

Bangs, Joann; Landon-Lane, John

Working Paper

International real business cycles and increasing returns to scale: A formal analysis using likelihood methods

Working Paper, No. 2002-12

Provided in Cooperation with:

Department of Economics, Rutgers University

Suggested Citation: Bangs, Joann; Landon-Lane, John (2002) : International real business cycles and increasing returns to scale: A formal analysis using likelihood methods, Working Paper, No. 2002-12, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:

<https://hdl.handle.net/10419/79172>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

International Real Business Cycles and Increasing
Returns to Scale: A Formal Analysis using Likelihood
Methods

Joann Bangs and John Landon-Lane*
Oakland University and Rutgers University

June 18, 2002

*Address for Correspondence: John Landon-Lane, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA. E-Mail: lane@econ.rutgers.edu. We would like to thank Jang-Ting Guo, who kindly provided us with his Gauss code for linearizing and solving the standard DSGE model used in this paper.

Abstract

One of the more common methods used to model international real business cycles is through the use of a dynamic stochastic general equilibrium (DSGE) model. Guo and Sturzenegger (1998) argue that an increasing returns to scale production technology can improve the performance of such a model. They also argue that if increasing returns are strong enough, then sunspot equilibria are possible. In this paper, we formally test the increasing returns to scale assumption and find that a model with a constant returns to scale technology has a superior out-of-sample prediction performance over a model with an increasing returns to scale production technology. Moreover, this result is robust to the degree of returns to scale and to the persistence and variance of the shocks in the model.

J.E.L. Classification E32, C11, C52

Keywords: International Real Business Cycles, Increasing Returns to Scale, Model Comparison

1 Introduction

One of the more common methods used to model international real business cycles is through the use of a dynamic stochastic general equilibrium (DSGE) model¹. Following from the work of Kydland and Prescott (1982), the parameters of a DGSE model are calibrated using values that reflect the modeler's prior beliefs. Once calibrated, the model is evaluated using simulation experiments (Kydland and Prescott 1996). A common practice in the literature is to evaluate a model by how well the artificial data generated from the calibrated model can match certain stylized facts. In the case of international real business cycle (IRBC) models, one of the most important stylized facts is that output is more highly correlated across countries than consumption. In fact, the standard IRBC model with constant returns to scale, perfect competition and complete markets has exactly the opposite result. This has been called the consumption-output anomaly. There have been many attempts to modify the standard IRBC model to eliminate the consumption-output anomaly. Models that incorporate incomplete markets, increasing returns, household production, or a combination of these features have been used to explain the consumption-output anomaly.

Previous studies have focused on the ability of an alternative model to adequately match the consumption-output stylized fact to determine if one model is "better" than a competing model. This method of model evaluation and comparison, while common, is not without criticism². In particular, determining a model's performance by its ability to mimic a small number of features in the data has been criticized for not comparing, or evaluating, models over the full dimension of the data. Likelihood methods do not suffer from this criticism. Therefore, the object of this paper is to use likelihood methods to formally compare a number of competing IRBC models across the full dimension of the

¹See Backus, Kehoe and Kydland (1992), Canova and Ubide (1998), and Devereux, Gregory and Smith (1991) for example.

²See Canova and Ortega (1995) for a good discussion of the methods used in evaluating DSGE models.

data.

Nearly all DSGE models cannot be simulated as they are written. This is the certainly the case for the models compared in this paper. It is necessary for certain aspects of the DSGE model to be approximated in order for it to be simulated. The approximation also allows the model to be written in state-space form. This (approximate) state-space form of the model is used to construct a likelihood function for the model in question. Principle components are used to solve the dimensionality problem and construct the likelihood function.³ Markov chain Monte Carlo (MCMC) methods are then used to formally compare two or more non-nested DSGE models. One of the benefits of using MCMC methods to compare models is that the models can then be formally compared across sub-samples of the data as well as across the whole data set. Also, it is possible to incorporate the practice of calibration into the framework through the description of appropriate prior distributions for the structural parameters of the model. It is then possible to not only compare different models, but also to formally compare a particular calibration experiment, such as matching specific moments of the data, with moments generated using a more standard likelihood estimation technique.

This paper uses likelihood methods to compare an IRBC model that has increasing returns to scale to a benchmark IRBC model with constant returns to scale production technology. The comparison is performed for the calibrated case, that is, where the parameters are chosen to be specific values, and in several cases where the prior distribution is relaxed on the parameters of, first, the equation governing the motion of the stochastic terms of the model, and second, the parameter determining the level of increasing returns. We have two major findings. First, the increasing returns assumption is overwhelmingly rejected in favor of a model with constant returns to scale, in the sense that the increasing returns to scale model has a worse out-of-sample prediction performance

³The dimensionality problem arises from the property that the number of stochastic components present in a model is less than the model's dimension. See Kim and Pagan (1995) for a more thorough discussion of the dimensionality issue.

than the constant returns to scale model. Second, allowing for prior uncertainty greatly improves the prediction performance of the models. The rest of the paper is as follows: Section 2 discusses the models that will be compared, Section 3 describes the method used in this paper, Section 4 describes the data and the various prior specifications used, Section 5 describes the results, and Section 6 concludes.

