by the benchmark, and has unit power under the alternative. It should perhaps also be noted that our approach can in principle be modified to allow for the evaluation of predictive densities (see e.g. Corradi and Swanson (2003a)).

3 Bootstrap Critical Values

In this subsection we outline how to obtain valid critical values for the asymptotic distribution of $\max_{j=2,...,m} \sqrt{T} \int_U (Z_{j,T,S}(u) - \left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2 \right)) \phi(u)du$, via use of an empirical process version of the block bootstrap that properly captures the contribution of parameter estimation error, simulation error, and the time series dynamics to the covariance kernel given in Proposition 1. In order to show the first order validity of the bootstrap, we derive the limiting distributions of appropriately formed bootstrap statistics and show that they coincide with the limiting distribution given in Proposition 1, recalling that as all candidate models are potentially misspecified under both hypotheses. the parametric bootstrap is not generally applicable in our context. We begin by resampling b blocks of length l, $bl = T - 1$. Let $Y_t^* = (\Delta \log X_t^*, \Delta \log X_{t-1}^*)$ be the resampled series, such that $Y_2^*, ..., Y_{l+1}^*, Y_{l+2}^*, ..., Y_{T-l+2}^*, ..., Y_T^*$ equals $Y_{I_1+1}, ..., Y_{I_1+l}, Y_{I_2+1}, ..., Y_{I_b+1}, ..., Y_{I_b+l}$, where I_i, $i = 1, ..., b$ are independent, discrete uniform random variates on $1, ..., T - l + 1$. That is, $I_i = i$, $i = 1, ..., T - l$ with probability $1/(T - l)$. Then, use Y_t^* to compute $\widehat{\theta}_{j,T}^*$ and plug in $\widehat{\theta}_{j,T}^*$ in order to simulate a sample under model j, $j = 1, ..., m$. Let $Y_{j,n}(\widehat{\theta}_{j,T}^*)$, $n = 2, ..., S$ denote the series simulated in this manner. At this point, we need to distinguish between the case where $\delta = 0$ (vanishing simulation error) and $\delta > 0$ (non vanishing simulation error). In

the former case, we do not need to resample the simulated series, as there is no need to mimic the contribution of simulation error to the covariance kernel. On the other hand, in the latter case we draw $\widetilde{b}$ blocks of length $\widetilde{l}$, with $\widetilde{b}\widetilde{l} = S-1$, and let $Y^*_{j,n}(\widehat{\theta}^*_{j,T})$, $j = 1, ..., m$, $n = 2, ..., S$ denote the resampled series under model j. Notice that $Y^*_{j,2}(\widehat{\theta}^*_{j,T}), ..., Y^*_{j,l+1}(\widehat{\theta}^*_{j,T}), ..., Y^*_{j,S}(\widehat{\theta}^*_{j,T})$ is equal to $Y_{j,\widetilde{I}_1}(\widehat{\theta}^*_{j,T}), ..., Y_{j,\widetilde{I}_1+l}(\widehat{\theta}^*_{j,T}), ..., Y_{j,\widetilde{I}_{\widetilde{b}}+l}(\widehat{\theta}^*_{j,T})$, where $\widetilde{I}_i$, $i = 1, ..., \widetilde{b}$ are independent discrete uniform random variates on $1, ..., S-\widetilde{l}$. Also, notice that for *each* of the m models, and for each bootstrap replication, we draw $\widetilde{b}$ discrete uniform random variates (the $\widetilde{I}_i$) on $1, ..., S-\widetilde{l}$, and that draws are independent across models. Thus, in our use of notation, we have suppressed the dependence of $\widetilde{I}_i$ on j.

As discussed above, we consider two different bootstrap versions of $Z_{T,S}$, the first of which is valid when $T/S \to \delta > 0$ and the second of which is valid when $T/S \to 0$, so that $Y^*_{j,n}(\widehat{\theta}^*_{j,T})$ in the first statistic is replaced with $Y_{j,n}(\widehat{\theta}^*_{j,T})$, in the second version. In particular, define $Z^{**}_{T,S} = \max_{j=2,...,m} \int_U Z^{**}_{j,T,S}(u)\phi(u)du$, where $Z^{**}_{j,T,S}(u) =$

$$\frac{1}{T}\sum_{t=1}^{T}\left(\left(1\{Y_t^* \le u\} - \frac{1}{S}\sum_{n=1}^{S} 1\{Y^*_{1,n}(\widehat{\theta}^*_{1,T}) \le u\}\right)^2 - \left(1\{Y_t \le u\} - \frac{1}{S}\sum_{n=1}^{S} 1\{Y_{1,n}(\widehat{\theta}_{1,T}) \le u\}\right)^2\right)$$
$$-\frac{1}{T}\sum_{t=1}^{T}\left(\left(1\{Y_t^* \le u\} - \frac{1}{S}\sum_{n=1}^{S} 1\{Y^*_{j,n}(\widehat{\theta}^*_{j,T}) \le u\}\right)^2 - \left(1\{Y_t \le u\} - \frac{1}{S}\sum_{n=1}^{S} 1\{Y_{j,n}(\widehat{\theta}_{j,T}) \le u\}\right)^2\right).$$

Also, define $Z^*_{T,S} =: \max_{j=2,...,m} \int_U Z^*_{j,T,S}(u)\phi(u)du$, where $Z^*_{j,T,S}(u) =$

$$\frac{1}{T}\sum_{t=1}^{T}\left(\left(1\{Y_t^* \le u\} - \frac{1}{S}\sum_{n=1}^{S} 1\{Y_{1,n}(\widehat{\theta}^*_{1,T}) \le u\}\right)^2 - \left(1\{Y_t \le u\} - \frac{1}{S}\sum_{n=1}^{S} 1\{Y_{1,n}(\widehat{\theta}_{1,T}) \le u\}\right)^2\right)$$
$$-\frac{1}{T}\sum_{t=1}^{T}\left(\left(1\{Y_t^* \le u\} - \frac{1}{S}\sum_{n=1}^{S} 1\{Y_{j,n}(\widehat{\theta}^*_{j,T}) \le u\}\right)^2 - \left(1\{Y_t \le u\} - \frac{1}{S}\sum_{n=1}^{S} 1\{Y_{j,n}(\widehat{\theta}_{j,T}) \le u\}\right)^2\right).$$

Now, let Ω_0 be the probability space underlying the historical variables, Ω_j the probability space underlying the series simulated under model j (i.e. the space from which we make random draws in order to form $Y_{j,n}(\widehat{\theta}_{j,T})$, where the draws are independent of $\widehat{\theta}_{j,T}$, by construction). Then, define the enlarged probability space, Ω, as $\Omega = \Omega_0 \times \Omega_1 \times ... \times \Omega_m$. Also, let P_0 be the probability law governing the actual sample, and P_j be the probability law of the samples simulated under model j. That is, P_j is the probability law governing $Y_{j,n}(\theta_j^\dagger)$. Now, define $P = P_0 \times P_1 \times ... \times P_m$.[9] Analogously, define P_0^* to be the probability law governing the resampled series, Y_t^*, conditional on the historical sample, Y_t, and define P_j^* to be the probability law governing the resampled series, $Y^*_{j,n}(\theta_j^\dagger)$, conditional on the simulated series, $Y_{j,n}(\theta_j^\dagger)$. Now, define $P^* = P_0^* \times P_1^* \times ... \times P_m^*$. Finally, let ω denote a draw from Ω, and

[9]Note that although we generate the simulated series using $\widehat{\theta}_{j,T}$, P_j denotes the probability law of the series simulated under $\theta_j^\dagger$, which is the probability limit of $\widehat{\theta}_{j,T}$.

let l_T and l_S denote the block length of the resampled series from the actual and simulated samples, respectively. Note that when constructing $Z^*_{j,T,S}$, $\delta = 0$ so that $P^* = P^*_0$.

Proposition 2: Let Assumptions A1-A3 hold. Also, assume that as $l_T, l_S \to \infty : l_T/T^{1/2} \to 0$ and $l_S/S^{1/2} \to 0$. (i) Assume that as $T, S \to \infty : T/S \to \delta$, $0 < \delta < \infty$, then:

$$P\left(\omega : \sup_{v \in \Re} \left| P^*\left(\max_{j=2,...,m} \sqrt{T} \int_U Z^{**}_{j,T,S}(u)\phi(u)du \leq v\right)\right.\right.$$

$$\left.\left. - P\left(\max_{j=2,...,m} \sqrt{T} \int_U \left(Z_{j,T,S}(u) - \left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2\right)\right)\phi(u)du \leq v\right)\right| > \varepsilon\right) \to 0.$$

(ii) Assume that as $T, S \to \infty : S/T^2 \to 0$ and $T/S \to 0$, then:

$$P\left(\omega : \sup_{v \in \Re} \left| P^*\left(\max_{j=2,...,m} \sqrt{T} \int_U Z^{*}_{j,T,S}(u)\phi(u)du \leq v\right)\right.\right.$$

$$\left.\left. - P\left(\max_{j=2,...,m} \sqrt{T} \int_U \left(Z_{j,T,S}(u) - \left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2\right)\right)\phi(u)du \leq v\right)\right| > \varepsilon\right) \to 0,$$

Thus, for any bootstrap replication, compute the bootstrap statistic, $\sqrt{T}Z^{**}_{T,S}\left(\sqrt{T}Z^{*}_{T,S}\right)$. Perform B bootstrap replications (B large) and compute the quantiles of the resultant empirical distribution. Reject H_0 if $\sqrt{T}Z_{T,S}$ is greater than the $(1-\alpha)th$-quantile. Otherwise, do not reject. Now, for all samples except a set with probability measure approaching zero, $\sqrt{T}Z_{T,S}$ has the same limiting distribution as the corresponding bootstrapped statistic, when $\int_U \left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2\right)\phi(u)du = 0$ for all $j = 2, ..., m$. In this case, the above approach ensures that the test has asymptotic size equal to α. On the other hand, when one (or more) competing models is (are) strictly dominated by the benchmark, the approach ensures that the test has an asymptotic size between 0 and α. Finally, under the alternative, $Z_{T,S}$ diverges to (plus) infinity, while the corresponding bootstrap statistic has a well defined limiting distribution. This ensures unit asymptotic power. As our bootstrap mimics the limiting distribution of the statistics in the least favorable case for the null, inference is conservative (Hansen (2001) suggests recentering the bootstrap statistics using the sample mean, when the latter is larger than (minus) a bound of order $\sqrt{2T \log\log T}$).

4 Empirical Illustration

Consider a stochastic growth model characterized by flexible labor supply, non-zero capital depreciation, and two shocks. The first is the usual permanent shock to technology, and

the second is a transitory preference (or taste) shock. The model is exactly that considered by Christiano (1988) and the accompanying appendix (i.e. Christiano (1987)), except that we use a Cobb-Douglas production function (i.e. we set $\sigma = 0$ in his equation (2)), so that inventories do not play a role in production. Assume that the representative agent chooses a contingent plan for consumption and labor supply which maximizes

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\exp(u_t) \log C_t - \gamma H_t \right), \tag{3}$$

where C_t and H_t are per capita consumption and hours worked, respectively. Additionally, u_t is a transitory taste shock defined in (6) below. The maximization above is subject to the production constraint,

$$C_t \leq X_t - K_t + \frac{1-\delta}{n} K_{t-1}, \tag{4}$$

where X_t is per capita output, K_t per capita capital stock, n is the rate of growth of population (assumed constant), δ is the depreciation rate. Additionally,

$$X_t = (z_t H_t)^{1-\theta} K_{t-1}^{\theta}, \tag{5}$$

where z_t is a permanent technological shock, defined as $z_t = z_{t-1} \exp(t_t)$, and where θ denotes the capital share. Now, assume that

$$\begin{pmatrix} u_t \\ t_t \end{pmatrix} = \begin{pmatrix} 0 \\ a_2 \end{pmatrix} + \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} u_{t-1} \\ t_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}, \tag{6}$$

where $-1 < a_{11}, A_{22} < 1$. The structural parameters in this model are: $\beta, \delta, n, \theta, \gamma, a_2, A_{11}$, and A_{22}. As we cannot find explicit solution of (3) subject to (4), we instead solve the following quadratic approximation:

$$\max_{d_t} E_0 \sum_{t=0}^{\infty} \beta^t \left(c_1' d_t + c_2' s_t + s_t' R s_t + d_t' Q d_t + 2 s_t' F d_t \right) \tag{7}$$

subject to $s_{t+1} = \phi_0 + \phi_1 s_t + B d_t + \epsilon_t$, where $s_t^* = (k_{t-1}^*, u_t, t_t, u_{t-1}, t_{t-1})'$ and $d_t^* = (k_t^*, h_t^*)'$,with $k_t^* = \log(K_t / z_{t-1})$ and $h_t^* = \log(H_t)$; and where $c_1, c_2, R, Q, F, \phi_0, \phi_1, B$ are parameters constructed using functions of first and second derivatives. (For complete details, see Appendix B of Corradi and Swanson (2003b).) Assume that our objective is to find the dynamics of X_t and H_t, as given by the constrained maximization problem above. The optimal policy function governing the dynamics of d_t^* are given by,

$$d_t^* = K_0 + K_1 s_t^*, \tag{8}$$

where $K_0 = -\frac{1}{2}\overline{Q}^{-1}\overline{c}_1$, $K_1 = -\overline{Q}^{-1}\overline{F}'$, K_0 is a 2×1 vector and K_1 is a 2×5 matrix. In this framework,it follows from Christiano (1988) that

$$\overline{c}_1' = c_1' + \beta V_s' B + 2\beta\phi_0' V_Q B, \tag{9}$$

$$\overline{Q} = Q + \beta B' V_Q B, \text{ and} \tag{10}$$

$$\overline{F} = F + \beta\phi_1' V_Q B. \tag{11}$$

In order to determine (8), we need to construct K_0 and K_1 by solving 2 Riccati equations for V_Q and V_s based on (9), (10), and (11). We begin with the determination of V_Q. V_Q is the solution of the following Riccati equation:

$$V_{Q,t} = R + \beta\phi_1' V_{Q,t+1}\phi_1 - (\beta\phi_1' V_{Q,t+1} B + F) \times (Q + \beta B' V_{Q,t+1} B)^{-1} (\beta B' V_{Q,t+1}\phi_1 + F') \tag{12}$$

An explicit closed-form solution to this Riccati equation does not exist. However we can find an arbitrarily accurate solution via the iteration method discussed in MacGrattan (1990). In particular, starting with an initial value, say $V_{Q,0}$ (typically set equal to the identity matrix time a small number), at the $n+1$-th iteration, we have

$$V_Q^{n+1} = R + \beta\phi_1' V_Q^n \phi_1 - (\beta\phi_1' V^n B + F) \times \left(Q + \beta B' V_Q^n B\right)^{-1} \left(\beta B' V_Q^n \phi_1 + F'\right). \tag{13}$$

The stopping point occurs at that value of V_Q^{n+1} for which $\left\| V_Q^{n+1} - V_Q^n \right\| < \varepsilon$, where $\|.\|$ denotes the euclidean norm, and where we set $\varepsilon = 0.0000001$. Now, in much the same manner, and by using V_Q^{n+1} we can obtain an (approximate) solution for the Riccati equation describing V_s. Namely,

$$V_s^{n+1} = (I - \beta(K_1' B' + \phi_1'))^{-1} \left(K_1' \left(c_1 + 2\beta B' V_Q^{n+1}\phi_0\right) + c_2 + 2\beta\phi_1' V_Q^{n+1}\phi_0\right). \tag{14}$$

Now,

$$d_t^* = \begin{pmatrix} k_t^* \\ h_t^* \end{pmatrix} = K_0 + K_1 s_t^* = \begin{pmatrix} k_{01} \\ k_{02} \end{pmatrix} + \begin{pmatrix} k_{11} & k_{12} & k_{13} & k_{14} & k_{15} \\ k_{21} & k_{22} & k_{23} & k_{24} & k_{25} \end{pmatrix} \begin{pmatrix} k_{t-1}^* \\ u_t \\ t_t \\ u_{t-1} \\ t_{t-1} \end{pmatrix}, \tag{15}$$

where K_0 and K_1 are defined as above. Thus,

$$\begin{pmatrix} k_t^* \\ h_t^* \end{pmatrix} = \begin{pmatrix} k_{01}+k_{13}a_2 \\ k_{02}+k_{23}a_2 \end{pmatrix} + \begin{pmatrix} k_{11} \\ k_{21} \end{pmatrix} k_{t-1}^* + \begin{pmatrix} k_{12}A_{11}+k_{14} & k_{13}A_{22}+k_{15} \\ k_{22}A_{11}+k_{24} & k_{23}A_{22}+k_{25} \end{pmatrix} \begin{pmatrix} u_{t-1} \\ t_{t-1} \end{pmatrix} + \begin{pmatrix} k_{12} & k_{13} \\ k_{22} & k_{23} \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \tag{16}$$