2 Increasing returns and International Real Business Cycles

Since the work of Baxter and King (1991), a number of authors have incorporated increasing returns to scale production technologies in DSGE models in an attempt to better explain the observed data. Examples of this can be found in Farmer (1993), Farmer and Guo (1994), and, in the IRBC context, Guo and Sturzenegger (1998). In these papers, the authors conclude that a model with increasing returns to scale technologies can better match selected moments of the observed data. These authors also claim that if increasing returns are strong enough, sunspot equilibria, where shocks are not real, are possible. The aim of this paper is to examine whether there is evidence for increasing returns to scale and if so whether the increasing returns to scale are large enough to permit sunspot equilibria.

The models we describe in this section are similar to those used in Guo and Sturzenegger (1998). Their basic model uses a Dixit-Stiglitz-type production. In each country, the final consumption good, Y , is produced from a continuum of intermediate inputs, $Y(i)$ with $i \in [0, 1]$. The production function for the perfectly competitive final goods sector is

$$Y = \left(\int_0^1 Y(i)^\lambda di \right)^{\frac{1}{\lambda}}. \quad (1)$$

The intermediate goods are all produced by the same technology represented by the production function

$$Y(i) = ZK(i)^\alpha L(i)^\beta, \quad (2)$$

so that aggregate output is

$$Y_t = Z_t K_t^\alpha L_t^\beta, \quad (3)$$

where the parameter Z_t is an aggregate technology shock with unit mean. Its law of motion is governed by

$$Z_{t+1} = \Omega Z_t + \eta_{t+1}, \quad (4)$$

where

$$\eta_t \sim N(\mathbf{I} - \mathbf{\Omega}, \mathbf{V}_\eta).$$

Preferences are represented by

$$U = \sum_{t=0}^{\infty} \rho^t \sum_{z_t \in H_t} \pi(z_t) \left\{ \frac{C_t^{1-\theta}(z_t)}{1-\theta} - \frac{L_t^{1-\gamma}(z_t)}{1-\gamma} \right\}, \quad (5)$$

where z_t is a particular realization of the state of nature, $\pi(z_t)$ is the probability of that state occurring, and all possible histories for the realization are given by H_t . Assuming complete markets, the representative household's budget constraint is

$$\sum_{t=0}^{\infty} \sum_{z_t \in H_t} q_t \{ C_t + K_{t+1} - (1 - \delta)K_t + p_t^1 \Theta_{t+1}^1 + p_t^2 \Theta_{t+1}^2 - r_t K_t - w_t L_t - (p_t^1 - d_t^1) \Theta_t^1 - (p_t^2 - d_t^2) \Theta_t^2 \} = 0. \quad (6)$$

The variable Θ_t^i represents the shares of equity holdings to a firm in country i at time t . Similarly, d_t^i represents dividends and p_t^i represents the share price.

Typically in IRBC models, only investment is mobile across countries, however, Guo and Sturzenegger (1998) also allow capital to be mobile. Hence the world resource constraint is

$$K_{t+1} = Y_{1t} + Y_{2t} + (1 - \delta)K_t - C_{1t} - C_{2t} \quad (7)$$

where $K_t = K_{1t} + K_{2t}$ is the world capital stock.

Guo and Sturzenegger (1998) show how this model can be reduced to a system of non-linear dynamic equations. Using the methods described in Farmer (1993), this set of non-linear equations can be linearized into the following system:

$$\begin{bmatrix} \hat{K}_t \\ \hat{C}_t \\ \hat{Z}_{1t} \\ \hat{Z}_{2t} \end{bmatrix} = \mathbf{J} \begin{bmatrix} \hat{K}_{t+1} \\ \hat{C}_{t+1} \\ \hat{Z}_{1t+1} \\ \hat{Z}_{2t+1} \end{bmatrix} + \mathbf{R} \begin{bmatrix} E_t [\hat{K}_{t+1}] - \hat{K}_{t+1} \\ E_t [\hat{C}_{t+1}] - \hat{C}_{t+1} \\ \hat{\eta}_{1t+1} \\ \hat{\eta}_{2t+1} \end{bmatrix}, \quad (8)$$

where Z_0 and K_0 are given and \hat{X}_t is the proportional deviation of variable X from its steady state. This system, (8), forms the basis of the models that we will compare in this paper. Note that this model assumes a symmetric equilibrium, $C_{1t} = C_{2t}$, and complete markets.