Finally, recalling that $k_t^* = \log(K_t/z_{t-1})$ and $h_t^* = \log(H_t)$, and recalling (6), we can simulate capital and hours using:

$$\begin{pmatrix} \log K_t \\ \log H_t \end{pmatrix} = \begin{pmatrix} \log z_{t-1} \\ 0 \end{pmatrix} + \begin{pmatrix} k_{01}+k_{13}a_2 \\ k_{02}+k_{23}a_2 \end{pmatrix} + \begin{pmatrix} k_{11} \\ k_{21} \end{pmatrix} (\log K_{t-1} - \log z_{t-2}) + \begin{pmatrix} k_{12}A_{11}+k_{14} & k_{13}A_{22}+k_{15} \\ k_{22}A_{11}+k_{24} & k_{23}A_{22}+k_{25} \end{pmatrix} \begin{pmatrix} u_{t-1} \\ t_{t-1} \end{pmatrix} + \begin{pmatrix} k_{12} & k_{13} \\ k_{22} & k_{23} \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \tag{17}$$

Note that all of the parameters in (17) are explicit functions of the structural parameters $\beta, \delta, n, \theta, \gamma, a_2, A_{11}$ andA_{22}. Note also that for $|k_{11}| < 1$, $\log K_t - \log z_{t-1}$ is $I(0)$, and therefore, for $|k_{21}| < 1$, $\log H_t$ is $I(0)$ too. Also, as $\log z_{t-1}$ is $I(1)$, $\log K_t$ is $I(1)$, so that both capital and output (see (18) below) follow unit root processes in logs. Finally, from the production function in (5), upon taking logs,

$$\log X_t = (1-\theta)\log z_t - \theta \log n + (1-\theta)\log H_t + \theta \log K_{t-1} \tag{18}$$

In the sequel, we take values for structural parameters $\beta, \delta, n, \theta, \gamma, a_2, A_{11}$, and A_{22} from King, Plosser and Rebelo (1988a,b) and Christiano (1988). In particular, and unless otherwise stated, we set $\beta = 0.99$, $n = 1.012$, $\delta = 0.025$, $\theta = 0.50$, $\gamma = 0.0028$, $a_2 = 0.0044$, $A_{22} = -0.10$, $A_{11} = 0.95$. Furthermore, and unless otherwise stated, we estimate $\sigma_{\epsilon 1}^2$ and $\sigma_{\epsilon 2}^2$ by simply constructing simulations with various values for these variances and minimizing the difference between the simulated and actual variance of per capita output and hours growth.[10,11] Given this setup, we evaluate the joint distribution of $Y_t = (\log \Delta X_t, \log \Delta H_t)$ as well as $Y_t = (\Delta \log X_t, \Delta \log H_t, \Delta \log X_{t-1}, \Delta \log H_{t-1}))$.

[10]Throughout, and for simplicity, we assume that the covariance between the shocks is zero. This assumption is not inconsistent with some of the empirical findings in the literature.

[11]Note that as we are interested in matching the covariance structure between historical and simulated data, we can use different specifications of the marginal distribution of the shocks. In the current context, the variance estimators we use depend also the simulated sample size S. Call these estimators $\widehat{\xi}_{T,S}$. If $S/T \to 0$, then simulation error is negligible and Assumption A2 is satisfied as $\widehat{\xi}_{T,S}$ is asymptotically equivalent to GMM. In the case where $S/T \to \delta > 0$, then we also require that the scaled difference between the simulated and actual sample moments is asymptotically normal. This is in general the case, provided that Assumption A3 holds.

Some illustrative results are given in Tables 1 and 2. The first column of numerical entries reports $\sqrt{T}Z_{T,S}$ for $T = 70$ (corresponding to the 70 annual U.S. per capita output growth observations used for the period 1933-2002), and for $S = \{T, 2T, 5T, 10T\}$.[12] All statistics are based on use of 20 uniformly distributed values for u across the range of the actual data. The next four columns of entries contain 5% and 10% bootstrap critical values constructed using both $\sqrt{T}Z^{**}_{T,S}$ and $\sqrt{T}Z^{*}_{T,S}$ (i.e. $T/S- > \delta > 0$ and $T/S- > 0$, respectively). For the case where $S = T$, we set the bootstrap block length equal to $l1 = 2$, $l2 = 5$, and $l3 = 7$. For all other cases, where $S = aT$, say, we set li equal to "a" times the corresponding value of li when $S = T$, for $i = 1, 2, 3$. Thus, "a" takes the values 2,5, and 10, depending on the magnitude of S. Bootstrap empirical distributions are constructed using 100 bootstrap replications (i.e. $B = 100$). Even though hours and output are examined, for the sake of brevity we only discuss output in the sequel. Various summary statistics (including *% <0, min, max,* and *standard error*) constructed using the simulated data are given in the remaining columns of the tables. Corresponding values based on the actual data are: *%<0 = 0.213; min = -0.128; max = 0.158; mean = 0.027*; and *standard error = 0.044*. Of course, the $Z_{T,S}$ statistics are based on the joint evaluation of *all* variables in Y_t, including output and hours. We consider four different scenarios: *Part I - Compare 3 models using* $Y_t = (\log \Delta X_t, \log \Delta H_t)$. *Set* $\beta = \{0.95, 0.90, 0.99\}$. *All shocks are normally distributed with variance estimated as discussed above. The model with* $\beta = 0.99$ *is the benchmark. Part II - Compare 3 models using* $Y_t = (\log \Delta X_t, \log \Delta H_t)$. *The benchmark has shocks drawn from a* $1/3t_3$ *distribution. One alternative assumes normality and uses estimated variances while the other assumes normality and sets* $\sigma^2_{\epsilon 1} = 0.039$ *and* $\sigma^2_{\epsilon 2} = 0.077$ *(as in Christiano (1988)). Part III - As Part I, except that* $Y_t = (\Delta \log X_t, \Delta \log H_t, \Delta \log X_{t-1}, \Delta \log H_{t-1}))$. *Part IV - As Part II, except that* $Y_t = (\Delta \log X_t, \Delta \log H_t, \Delta \log X_{t-1}, \Delta \log H_{t-1}))$.

Turning to our results, notice in Part I of Table 1 that the null hypothesis fails to reject in almost all cases, and never rejects when the simulation sample is 5 or 10 times as large as the sample size. This suggests that the discount rate can equally take values between 0.90 and 0.99, at least when the objective is to match the joint dynamics of simulated output and hours as closely as possible with the historical record. However, notice in Part II of Table

[12]Note that in all case $S/T > 0$, however, when for example $T = 70$ and $S = 700$, then simulation error is likely to be negligible with respect to parameter estimation error.

1 that the benchmark where shocks follow Student's t-distributions is soundly rejected in every case. Furthermore, the dynamics of output are clearly far from reasonable under the benchmark as well as under the first alternative model (see Sim1 and Sim2 summary statistics), corresponding to the t-distribution assumption (Sim1) and the case where we assume normality and fix $\sigma^2_{\epsilon 1} = 0.039$ *and* $\sigma^2_{\epsilon 2} = 0.077$ (Sim2). This supports the use of our simple simulated GMM type scheme for estimating the shock variances (see Sim3 results). Table 2 contains results for experiments analogous to those reported in Table 1, except that the joint dynamics of four variables are examined. The results of these experiments are the same as those reported on in Table 1, suggesting that generalizing our approach to the assessment of more variables does not affect test performance.

5 Concluding Remarks

In this paper we propose a test for comparing the joint distributions of historical time series with those simulated under a given DSGE model via a distributional generalization of the reality check of White (2000) in which we assess whether competing models are more accurate approximations to the "true" distribution than a given benchmark model, in a squared error sense. Two empirical versions of the block bootstrap are used to construct valid asymptotic critical values. Finally, an illustrative example is given in which the testing approach is applied to an RBC model, and it is found that RBC models are quite sensitive to distributional and error variance magnitude assumptions, but are less sensitive to small changes in primitive parameter values.

6 Appendix

Hereafter, P, P^*, Ω, P_i, P_i^* and Ω_i, $i = 0, 1, ..., m$ are defined above the statement of Proposition 2. Also, $o_{P^*}(1)$ Pr $-P$ and $o_{P_i^*}(1)$ Pr $-P$ denote terms approaching zero in probability P^*, $\forall \omega \in \Omega$ and P_i^*, $\forall \omega_i \in \Omega_i$, respectively. Furthermore, E_{P^*} and Var_{P^*} are mean and variance operators with respect to P^*, with $E_{P_i^*}$, and $Var_{P_i^*}$ defined analogouly. Finally, *czp* is meant to mean "conditional on the sample and for all samples except a set of zero probability measure", while *cszp* is the same, except applies to all samples, simulated and empirical.

Proof of Proposition 1: **(i)** We begin by examining the limiting distribution, pointwise in u. Now,

$$\sqrt{T}Z_{j,T,S}(u) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Bigg((1\{Y_t \leq u\} - F_1(u)) - \Bigg(\frac{1}{S}\sum_{n=1}^{S} 1\{Y_{1,n}(\theta_1^\dagger) \leq u - (Y_{1,n}(\widehat{\theta}_{1,T}) - Y_{1,n}(\theta_1^\dagger))\}$$

$$-F_1(u - (Y_{1,n}(\widehat{\theta}_{1,T}) - Y_{1,n}(\theta_1^\dagger)))\Big) - \Big(F_1(u - (Y_{1,n}(\widehat{\theta}_{1,T}) - Y_{1,n}(\theta_1^\dagger))) - F_1(u)\Big)\Big)^2$$

$$-\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Bigg((1\{Y_t \leq u\} - F_j(u)) - \frac{1}{S}\sum_{n=1}^{S}\Big(1\{Y_{j,n}(\theta_j^\dagger) \leq u - (Y_{j,n}(\widehat{\theta}_{j,T}) - Y_{j,n}(\theta_j^\dagger))\}$$

$$-F_j(u - (Y_{j,n}(\widehat{\theta}_{j,T}) - Y_{j,n}(\theta_j^\dagger)))\Big) - \frac{1}{S}\sum_{n=1}^{S}\Big(F_j(u - (Y_{j,n}(\widehat{\theta}_{j,T}) - Y_{j,n}(\theta_j^\dagger))) - F_j(u)\Big)\Bigg)^2 \quad (19)$$

Given A3(i), $\frac{1}{S^{1/2}}\sum_{n=1}^{S}(1\{Y_{j,n}(\theta_j^\dagger) \leq u\} - F_j(u))$ is stochastic equicontinuous in $u \in U$, from Theorem 1 (Application 1) of Doukhan, Massart and Rio (1995). Also, given A2 and A3(ii),

$$Y_{j,n}(\widehat{\theta}_{j,T}) - Y_{j,n}(\theta_j^\dagger) = \nabla_{\theta_j} Y_{j,n}(\overline{\theta}_{j,T})\left(\widehat{\theta}_{j,T} - \theta_j^\dagger\right) = \frac{1}{\sqrt{T}}\nabla_{\theta_j} Y_{j,n}(\overline{\theta}_{j,T}) O_P(1),^{13}$$

with $\overline{\theta}_{j,T} \in \left(\widehat{\theta}_{j,T}, \theta_j^\dagger\right)$. Now, given the domination conditions in A3(iii), by Chebyshev's inequality,

$$\Pr\left(\frac{1}{\sqrt{T}}\sup_{n\leq S}\left|\nabla_{\theta_j} Y_{j,n}(\theta_j)\right| > \varepsilon\right) \leq \sum_{i=1}^{S}\Pr\left(\frac{1}{\sqrt{T}}\left|\nabla_{\theta_j} Y_{j,n}(\theta_j)\right| > \varepsilon\right)$$

$$\leq \sum_{i=1}^{S}\Pr\left(\frac{1}{\sqrt{T}}D_j > \varepsilon\right) \leq \frac{1}{\varepsilon^4}\frac{S}{T^2}E(D_j^4) \to 0$$

as $S/T^2 \to 0$, given that $S/T \to \delta > 0$. Thus,

$$\begin{aligned}
&\sqrt{\frac{T}{S}}\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{1,n}(\theta_1^\dagger) \le u - (Y_{1,n}(\widehat{\theta}_{1,T}) - Y_{1,n}(\theta_1^\dagger))\} - \left(F_1(u - (Y_{1,n}(\widehat{\theta}_{1,T}) - Y_{1,n}(\theta_1^\dagger))) - F_1(u)\right)\right) \\
= \;& \sqrt{\frac{T}{S}}\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{1,n}(\theta_1^\dagger) \le u\} - F_1(u)\right) + o_P(1),
\end{aligned}$$

with the $o_P(1)$ term holding uniformly in u. Noticing that $F_j(u - (Y_{j,n}(\widehat{\theta}_{j,T}) - Y_{j,n}(\theta_1^\dagger))) = f_j(u-(Y_{j,n}(\overline{\theta}_{j,T})-Y_{j,n}(\theta_j^\dagger)))\nabla_{\theta_j}Y_{j,n}(\overline{\theta}_{j,T})'\left(\widehat{\theta}_{j,T} - \theta_j^\dagger\right)$, where f_j is the density associated with F_j, after some elementary manipulations, the right hand side (RHS) of (19) can be written as:

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left((1\{Y_t \le u\} - F_0(u)) - (F_1(u) - F_0(u)) - \left(\frac{1}{S}\sum_{n=1}^{S}1\{Y_{1,n}(\theta_1^\dagger) \le u\} - F_1(u)\right)\right. \\
&\left.-\frac{1}{S}\sum_{n=1}^{S}f_1(u - (Y_{1,n}(\overline{\theta}_{1,T}) - Y_{1,n}(\theta_1^\dagger)))\nabla_{\theta_1}Y_{1,n}(\overline{\theta}_{1,T})'\left(\widehat{\theta}_{1,T} - \theta_1^\dagger\right)\right)^2 \\
&-\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left((1\{Y_t \le u\} - F_0(u)) - (F_j(u) - F_0(u)) - \left(\frac{1}{S}\sum_{n=1}^{S}1\{Y_{j,n}(\theta_j^\dagger) \le u\} - F_j(u)\right)\right. \\
&\left.-\frac{1}{S}\sum_{n=1}^{S}f_j(u - (Y_{j,n}(\overline{\theta}_{j,T}) - Y_{j,n}(\theta_j^\dagger)))\nabla_{\theta_j}Y_{j,n}(\overline{\theta}_{j,T})'\left(\widehat{\theta}_{j,T} - \theta_j^\dagger\right)\right)^2 + o_P(1),
\end{aligned}$$

where the $o_P(1)$ term holds uniformly in u. Now,

$$\sqrt{T}\left(\frac{1}{S}\sum_{n=1}^{S}f_j(u - (Y_{j,n}(\overline{\theta}_{j,T}) - Y_{j,n}(\theta_j^\dagger)))^2\nabla_{\theta_j}Y_{j,n}(\overline{\theta}_{j,T})'\left(\widehat{\theta}_{j,T} - \theta_j^\dagger\right)\right)^2 = o_P(1), \tag{20}$$

uniformly in u, as $\left(\widehat{\theta}_{j,T} - \theta_j^\dagger\right) = O_P(T^{-1/2})$. Given A1 and A3, uniformly in u,

$$\sqrt{T}\left(\frac{1}{S}\sum_{n=1}^{S}1\{Y_{j,n}(\theta_j^\dagger) \le u\} - F_j(u)\right)^2 = O_P\left(\frac{\sqrt{T}}{S}\right) = o_P(1), \tag{21}$$

and

$$\begin{aligned}
&\frac{1}{S}\sum_{n=1}^{S}\left(1\{Y_{j,n}(\theta_j^\dagger) \le u\} - F_j(u)\right)\frac{1}{S}\sum_{n=1}^{S}f_j(u - (Y_{j,n}(\overline{\theta}_{j,T}) - Y_{j,n}(\theta_j^\dagger)))\nabla_{\theta_j}Y_{j,n}(\overline{\theta}_{j,T})'\sqrt{T}\left(\widehat{\theta}_{j,T} - \theta_j^\dagger\right) \\
= \;& o_P(1).
\end{aligned} \tag{22}$$

Now, let $\mu_{F_j}(u) = E(1\{Y_t \leq u\} - F_j(u))$, and $\mu_{f_j,\theta_j^\dagger}(u') = E\left(f_j(u)\nabla_{\theta_j} Y_{j,n}(\theta_j^\dagger)\right)$, where f_j is the density associated with F_j. Given A1 and A2, by the uniform law of large numbers for mixing processes,

$$\begin{aligned}
&\sqrt{T}Z_{j,T,S}(u) - \sqrt{T}\left(\left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2\right)\right) \\
= \; & -\frac{2}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{Y_t \leq u\} - F_0(u)\right)\left(F_1(u) - F_j(u)\right)
\end{aligned}$$

$$\begin{aligned}
&-2\mu_{F_1}(u)\delta\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{1,n}(\theta_1^\dagger) \leq u\} - F_1(u)\right) + 2\mu_{F_j}(u)\delta\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{j,n}(\theta_j^\dagger) \leq u\} - F_j(u)\right) \\
&-2\mu_{F_1}(u)\mu_{f_1,\theta_1^\dagger}(u)'\sqrt{T}\left(\widehat{\theta}_{1,T} - \theta_1^\dagger\right) + 2\mu_{F_j}(u)\mu_{f_j,\theta_j^\dagger}(u)'\sqrt{T}\left(\widehat{\theta}_{j,T} - \theta_j^\dagger\right) + o_P(1).
\end{aligned}$$

Therefore, pointwise in u, $\sqrt{T}\left(Z_{j,T,S}(u) - \left((F_0(u) - F_1(u))^2 - (F_0(u) - F_j(u))^2\right)\right)$ satisfies a CLT, and the asymptotic variance is as given in the statement of the proposition. Convergence of the finite dimensional distribution follows immediately upon application of the Cramer Wold device. Finally, given A1 and A3, $\frac{2}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{Y_t \leq u\} - F_0(u)\right)$ and $\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{j,n}(\theta_1^\dagger) \leq u\} - F_j(u)\right)$ are stochastic equicontinuous in u (e.g. by Theorem 1, Application 1 in Doukhan, Massart and Del Rio (1995)). The desired result follows by the continuous mapping theorem.