For most of the structural parameters, the restrictions placed on them are the same in

both the constant returns to scale and the increasing returns to scale versions of the model. The parameter γ is the negative inverse of the labor supply elasticity. Values are restricted to $\gamma \leq 0$. The parameter ρ is the discount factor and is restricted to $0 < \rho < 1$. The rate of depreciation of the capital stock must have, $0 < \delta < 1$. Finally, θ represents the intertemporal elasticity of substitution for consumption. This elasticity is restricted to be positive and $\theta \neq 1$.

The values of the remaining structural parameters determine whether the model has constant or increasing returns to scale. The parameters α and β represent the capital and labor elasticities in production. The parameter $\lambda \in (0, 1]$ represents the degree of monopoly power held by the monopolistically competitive intermediate goods producers. The factor shares in total output can then be defined as $a = \lambda\alpha$ and $b = \lambda\beta$ for capital and labor respectively. Note that if $\lambda = 1$, the various intermediate inputs are all perfect substitutes. In that case, the intermediate goods producers do not have any monopoly power, and behave as perfectly competitive firms. The constant returns to scale version of the model uses $\lambda = 1$. The factor shares are then simply $a = \alpha$ and $b = \beta$. Then with $\alpha + \beta = 1$, the model has constant returns to scale. To incorporate increasing returns to scale into the model requires $\alpha + \beta > 1$ and $\lambda < 1$.

3 Model Comparison and Construction of Likelihood Function

The model comparison methods used in this paper are based on the methods described in Landon-Lane((1998) and (2002)). In particular, the linearized state-space approximation to the model described in Section 2 is used to construct a likelihood function which is then used to compare the two versions of the model.

The linear system given in (8) is used to construct the error equation of the state-space system. As long as there is only one eigenvalue of \mathbf{J} that has magnitude less than one, (8) can be iterated into the future to get \hat{C}_t as a linear function of the state vector, $\hat{s}_t = (\hat{K}_t, \hat{Z}_{1t}, \hat{Z}_{2t})'$. Thus (8) can be rewritten as

$$\begin{bmatrix} \hat{K}_{t+1} \\ \hat{Z}_{1t+1} \\ \hat{Z}_{2t+1} \end{bmatrix} = \mathbf{J}_1 \begin{bmatrix} \hat{K}_t \\ \hat{Z}_{1t} \\ \hat{Z}_{2t} \end{bmatrix} + \begin{bmatrix} 0 \\ \hat{\eta}_{1t+1} \\ \hat{\eta}_{2t+1} \end{bmatrix}; \quad (9)$$

the error equation of the state-space approximation.

The data equation can be obtained by linearizing the equations that relate the observed variables of the model to the state variables. The observed variables of the model are output, hours worked, and savings, all for both countries, and total consumption and total investment across both countries.⁴ The equation that governs output is described in (3) while total labor hours used is described by

$$L_{it} = \left[\frac{1}{b} \frac{C_{it}^\theta}{Z_{it} K_{it}^\alpha} \right]^\omega. \quad (10)$$

Total investment is given by

$$I_t = K_{t+1} - \delta K_t, \quad (11)$$

and savings is defined as

$$SAV_{it} = Y_{it} - C_{it}. \quad (12)$$

Linearizing, where necessary, and substituting out all non-state variables it is possible to write the data equation as

$$\hat{D}_t = \mathbf{H}\hat{s}_t \quad (13)$$

⁴We use total consumption, $C_{1t} + C_{2t}$, and total investment, $I_{1t} + I_{2t}$, as we assume a symmetric equilibrium for consumption, $C_{1t} = C_{2t}$, and we assume that investment is mobile between the two countries.

where $\hat{D}_t = (\hat{Y}_{1t}, \hat{Y}_{2t}, \hat{L}_{1t}, \hat{L}_{2t}, S\hat{A}V_{1t}, S\hat{A}V_{2t}, \hat{C}_{1t} + \hat{C}_{2t}, \hat{I}_{1t} + \hat{I}_{2t})'$, and \mathbf{H} is a (8×3) matrix of coefficients. Thus, equations (9) and (13) form the state-space approximation to the model described in Section 2.

Here we see the dimensionality problem in stark contrast. There are eight observed variables but only two random components of the model. A number of papers, such as Sargent (1989), Anderson, Hansen, McGratten and Sargent (1996), and Ireland (1999), have attempted to solve this problem by assuming that the variables of the model are measured with error thus increasing the number of stochastic terms in the model. Other papers, such as DeJong, Ingram and Whiteman (2000) and Landon-Lane (1998) construct likelihoods based on a subset of the data. Both methods suffer from a non-uniqueness problem. In the model used in this paper, six additional shocks need to be added to make the model of full dimension. There are twenty-eight ways in which this can happen. The same issue applies to the problem of only using two variables to calculate the likelihood function. There are also twenty-eight different combinations of two variables that can be used.