(ii) The proof follows using the same argument as in part (i), and when the expression for $K_j(u,u')$ in part (i) is adjusted by setting $\delta = 0$.

Proof of Proposition 2: (i) Note that:

$$\begin{aligned}
\sqrt{T}Z^{**}_{j,T,S}(u) \; = \; & \frac{1}{T^{1/2}}\sum_{t=1}^{T}\left(\left((1\{Y_t^* \leq u\} - F_1(u)) - \frac{1}{S}\sum_{n=1}^{S}(1\{Y^*_{1,n}(\widehat{\theta}^*_{1,T}) \leq u\} - F_1(u))\right)^2\right. \\
& \left. - \left((1\{Y_t \leq u\} - F_1(u)) - \frac{1}{S}\sum_{n=1}^{S}(1\{Y_{1,n}(\widehat{\theta}_{1,T}) \leq u\} - F_1(u))\right)^2\right) \\
& - \frac{1}{T^{1/2}}\sum_{t=1}^{T}\left(\left((1\{Y_t^* \leq u\} - F_j(u)) - \frac{1}{S}\sum_{n=1}^{S}(1\{Y^*_{j,n}(\widehat{\theta}^*_{j,T}) \leq u\} - F_j(u))\right)^2\right. \\
& \left. - \left((1\{Y_t \leq u\} - F_j(u)) - \frac{1}{S}\sum_{n=1}^{S}(1\{Y_{j,n}(\widehat{\theta}_{j,T}) \leq u\} - F_j(u))\right)^2\right). \qquad (23)
\end{aligned}$$

Now, we begin by considering the first term on the RHS of (23), which can be rewritten as,

$$= \frac{1}{T^{1/2}}\sum_{t=1}^{T}\left((1\{Y_t^* \leq u\} - F_1(u)) - \frac{1}{S}\sum_{n=1}^{S}\left(1\{Y^*_{1,n}(\theta_1^\dagger) \leq u - (Y^*_{1,n}(\widehat{\theta}^*_{1,T}) - Y^*_{1,n}(\theta_1^\dagger))\right.\right.$$

$$-F_1(u-(Y^*_{1,n}(\widehat{\theta}^*_{1,T})-Y^*_{1,n}(\theta^\dagger_1)))\Big)-\frac{1}{S}\sum_{n=1}^{S}\Big(F_1(u-(Y^*_{1,n}(\widehat{\theta}^*_{1,T})-Y^*_{1,n}(\theta^\dagger_1)))-F_1(u)\Big)\Bigg)^2$$
$$-\frac{1}{T^{1/2}}\sum_{t=1}^{T}\Bigg((1\{Y_t\le u\}-F_1(u))-\frac{1}{S}\sum_{n=1}^{S}\Big(1\{Y_{1,n}(\theta^\dagger_1)\le u-(Y_{1,n}(\widehat{\theta}_{1,T})-Y_{1,n}(\theta^\dagger_1))\}$$
$$-F_1(u-(Y_{1,n}(\widehat{\theta}_{1,T})-Y_{1,n}(\theta^\dagger_1)))\Big)-\frac{1}{S}\sum_{n=1}^{S}\Big(F_1(u-(Y_{1,n}(\widehat{\theta}_{1,T})-Y_{1,n}(\theta^\dagger_1)))-F_1(u)\Big)\Bigg)^2$$

$$\begin{aligned}
= \;& \frac{1}{T^{1/2}}\sum_{t=1}^{T}\Bigg((1\{Y^*_t\le u\}-F_1(u))-\frac{1}{S}\sum_{n=1}^{S}\Big(1\{Y^*_{1,n}(\theta^\dagger_1)\le u\}-F_1(u)\Big)\\
& -\frac{1}{S}\sum_{n=1}^{S}f_1\Big(u-(Y^*_{1,n}(\overline{\theta}^*_{1,T})-Y^*_{1,n}(\theta^\dagger_1))\Big)\nabla_{\theta_1}Y^*_{1,n}(\overline{\theta}^*_{1,T})\sqrt{T}\Big(\widehat{\theta}^*_{1,T}-\theta^\dagger_1\Big)\Bigg)^2\\
& -\frac{1}{T^{1/2}}\sum_{t=1}^{T}\Bigg((1\{Y_t\le u\}-F_1(u))-\frac{1}{S}\sum_{n=1}^{S}\Big(1\{Y_{1,n}(\theta^\dagger_1)\le u\}-F_1(u)\Big)\\
& -\frac{1}{S}\sum_{n=1}^{S}f_1\Big(u-(Y_{1,n}(\overline{\theta}_{1,T})-Y_{1,n}(\theta^\dagger_1))\Big)\nabla_{\theta_1}Y^*_{1,n}(\overline{\theta}_{1,T})\sqrt{T}\Big(\widehat{\theta}_{1,T}-\theta^\dagger_1\Big)\Bigg)^2.
\end{aligned}$$

The second term on the RHS of (23) can be treated in the same way, by replacing 1 with j. Therefore, after some elementary manipulations, $Z^{**}_{j,T,S}(u)$ can be written as:

$$Z^{**}_{j,T,S}(u)=-\frac{2}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{Y^*_t\le u\}-1\{Y_t\le u\}\right)\left(F_1(u)-F_j(u)\right)$$

$$-d_{1,1}\sqrt{\frac{T}{S}}\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\Big(1\{Y^*_{1,n}(\theta^\dagger_1)\le u\}-F_1(u)\Big)+d_{2,1}\sqrt{\frac{T}{S}}\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\Big(1\{Y_{1,n}(\theta^\dagger_1)\le u\}-F_1(u)\Big)$$

$$+d_{1,j}\sqrt{\frac{T}{S}}\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\Big(1\{Y^*_{j,n}(\theta^\dagger_j)\le u\}-F_j(u)\Big)-d_{2,j}\sqrt{\frac{T}{S}}\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\Big(1\{Y_{j,n}(\theta^\dagger_j)\le u\}-F_j(u)\Big)$$
$$-d_{1,1}\frac{1}{S}\sum_{n=1}^{S}f_1\Big(u-(Y^*_{1,n}(\overline{\theta}^*_{1,T})-Y^*_{1,n}(\theta^\dagger_1))\Big)\nabla_{\theta_1}Y^*_{1,n}(\overline{\theta}^*_{1,T})\sqrt{T}\Big(\widehat{\theta}^*_{1,T}-\theta^\dagger_1\Big)$$
$$+d_{2,1}\frac{1}{S}\sum_{n=1}^{S}f_1\Big(u-(Y_{1,n}(\overline{\theta}_{1,T})-Y_{1,n}(\theta^\dagger_1))\Big)\nabla_{\theta_1}Y_{1,n}(\overline{\theta}_{1,T})\sqrt{T}\Big(\widehat{\theta}_{1,T}-\theta^\dagger_1\Big)$$
$$+d_{1,j}\frac{1}{S}\sum_{n=1}^{S}f_j\Big(u-(Y^*_{j,n}(\overline{\theta}^*_{j,T})-Y^*_{j,n}(\theta^\dagger_j))\Big)\nabla_{\theta_j}Y^*_{j,n}(\overline{\theta}^*_{j,T})\sqrt{T}\Big(\widehat{\theta}^*_{j,T}-\theta^\dagger_j\Big)$$
$$-d_{2,j}\frac{1}{S}\sum_{n=1}^{S}f_j\Big(u-(Y_{j,n}(\overline{\theta}_{j,T})-Y_{j,n}(\theta^\dagger_j))\Big)\nabla_{\theta_j}Y_{j,n}(\overline{\theta}_{j,T})\sqrt{T}\Big(\widehat{\theta}_{j,T}-\theta^\dagger_j\Big),$$

where $d_{1,x} = \frac{2}{T}\sum_{t=1}^{T}\left(1\{Y_t^* \leq u\} - F_x(u)\right)$ and $d_{2,x} = \frac{2}{T}\sum_{t=1}^{T}\left(1\{Y_t \leq u\} - F_x(u)\right)$, $x = 1, j$. Now, conditional on the sample, $Y_t^*, Y_{t+1}^*, ..., Y_{t+l-1}^*$, is *iid* uniform, for $t = 2, ...T - l$. Analogously, $Y_{j,t}^*, Y_{j,t+1}^*, ..., Y_{j,t+l-1}^*$, conditional on the (simulated) sample, is also *iid* uniform. Thus, given A1-A3, and by the basic properties of the block bootstrap (Künsch (1989), theorem 3.5), we also have that pointwise in u :

$$\frac{1}{T}\sum_{t=1}^{T}\left(1\{Y_t^* \leq u\} - F_j(u)\right) = \frac{1}{T}\sum_{t=1}^{T}\left(1\{Y_t \leq u\} - F_j(u)\right) + o_{P_0^*}(1), \ \Pr - P_0, \tag{24}$$

and as $l_T/T^{1/2} \to 0$,

$$\frac{1}{T}\sum_{t=1}^{T}\left(1\{Y_t^* \leq u\} - F_j(u)\right) = \mu_{F_j}(u) + o_{P_0^*}(1), \ \Pr - P_0. \tag{25}$$

Uniform convergence on U follows because of the stochastic equicontinuity of the left hand side of (24) and (25). Also, by the same argument used in Lemma A4 of Goncalves and White (2002: GW), and recalling A3(iv), it follows that for $j = 1, ..., m$:

$$P_j^*\left(\sup_{\theta_j \in \Theta_j}\sup_{u \in U}\left|\frac{1}{S}\sum_{n=1}^{S} f_j\left(u - \left(Y_{j,n}^*(\theta_j) - Y_{j,n}^*(\theta_j^\dagger)\right)\right)\nabla_{\theta_j} Y_{j,n}^*(\theta_j) - \mu_{f_{j,\theta_j^\dagger}}(u)\right| > \varepsilon\right) \to 0, \ \Pr - P_j, \tag{26}$$

Thus, given (24)-(26), and using arguments similar to those used in the proof of Proposition 1, we see that after some elementary manipulations $\sqrt{T}Z_{j,T,S}^{**}(u)$ can be written as:

$$\sqrt{T}Z_{j,T,S}^{**}(u) = -\frac{2}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{Y_t^* \leq u\} - 1\{Y_t \leq u\}\right)\left(F_1(u) - F_j(u)\right)$$

$$-2\mu_{F_1}(u)\delta\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{1,n}^*(\theta_1^\dagger) \leq u\} - 1\{Y_{1,n}(\theta_1^\dagger) \leq u\}\right)$$
$$+2\mu_{F_j}(u)\delta\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{j,n}^*(\theta_j^\dagger) \leq u\} - 1\{Y_{j,n}(\theta_j^\dagger) \leq u\}\right)$$
$$+2\mu_{F_1}(u)\mu_{f_{1,\theta_1^\dagger}}(u)\sqrt{T}\left(\widehat{\theta}_{1,T}^* - \widehat{\theta}_{1,T}\right) - 2\mu_{F_j}(u)\mu_{f_{j,\theta_j^\dagger}}(u)\sqrt{T}\left(\widehat{\theta}_{j,T}^* - \widehat{\theta}_{j,T}\right) + o_{P^*}(1), \ \Pr - P.$$

By Theorem 2.2 in GW, $\sqrt{T}\left(\widehat{\theta}_{j,T}^* - \widehat{\theta}_{j,T}\right)$ has the same limiting distribution as $\sqrt{T}\left(\widehat{\theta}_{j,T}^* - \theta_j^\dagger\right)$, *czp*. Thus, by Theorem 3.5 in Künsch (1989), we see that pointwise in u, $Z_{j,T,S}^{**}(u)$ has the same limiting distribution as $Z_{j,T,S}(u)$, *cszp* P_0 and P_j, $j = 1, ..., m$ probability measure, respectively. By the Cramer Wold device, the same holds for any finite set of points in U.

Finally, by the empirical version of the block bootstrap (see e.g. Naik-Nimbalkar and Rajarshi (1994, Theorem 2.1)), it follows that $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{Y_t^* \leq u\} - 1\{Y_t \leq u\}\right)$ has the same limiting distribution as $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{Y_t \leq u\} - F_0(u)\right)$, as a process on U, *czp* P_0-probability measure. Also, $\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{j,n}^*(\theta_j^\dagger) \leq u\} - 1\{Y_{j,n}(\theta_j^\dagger) \leq u\}\right)$ has the same limiting distribution as $\frac{1}{\sqrt{S}}\sum_{n=1}^{S}\left(1\{Y_{j,n} \leq u\} - F_j(u)\right)$, as a process on U, conditional on the simulated sample and for all simulated samples except a set of zero P_j-probability measure, $j = 1, ..., m$. The statement in part (i) of the theorem then follows as a straightforward consequence of the continuous mapping theorem.

(ii) The proof follows as a special case of part (i).

7 References

Andrews, D.W.K., (1997), A Conditional Kolmogorov Test, *Econometrica*, 65, 1097-1128.

Andrews, D.W.K., (2002), Higher-Order Improvements of a Computationally Attractive $k-$step Bootstrap for Extremum Estimators, *Econometrica,* 70, 119-162.

Bai, J., (2003), Testing Parametric Conditional Distributions of Dynamic Models, *Review of Economics and Statistics,* 85, 531-549

Bierens, H.J., (2003), Econometric Analysis of Singular Dynamic Stochastic General Equilibrium Models with an Application to the King-Plosser-Rebelo Stochastic Growth Model, Working Paper, Pennsylvania State University.

Bierens, H.J., and N.R. Swanson, (2000), The Econometric Consequences of the Ceteris Paribus Condition in Theoretical Economics, *Journal of Econometrics*, 95, 223-253.

Chang, Y.S., J.F. Gomes, and F. Schorfheide, (2002), Learning-by-Doing as a Propagation Mechanism, *American Economic Review, 92, 1498-1520.*

Christiano, L.J., (1987), Technical Appendix to 'Why Does Inventory Investment Fluctuate So Much', Federal Reserve Bank of Minneapolis, Working Paper 380.

Christiano, L.J., (1988), Why Does Inventory Investment Fluctuate So Much, *Journal of Monetary Economics,* 21, 247-280.

Christiano, L.J., and M. Eichenbaum, (1992), Current Real Business Cycles Theories and Aggregate Labor Market Fluctuations, *American Economic Review,* 82-430-450.

Cogley, T., and J.M. Nason, (1995), Output Dynamics for Real Business Cycles Models, *American Economic Review,* 85, 492-511.

Corradi, V., N.R. Swanson, (2002), A Test For Comparing Multiple Misspecified Conditional Distribution Models, Working Paper, Queen Mary, University of London and Rutgers University.

Corradi, V., N.R. Swanson, (2003a), The Block Bootstrap for Recursive m-Estimators With Applications to Predictive Evaluation, Working Paper, Queen Mary, University of London and Rutgers University.

Corradi, V., N.R. Swanson, (2003b), Evaluation of Dynamic Stochastic General Equilibrium Models Based on Distributional Comparison of Simulated and Historical Data, Working Paper, Queen Mary, University of London and Rutgers University.

DeJong, D.N., B.F. Ingram, and C.H. Whiteman (2000), A Bayesian Approach to Dynamic Macroeconomics, *Journal of Econometrics,* 98, 203-223.

Danielsson, J., and J.F. Richard, (1993), Accelerated Gaussian Importance Sampler With Applications to Dynamic Latent Variable Models, *Journal of Applied Econometrics*, 8, S153-S173.

Diebold, F.X., and R.S. Mariano, (1995), Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.

Diebold, F.X., L.E. Ohanian, and J. Berkowitz, (1998), Dynamic Equilibrium Economies: A Framework for Comparing Models and Data, *Review of Economic Studies,* 65, 433-451.

Doukhan, P., P. Massart and E. Rio, (1995), Invariance Principles for Absolutely Regular Empirical Processes, *Annales de l'Institute Henri Poincare: Probabilities et Statistique,* 31, 393-427.