In order to solve the uniqueness problem, independent linear combinations of the eight observed variables are used to construct the likelihood function. Two independent linear combinations of the observed variables need to be found. The obvious approach is to use the two principle components that account for the highest proportion of the observed variance in the data.⁵ This allows us to uniquely determine the values of the shocks, and hence calculate the likelihood function.

The approach is as follows: Let $\hat{\mathbf{D}}^T = \left\{ \hat{D}_t \right\}_{t=1}^T$ be the T (8×1) dimensioned observed data vectors and let $\psi = (\psi_s, \psi_z)$ be the vector of parameters of the model⁶. Next, let \mathbf{P} be the (8×2) dimensioned matrix containing the two eigenvectors, normalized to have

⁵See Johnson and Wichern (1988) for more details on the method of principle components.

⁶The structural parameters of the model are represented by ψ_s , and ψ_z refers to the parameters that characterize the law of motion for the shock vector Z_t .

length one, associated with the two highest eigenvalues of the correlation matrix⁷ of the observed data, $\hat{\mathbf{D}}^T$. Thus the data equation, (13), becomes

$$\hat{X}_t = \mathbf{P}'\hat{D}_t = \mathbf{P}'\mathbf{H}\hat{s}_t. \quad (14)$$

We then construct the likelihood function using the modified data, $\hat{\mathbf{X}}^T = \{\hat{X}_t\}_{t=1}^T$.

The construction of the likelihood function uses equations (9) and (14), the modified state-space approximation to the model described in Section 2. The basic idea is that if the value of \hat{K}_t is known then (14) can be solved uniquely for $\hat{Z}_t = (\hat{Z}_{1t}, \hat{Z}_{2t})'$. Once \hat{Z}_t is known then, from (9), using the fact that \hat{K}_{t+1} is a deterministic function of the current state, \hat{s}_t , it is possible to uniquely determine the value of \hat{K}_{t+1} . Thus, given the value of \hat{K}_0 , it is possible to uniquely solve for $\mathbf{Z}^T = \{Z_t\}_{t=1}^T$. Note that \mathbf{Z}^T is a function of the observed data, $\hat{\mathbf{X}}^T$, the unknown parameters of the model, ψ , and the initial value of \hat{K} , \hat{K}_1 . That is, $\hat{\mathbf{Z}}^T \equiv \mathbf{Z}^T(\hat{\mathbf{X}}^T, \hat{K}_1, \psi)$. For the sake of brevity, however, we will use the notation \mathbf{Z}^T rather than $\mathbf{Z}^T(\hat{\mathbf{X}}^T, \hat{K}_1, \psi)$ in what follows.

One consequence of the method used to construct the likelihood function is that the initial value of the deviation of capital from its steady-state (or balanced growth path) value is made a parameter of the model. Thus the parameter vector of the model becomes $\tilde{\psi} = (\psi, \hat{K}_1)$. For $t = 2, \dots, T$ the conditional density function of $\hat{\mathbf{Z}}_t$, from (4) is

$$p(\hat{\mathbf{Z}}_t | \hat{\mathbf{Z}}_{t-1}, \hat{\mathbf{X}}^t, \tilde{\psi}) = (2\pi)^{-1} |\mathbf{V}_\eta|^{-1} \exp \left\{ (\hat{\mathbf{Z}}_t - \Omega \hat{\mathbf{Z}}_{t-1})' \mathbf{V}_\eta^{-1} (\hat{\mathbf{Z}}_t - \Omega \hat{\mathbf{Z}}_{t-1}) \right\}. \quad (15)$$

For $t = 1$ we use the unconditional density

$$p(\hat{\mathbf{Z}}_1 | \hat{\mathbf{X}}^1, \tilde{\psi}) = (2\pi)^{-1} |\mathbf{V}_{Z_1}|^{-1} \exp \left\{ (\hat{\mathbf{Z}}_1)' \mathbf{V}_{Z_1}^{-1} (\hat{\mathbf{Z}}_1) \right\}, \quad (16)$$

⁷Note that the correlation matrix is used here instead of the covariance matrix so that the linear combinations are not weighted towards the variable with the highest variance.

where \mathbf{V}_{Z_1} solves

$$\mathbf{V}_{Z_1} - \Omega \mathbf{V}_{Z_1} \Omega' = \mathbf{V}_\eta.$$

Therefore the likelihood function for model \mathcal{M} is

$$p(\hat{\mathbf{X}}^T | \tilde{\psi}, \mathcal{M}) = p(\hat{\mathbf{Z}}_1 | \hat{\mathbf{X}}^1, \tilde{\psi}, \mathcal{M}) \prod_{t=2}^T p(\hat{\mathbf{Z}}_t | \hat{\mathbf{Z}}_{t-1}, \hat{\mathbf{X}}^t, \tilde{\psi}, \mathcal{M}) |\mathbf{Jac}|, \quad (17)$$