Duffie, D. and K. Singleton, (1993), Simulated Moment Estimation of Markov Models of Asset Prices, *Econometrica,* 61, 929-952.

Fernandez-Villaverde, J., and J.F. Rubio-Ramirez, (2001), Comparing Dynamic Equilibrium Models to Data, Manuscript, University of Pennsylvania.

Gallant A.R. and H. White, *Inference for Nonlinear Dynamic Models,* Blackwell, Oxford.

Giacomini, R. (2002), Comparing Density Forecasts via Weighted Likelihood Ratio Tests: Asymptotic and Bootstrap Methods, Working Paper, University of California, San Diego.

Goncalves, S., and H. White, (2002), Maximum Likelihood and the Bootstrap for Nonlinear Dynamic Models, *Journal of Econometrics*, forthcoming.

Hall, P., and J.L. Horowitz, (1996), Bootstrap Critical Values for Tests Based on Generalized Method of Moments Estimators, *Econometrica,* 64, 891-916.

Hansen, P.R., (2001), An Unbiased Test for Superior Predictive Ability, Working Paper, Brown University.

Horowitz, J.L., (2002), Bootstrap Methods for Markov Processes, Working Paper, *Econometrica,* forthcoming.

Inoue, A., and M. Shintani, (2003), Bootstrapping GMM Estimators for Time Series, *Journal of Econometrics*, forthcoming.

King, R.G., C.I. Plosser, and S.T. Rebelo (1988a), Production, Growth and Business Cycles 1: The Basic Neoclassical Model, *Journal of Monetary Economics,* 21, 195-232.

King, R.G., C.I. Plosser, and S.T. Rebelo (1988b), Production, Growth and Business Cycles 2: New Directions, *Journal of Monetary Economics,* 21, 309-341.

Kitamura, Y., (2002), Econometric Comparisons of Conditional Models, Working Paper, University of Pennsylvania.

Künsch H.R., (1989), The Jackknife and the Bootstrap for General Stationary Observations, *Annals of Statistics,* 17, 1217-1241.

Kydland, F.E., and E.C. Prescott, (1982), Time to Build and Aggregate Fluctuations, *Econometrica,* 50, 1345-1370.

Long, J.B. and C.I. Plosser, (1983), Real Business Cycles, *Journal of Political Economy,* 91, 39-69.

McGrattan, E.R., (1990), Solving the Stochastic Growth Model by Linear-Quadratic Approximation, *Journal of Business and Economic Statistics*, 8, 41-44.

Naik-Nimbalkar U.V. and M.B. Rajarshi, (1994), Validity of Blockwise Bootstrap for Empirical Processes with Stationary Observations, *Annals of Statistics,* 22, 980-994.

Schmitt-Grohe, S., (2000), Endogenous Business Cycles and the Dynamics of Output, Hours and Consumption, *American Economic Review,* 90, 1136-1159.

Vuong, Q. (1989), Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses, *Econometrica,* 57, 307-333.

Watson, M.W., (1993), Measure of Fit for Calibrated Models, *Journal of Political Economy,* 101, 1011-1041.

White, H., (2000), A Reality Check for Data Snooping, *Econometrica*, 68, 1097-1126.

Table 1: Comparison of Three Model Parameterizations Using Two Variables

S,l	$\sqrt{T}Z_{T,S}$	Critval ($Z^{**}_{T,S}$)		Critval ($Z^{*}_{T,S}$)		$\%<0$			Min			Max			$Stderr$		
		5%	10%	5%	10%	Sim1	Sim2	Sim3	Sim1	Sim2	Sim3	Sim1	Sim2	Sim3	Sim1	Sim2	Sim3
Part I: Technology Shocks Assumed Normally Distributed; Vary $\beta = 0.90, 0.95, 0.99$; Benchmark sets $\beta = 0.99$																	
T,l1	0.0790	0.1037	0.0844	0.0917	0.0726	51.42	55.71	40.00	-0.0836	-0.1427	-0.0964	0.0999	0.0735	0.0949	0.0410	0.0382	0.0447
T,l2	0.0804	0.0877	0.0659	0.0653	0.0472	57.14	47.14	30.00	-0.0878	-0.0816	-0.0969	0.0976	0.1185	0.1253	0.0383	0.0371	0.0455
T,l3	0.0466	0.1132	0.0773	0.0720	0.0462	51.42	50.00	41.42	-0.1178	-0.0891	-0.1030	0.1390	0.0955	0.1211	0.0594	0.0430	0.0521
2T,l1	-0.0011	0.1234	0.0951	0.0909	0.0846	41.42	49.28	40.00	-0.0862	-0.1190	-0.1286	0.1310	0.1264	0.1437	0.0401	0.0514	0.0470
2T,l2	0.0645	0.0395	0.0170	0.0063	-0.0032	58.57	42.85	38.57	-0.1327	-0.1029	-0.0773	0.0849	0.1065	0.1108	0.0414	0.0449	0.0392
2T,l3	0.0504	0.0540	0.0422	0.0161	0.0027	53.57	48.57	40.71	-0.1175	-0.1401	-0.1424	0.1025	0.1009	0.1607	0.0440	0.0432	0.0455
5T,l1	-0.0062	0.0734	0.0647	0.0594	0.0521	46.00	42.85	48.00	-0.1240	-0.1484	-0.1084	0.1149	0.1079	0.1555	0.0410	0.0430	0.0420
5T,l2	-0.0069	0.0866	0.0655	0.0585	0.0443	47.42	48.85	50.85	-0.1244	-0.1248	-0.1103	0.1555	0.1434	0.1140	0.0447	0.0441	0.0432
5T,l3	0.0060	0.0518	0.0422	0.0426	0.0312	47.71	45.71	45.14	-0.1124	-0.1126	-0.1326	0.1592	0.1159	0.1631	0.0488	0.0395	0.0451
10T,l1	0.0056	0.0529	0.0404	0.0307	0.0294	45.00	44.57	46.28	-0.1325	-0.1340	-0.1074	0.1525	0.1295	0.1617	0.0457	0.0439	0.0425
10T,l2	-0.0049	0.0560	0.0484	0.0448	0.0379	44.71	45.85	47.42	-0.1323	-0.1153	-0.1323	0.1498	0.1105	0.1428	0.0436	0.0415	0.0449
10T,l3	-0.0011	0.0685	0.0588	0.0531	0.0428	42.14	48.14	44.14	-0.1288	-0.1433	-0.1586	0.1240	0.1180	0.1450	0.0397	0.0432	0.0473
Part II: Technology Shocks Assumed to follow Student's t Distribution Under Benchmark and Normal Distributions Under Alternatives																	
T,l1	0.5030	0.2961	0.2346	0.2186	0.1902	47.14	51.42	45.71	-1.8647	-0.3304	-0.0859	1.4168	0.3734	0.0947	0.6594	0.1856	0.0438
T,l2	0.5430	0.5044	0.4019	0.3493	0.3221	54.28	58.57	44.28	-2.3253	-0.4187	-0.1190	1.3294	0.3824	0.1538	0.6961	0.1942	0.0463
T,l3	0.4639	0.2503	0.2135	0.1960	0.1455	48.57	40.00	58.57	-2.7705	-0.3456	-0.1449	1.0149	0.5267	0.1376	0.7137	0.2052	0.0609
2T,l1	0.5222	0.1726	0.1066	0.1002	0.0771	55.71	50.00	55.71	-3.3536	-0.4719	-0.1104	1.4888	0.5971	0.1018	0.7423	0.2055	0.0453
2T,l2	0.4725	0.2196	0.1866	0.2243	0.1887	46.42	47.14	44.28	-1.7797	-0.7331	-0.1101	2.0883	0.6130	0.1237	0.5650	0.1981	0.0467
2T,l3	0.5568	0.1408	0.0832	0.1016	0.0568	55.00	46.42	45.71	-2.7147	-0.3928	-0.1138	2.1707	0.4860	0.1002	0.7059	0.1928	0.0451
5T,l1	0.4742	0.1401	0.1098	0.1077	0.0887	56.28	49.71	46.85	-3.0395	-0.4408	-0.1386	2.5570	0.6628	0.1113	0.5673	0.1745	0.0462
5T,l2	0.5214	0.1114	0.1048	0.0893	0.0773	50.00	51.71	47.14	-2.6345	-0.5936	-0.1419	3.2823	0.6759	0.1608	0.7034	0.1932	0.0482
5T,l3	0.5269	0.1079	0.0817	0.0878	0.0709	50.00	50.85	49.42	-2.2623	-0.4754	-0.1153	2.0817	0.5578	0.1337	0.6647	0.1861	0.0399
10T,l1	0.5126	0.0707	0.0532	0.0632	0.0435	50.14	49.14	46.71	-2.3122	-0.6121	-0.1035	4.2879	0.5604	0.1627	0.6723	0.1919	0.0405
10T,l2	0.5200	0.1012	0.0734	0.0895	0.0680	48.85	48.28	44.85	-3.0551	-0.7013	-0.1383	2.8569	0.5721	0.1422	0.6977	0.1987	0.0452
10T,l3	0.4879	0.1394	0.1247	0.1211	0.1129	47.85	51.00	46.42	-1.9199	-0.5986	-0.1552	3.7317	0.7198	0.1554	0.6494	0.2114	0.0456