where

$$\mathbf{Jac} = \left(\frac{\partial(\hat{X}_1, \dots, \hat{X}_T)}{\partial(\hat{Z}_1, \dots, \hat{Z}_T)'} \right)^{-1}$$

is the Jacobian of the transformation from $(\hat{Z}_1, \dots, \hat{Z}_T)$ to $(\hat{X}_1, \dots, \hat{X}_T)$. From (14) it follows that

$$\hat{X}_t = \mathbf{P}' \mathbf{H}_{.,1} \hat{K}_t + \mathbf{P}' \mathbf{H}_{.,23} \mathbf{Z}_t, \quad (18)$$

where $\mathbf{H}_{.,i}$ refers to the i^{th} column of the matrix \mathbf{H} , so that

$$\mathbf{Jac} = (\mathbf{P}' \mathbf{H}_{.,23})^{-1}. \quad (19)$$

3.1 Estimation and Model Comparison

For a complete discussion of estimation and model comparison the context of general DSGE models see Landon-Lane (2002). The methods derived in that paper are used extensively here. The basic approach is to use Bayesian model comparison methods to formally compare competing models⁸, as these methods allow for the formal treatment of model uncertainty and prior beliefs. As discussed in Sims (1996), another reason for using Bayesian methods is that there is a fair degree of prior information used in the DSGE literature. Models are calibrated to certain values based on the prior beliefs of the authors. Bayesian methods are the natural way to formally incorporate prior information

⁸See Geweke (1995) for a good discussion on Bayesian model comparison.

with information from the observed data in the estimation of the models.

Given the likelihood function, sometimes referred as the data density, $p(\hat{\mathbf{X}}^T|\tilde{\psi}, \mathcal{M})$ and the prior density, $p(\tilde{\psi}|\mathcal{M})$, it follows from Bayes Theorem that the posterior density, $p(\tilde{\psi}|\hat{\mathbf{X}}^T, \mathcal{M})$ is defined as

$$p(\tilde{\psi}|\hat{\mathbf{X}}^T, \mathcal{M}) \propto p(\tilde{\psi}|\mathcal{M})p(\hat{\mathbf{X}}^T|\tilde{\psi}, \mathcal{M}). \quad (20)$$

The marginal likelihood, $p(\hat{\mathbf{X}}^T|\mathcal{M})$, is then defined as

$$p(\hat{\mathbf{X}}^T|\mathcal{M}) = \int_{\Psi} p(\tilde{\psi}|\mathcal{M})p(\hat{\mathbf{X}}^T|\tilde{\psi}, \mathcal{M})d\tilde{\psi}, \quad (21)$$

which is interpreted as the probability of observing $\hat{\mathbf{X}}^T$ under model \mathcal{M} . Geweke (1995) shows how the marginal likelihood can be interpreted as representing the cumulative out-of-sample prediction performance of the model so that the Bayes factor, defined as

$$\mathbf{BF}_{ij} = \frac{p(\hat{\mathbf{X}}^T|\mathcal{M}_i)}{p(\hat{\mathbf{X}}^T|\mathcal{M}_j)}, \quad (22)$$

can be interpreted as the cumulative out-of-sample prediction performance of model \mathcal{M}_i relative to that of model \mathcal{M}_j .

For any well defined function, $g(\cdot)$, Geweke (1999) shows that, under certain regularity conditions,

$$M^{-1} \sum_{m=1}^M g(\tilde{\psi}^{(m)}) \longrightarrow E(g(\tilde{\psi})|\hat{\mathbf{X}}^T, \mathcal{M}_i),$$

where $\{\tilde{\psi}^{(1)}, \dots, \tilde{\psi}^{(M)}\}$ is a series of random draws, using a MCMC algorithm, from $p(\tilde{\psi}|\hat{\mathbf{X}}^T, \mathcal{M})$. Geweke (1999) also shows how the method of Gelfand and Dey (1994) can be modified so that $\{\tilde{\psi}^{(1)}, \dots, \tilde{\psi}^{(M)}\}$ can be used to calculate the marginal likelihood,

$p(\hat{\mathbf{X}}^T|\mathcal{M}_i)$.

Therefore, in order to perform inference on the parameters, $\tilde{\psi}$, and also to calculate $p(\hat{\mathbf{X}}^T|\mathcal{M}_i)$, thus allowing a formal comparison the models presented in Section 2, a method of drawing randomly from $p(\tilde{\psi}|\hat{\mathbf{X}}^T, \mathcal{M})$ is needed. One such algorithm is the random walk Hastings-Metropolis MCMC algorithm described in Chib and Greenberg (1995) and Tierney (1994). This algorithm is used to obtain the results that follow.

4 Data and Prior Specifications

Before describing the results we first describe the data and prior distributions used in this exercise. The data we use are from the OECD⁹. They provide quarterly data from the United States and an aggregate for the fifteen European Union countries. Data on consumption, output, total hours worked, investment and savings are used to construct the likelihood function. Data for the European Union are in terms of 1995 US dollars. In order to have a comparable data set, a GDP deflator was used to convert the data in current dollars from the United States into 1995 US dollars. The specific series taken from the OECD data (codes are listed in parentheses) are as follows. For consumption, we used the private final consumption expenditure series (USA1001S2, 021023HSA). The gross fixed capital formation series (USA1003S2, 0210035HSA) was used for investment. For output, we used the gross domestic product by expenditure series (USA1008S2, 0210037HSA). The OECD do not provide data on total hours worked for the European aggregate. Therefore, German data on total hours worked is used as a proxy for European total hours worked.

In order to use the methods described in Section 3.1, the data are converted to proportional deviations from their balanced growth path. For a variable, X_t , the Hodrick and

⁹<http://www.sourceoecd.org>

Prescott (1997) filter is used to calculate the underlying trend, X_t^* , and the deviation from trend, $X_t - X_t^*$. Then, \hat{X}_t is defined as

$$\hat{X}_t = \frac{X_t - X_t^*}{X_t^*}. \quad (23)$$

4.1 Prior Specifications

The parameter vector, $\tilde{\psi}$, is made up of three components; a vector of the structural parameters of the model, ψ_s , a vector of the parameters that describe the law of motion of the shocks that hit the model, ψ_z , and the initial value of \hat{K} , \hat{K}_1 .

The structural parameter vector, ψ_s is made up of the parameters of the model that do not affect the law of motion of the shock process. That is, $\psi_s = (\gamma, \rho, \delta, \theta, \lambda, a, b)^{10}$. The parameters that determine the law of motion of the shock vector, \hat{Z}_t , are the (2×2) matrices $\mathbf{\Omega}$ and \mathbf{V}_η . It is assumed that the shock vector follows a vector autoregressive process of order one where $\mathbf{\Omega}$ is the coefficient on the first lag of \hat{Z}_t , the eigenvalues of which determine the persistence properties of the shock process, and \mathbf{V}_η is the symmetric positive definite variance covariance matrix of η , the independent and identically distributed innovations to the shock process. Therefore, $\psi_z = (\sigma_{\eta_1}^2, \sigma_{\eta_2}^2, \rho_{(\eta_1, \eta_2)}, \text{vec}(\mathbf{\Omega})')'$, where $\rho_{(\eta_1, \eta_2)}$ is the correlation between η_1 and η_2 and $\text{vec}(\mathbf{\Omega})$ is the column-wise vectorization of $\mathbf{\Omega}$.

The comparison exercise involves the comparison of two variants of the model described in Section 2. One version of the model is for the case where the production technology exhibits constant returns to scale (CRS version) and the other version is where the aggregate production technology exhibits increasing returns to scale (IRS version). There are three separate comparisons made. The first comparison involves comparing the CRS model with the IRS model at the calibrated values of the parameters used in Guo and Sturzenegger (1998). The second comparison compares the two versions allowing for the

¹⁰The parameters α and β are determined by $\alpha = a/\lambda$ and $\beta = b/\lambda$.

parameters of the shock process to vary, while the third comparison involves allowing the degree of increasing returns to scale to vary with ψ_z . That is, in the first two comparisons, the values of ψ_s are fixed at their calibrated values, while in the third comparison, only λ is allowed to vary.

The calibrated values used in Guo and Sturzenegger (1998) are in Tables 1 and 2 below.

Table 1: Calibrated Values for ψ_s

Parameter	CRS	IRS
γ	-0.25	-0.25
ρ	0.99	0.99
δ	0.025	0.025
θ	2	2
λ	1	0.77
a	0.30	0.24
b	0.70	0.70

Table 2: Calibrated Values for ψ_z

Parameters	CRS	IRS
$\text{var}(\eta_1)$	0.00862 ²	0.00923 ²
$\text{var}(\eta_2)$	0.00862 ²	0.00923 ²
$\text{corr}(\eta_1, \eta_2)$	0.396	0.289
Ω	$\begin{bmatrix} 0.940 & 0.058 \\ 0.058 & 0.940 \end{bmatrix}$	$\begin{bmatrix} 0.947 & 0.047 \\ 0.047 & 0.947 \end{bmatrix}$

In all cases, the prior means for the parameters are set to the calibrated values given above. The choice of prior variances is more troublesome. Typically, one would choose prior variances to reflect any disagreement in the literature regarding the values of parameters. However, one of the striking aspects of the DSGE literature is the uniformity in the values of the parameters used. Therefore it is hard to discern any prior uncertainty

over the values of the parameters that make up the model. Therefore, prior variances are used to represent small uncertainty in the calibration. Ninety-five percent prior coverage intervals for the parameters are reported in Table 3 below. Note that prior coverage intervals that are degenerate imply a point-mass prior for that parameter. For brevity, degenerate coverage intervals are not reported.