Notes: Results are based on examination of a version of the Christiano (1988) model. Simulated output and hours growth rates are compared. The first five columns of numerical entries report test statistic values (column 1) as well as 5% and 10% bootstrap critical values constructed using $\sqrt{T}Z^{*}_{T,S}$ and $\sqrt{T}Z^{**}_{T,S}$ (i.e. $T/S- > \delta > 0$ and $T/S- > 0$, respectively, where T denotes the actual data sample size and S is the length of the simulated sample). For the case where $S = T$, we set $l1 = 2, l2 = 5$ and $l3 = 10$. For all other cases, where $S = aT$, say, we set li equal to "a" times the corresponding value of li when $S = T$. All statistics are based on use of 20 uniformly distributed values for u across the range of the actual data; and bootstrap empirical distributions are constructed using 100 bootstrap replications. In addition to test statistics and critical values, various summary statistics constructed using the simulated data are given in the remaining columns of the table, for simulated output growth. Corresponding values based on the actual data (70 annual U.S. per capita output growth observations for the period 1933-2002) are: %<0 = 21.3; min = -0.128; max = 0.158; and $stderr$ = 0.044. We illustrate that the test has power against unreasonable alternatives in Part II. Under the benchmark scenario (Sim1), the errors driving the model are assumed to follow Student's t distributions with unit variance, while under the alternatives, the errors are normally distributed (with variances 0.039 and 0.077 in Sim2 - see Christiano (1988) - and with variances estimated as discussed above in Sim3). See above for further details.

Table 2: Comparison of Three Model Parameterizations Using Four Variables

S,l	$\sqrt{T}Z_{T,S}$	Critval ($Z^{**}_{T,S}$)		Critval ($Z^{*}_{T,S}$)		$\% < 0$			Min			Max			$Stderr$		
		5%	10%	5%	10%	Sim1	Sim2	Sim3	Sim1	Sim2	Sim3	Sim1	Sim2	Sim3	Sim1	Sim2	Sim3
Part III: Technology Shocks Assumed Normally Distributed; Vary $\beta = 0.90, 0.95, 0.99$; Benchmark sets $\beta = 0.99$																	
T,l1	0.0194	0.0266	0.0162	0.0145	0.0054	55.71	57.14	44.28	-0.1278	-0.1272	-0.1091	0.1044	0.1314	0.1391	0.0535	0.0499	0.0427
T,l2	0.0056	0.0379	0.0275	0.0266	0.0209	44.28	45.71	45.71	-0.1224	-0.0962	-0.1196	0.1400	0.0835	0.1490	0.0556	0.0356	0.0516
T,l3	-0.0079	0.0613	0.0525	0.0537	0.0497	54.28	47.14	58.57	-0.1143	-0.1220	-0.0874	0.1050	0.1287	0.0780	0.0477	0.0491	0.0379
2T,l1	-0.0001	0.0269	0.0233	0.0229	0.0172	43.57	40.00	42.85	-0.1454	-0.0797	-0.0864	0.1484	0.1224	0.1283	0.0491	0.0433	0.0413
2T,l2	-0.0042	0.0402	0.0332	0.0342	0.0299	44.28	47.14	55.71	-0.1079	-0.0826	-0.1304	0.1054	0.1193	0.1059	0.0373	0.0411	0.0420
2T,l3	0.0092	0.0176	0.0116	0.0152	0.0097	53.57	41.42	47.14	-0.1283	-0.1754	-0.1372	0.1056	0.1077	0.1453	0.0456	0.0392	0.0499
5T,l1	0.0176	0.0039	-0.0022	-0.0054	-0.0074	52.57	44.00	44.00	-0.1259	-0.1479	-0.1041	0.1133	0.1357	0.1068	0.0455	0.0454	0.0401
5T,l2	-0.0007	0.0185	0.0166	0.0131	0.0109	46.85	48.85	46.28	-0.1224	-0.1351	-0.1233	0.1611	0.1026	0.1394	0.0466	0.0445	0.0455
5T,l3	0.0026	0.0148	0.0118	0.0102	0.0072	43.42	43.71	40.85	-0.1380	-0.1045	-0.1355	0.1278	0.1288	0.1439	0.0451	0.0420	0.0424
10T,l1	-0.0026	0.0205	0.0182	0.0146	0.0141	42.85	49.71	43.28	-0.1418	-0.1277	-0.0991	0.1396	0.1243	0.1387	0.0477	0.0427	0.0416
10T,l2	-0.0018	0.0187	0.0163	0.0142	0.0127	43.28	46.57	46.71	-0.1211	-0.1193	-0.1285	0.1301	0.1401	0.1623	0.0408	0.0434	0.0423
10T,l3	0.0076	0.0033	0.0018	0.0022	0.0011	50.28	46.71	46.14	-0.1471	-0.1124	-0.1311	0.1352	0.1361	0.1681	0.0435	0.0469	0.0447
Part IV: Technology Shocks Assumed to follow Student's t Distribution Under Benchmark and Normal Distributions Under Alternatives																	
T,l1	0.1619	0.1033	0.0874	0.1118	0.0860	42.85	42.85	61.42	-1.7573	-0.3986	-0.1227	0.8738	0.4134	0.1049	0.6116	0.1827	0.0400
T,l2	0.1912	0.0801	0.0623	0.0652	0.0499	44.28	45.71	55.71	-2.6478	-0.4923	-0.1059	1.8229	0.3858	0.0810	0.7915	0.1808	0.0440
T,l3	0.1419	0.1707	0.1458	0.1346	0.1313	47.14	64.28	51.42	-1.5118	-0.7097	-0.1052	2.9100	0.4704	0.1088	0.7647	0.2412	0.0529
2T,l1	0.1911	0.0549	0.0487	0.0353	0.0301	47.85	56.42	42.14	-2.4122	-0.4641	-0.0969	2.0273	0.5378	0.1227	0.7063	0.2090	0.0453
2T,l2	0.1909	0.0794	0.0456	0.0398	0.0353	43.57	48.57	47.14	-3.8324	-0.5283	-0.1102	4.7029	0.6195	0.1073	0.9640	0.2262	0.0410
2T,l3	0.1973	0.0791	0.0697	0.0496	0.0370	45.71	50.00	47.14	-11.000	-0.5243	-0.1185	3.5188	0.4233	0.1240	1.2714	0.2019	0.0470
5T,l1	0.1803	0.0516	0.0420	0.0477	0.0409	46.28	51.14	47.14	-1.8517	-0.4638	-0.1806	5.0311	0.6683	0.1306	0.7669	0.2032	0.0467
5T,l2	0.1799	0.0513	0.0410	0.0481	0.0364	51.14	48.85	41.14	-2.1991	-0.6537	-0.1738	2.6569	0.6211	0.1394	0.6158	0.2018	0.0429
5T,l3	0.1724	0.0566	0.0420	0.0514	0.0397	49.71	47.42	41.42	-2.5257	-0.4934	-0.1442	3.9083	0.6733	0.1546	0.7452	0.1851	0.0421
10T,l1	0.1793	0.0348	0.0322	0.0354	0.0313	48.00	50.42	46.57	-6.8205	-0.7557	-0.1367	3.5213	0.4993	0.1531	0.7939	0.1928	0.0455
10T,l2	0.1662	0.0609	0.0456	0.0532	0.0427	51.42	52.14	47.42	-2.8809	-0.6292	-0.1182	2.9678	0.6587	0.1463	0.6471	0.2016	0.0408
10T,l3	0.1777	0.0550	0.0411	0.0500	0.0421	48.28	47.14	47.28	-2.8273	-0.6831	-0.1352	9.0227	0.7221	0.1662	0.8134	0.2070	0.0444

Notes: See notes to Table 1. Simulated data being compared are output and hours growth rates, as well as the first lags of each of these variables. See above for further details.