Table 3: 95% Highest Prior Density Coverage Intervals

Parameter	Comparison		
	I	II	III
γ	-	-	-
ρ	-	-	-
δ	-	-	-
θ	-	-	-
λ	-	-	[0.6, 0.88]
a	-	-	-
b	-	-	-
$\text{var}(\eta_1)$	-	[0.00780 ² , 0.00952 ²]	[0.00780 ² , 0.00952 ²]
$\text{var}(\eta_2)$	-	[0.00835 ² , 0.0102 ²]	[0.00835 ² , 0.0102 ²]
$\text{corr}(\eta_1, \eta_2)$	-	[-0.141, 0.649]	[-0.172, 0.639]
$\mathbf{\Omega}$	-	$\begin{bmatrix} [0.85, 0.98] & [0.01, 0.12] \\ [0.01, 0.12] & [0.85, 0.98] \end{bmatrix}$	$\begin{bmatrix} [0.86, 0.98] & [0.01, 0.11] \\ [0.01, 0.11] & [0.86, 0.98] \end{bmatrix}$
\hat{K}_1	[-0.70, 0.70]	[-0.70, 0.70]	[-0.70, 0.70]

Notes:

1. A dash, -, represents a degenerate prior distribution. That is, the ninety-five percent prior coverage interval is just the prior mean.
2. The prior coverage intervals for $\mathbf{\Omega}$ reported above are the unrestricted coverage intervals. The actual prior for $\mathbf{\Omega}$ is restricted to those $\mathbf{\Omega}$ that have eigenvalues whose magnitude are strictly less than one.

In order to apply the random walk Hastings-Metropolis algorithm it was convenient to work with transformed parameters. For example, the value of $\sigma_{\eta_1}^2$ is constrained to be greater than zero. Therefore, the model was re-parameterized to include $\log(\sigma_{\eta_1}^2)$ as a parameter. The correlation between η_1 and η_2 is constrained to be in the interval [-1,1].

Therefore, $\rho_{(\eta_1, \eta_2)}$ was re-parameterized using the modified logistic transformation

$$\tilde{\rho}_{(\eta_1, \eta_2)} = \frac{\rho_{(\eta_1, \eta_2)} + 1}{\rho_{(\eta_1, \eta_2)} - 1}. \quad (24)$$

All transformations that were applied were done so as to transform the parameter space to be the real line. Given these transformations, it was convenient to define Gaussian priors for the transformed variables. The prior coverage intervals reported in Table 3 are obtained by inverting the transformation and calculating a highest prior density region for the original parameter.

Finally, the prior for $\mathbf{\Omega}$ is defined similarly except that the prior is truncated to the region of the space for which the eigenvalues of $\mathbf{\Omega}$ are all less than one in magnitude. The random-walk step for $\mathbf{\Omega}$ is therefore modified using an accept/reject step.

5 Results

As described in the previous section, three separate comparisons are carried out. The first comparison compares the calibrated constant returns to scale(CRS) version of the IRBC model described in Section 2 with the calibrated increasing returns to scale(IRS) version. In this case only \hat{K}_1 is free to vary.

The second comparison involves allowing the parameters that determine the law of motion of \mathbf{Z} , ψ_z , to vary. The reason for this comparison is that in the calibration experiment in Guo and Sturzenegger (1998), the values for ψ_z are chosen so that the model mimics the underlying properties of observed output only. This is done by calculating Solow residuals using the aggregate production function given in (3). In this comparison, the values of the structural parameters are still fixed at their calibrated values. The reason for allowing the parameters that make up ψ_z to vary is to see whether the result in the

first comparison is robust to how the shock process is defined. Note that in the calibration experiment, except for λ , the structural parameters are identical. However, the values of ψ_z differ. Allowing the values of ψ_z to be chosen using likelihood methods will eliminate any suspicion that the result in the first comparison is due to the choice of ψ_z rather than the validity of the models.

Finally, the third comparison involves allowing, along with ψ_z , λ to be chosen freely in the IRS version of the model. Thus, we will be able to check the robustness of the result in the first comparison to the choice of λ , the parameter that determines the degree of increasing returns to scale. Therefore the three comparisons are as follows:

Table 4: Summary of Comparisons

Comparison	Free Parameters
I:	full calibration - only \hat{K}_1 free to vary
II:	(ψ_z, \hat{K}_1) allowed to vary for both models
III:	(ψ_z, \hat{K}_1) allowed to vary for the CRS model, $(\psi_z, \hat{K}_1, \lambda)$ allowed to vary for the IRS model

In all comparisons, the approach is as follows: A random-walk Hastings-Metropolis algorithm is used to make serially dependent draws from the posterior distribution of the free parameters of both models. These draws are then used to calculate the log-marginal likelihood using the modified method of Gelfand and Dey (1994) as suggested by Geweke (1999). The log-marginal likelihoods are then used to calculate the log-Bayes factor which is used to compare the two models. In all cases, the random-walk Hastings-Metropolis algorithm is tuned so that it efficiently draws from the posterior of choice. In all cases the numerical standard error is less than ten percent of the posterior standard error for each parameter estimated.

The posterior means of the free parameters for each comparison can be found in Table 7 at the end of this section and the log Bayes factors can be found in Table 5 below.

Table 5: Log Bayes Factor in favor of CRS Model over IRS Model

Comparison	I	II	III
log BF	228.6	102.6	99.6
	(0.09)	(0.24)	(0.64)

Note:

The numbers in the brackets reflect the numerical standard error of the log-Bayes factor estimates.

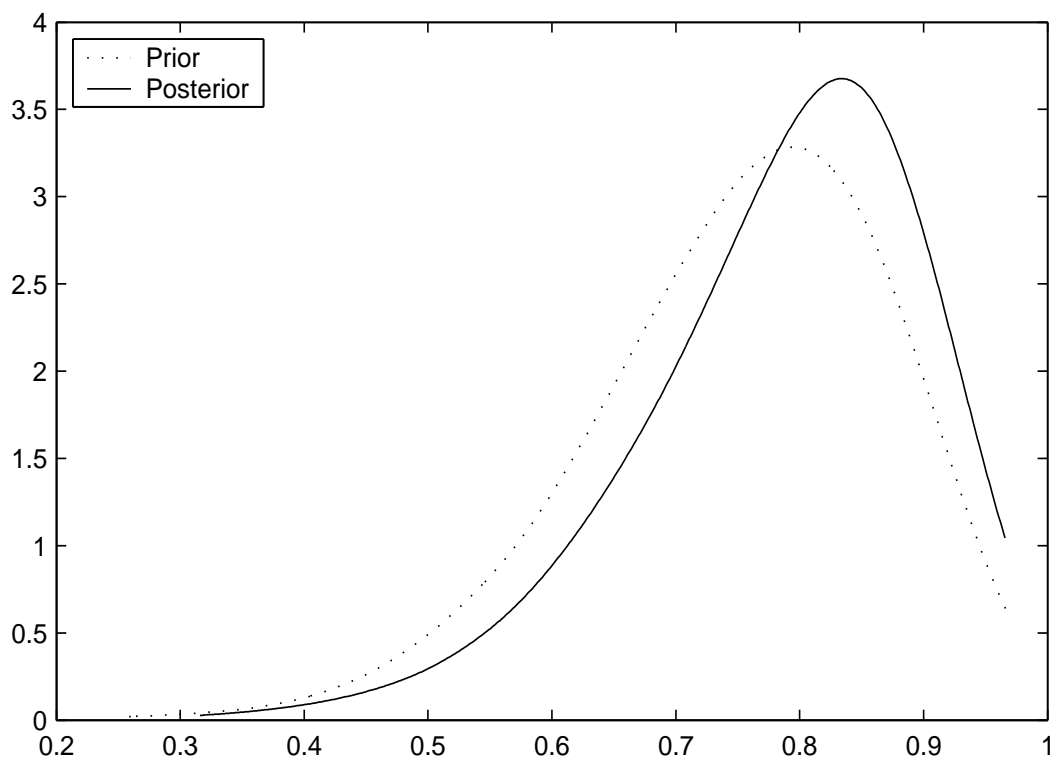
It is clear from the results that the CRS model is strongly favored in comparison to the IRS model and this result is robust to the choice of ψ_z and λ . The Bayes factor in favor of the CRS model over the IRS model is $\exp(228.6)$, $\exp(102.6)$, and $\exp(99.6)$ for comparisons I, II and III respectively. That is, in the case of comparison III, for example, the CRS model is $\exp(99.6)$ times more likely than the IRS model. Or, another way of stating this result is that the out-of-sample prediction performance of the CRS model is $\exp(99.6)$ times better than the IRS model.

When ψ_z is estimated we find that the estimated value of Ω is close to its calibrated value but the variance-covariance matrix, \mathbf{V}_η is very different. The variance of the shocks are estimated to be smaller than their calibrated values, by a factor of about 70%, and the correlation between the two shocks is estimated to be very close to one. In fact, we constrain the value of $\rho_{Z_1, Z_2} \in (-1, 1)$ so we take the fact that the estimate is very close to one as evidence that the correlation between the two shocks is one.

Moreover, inspection of the posterior mean for the parameter λ in comparison III shows that the data contains some information about its value. Evidence of this can be seen in the plot of the prior and posterior distributions for λ , given in Figure 1 which clearly shows that the posterior distribution of λ is shifted to the right and more concentrated than the prior distribution. Thus, the estimated value of λ is larger than the value calibrated by Guo and Sturzenegger and so the estimated value of the degree of increasing returns to scale¹¹ is smaller than that calibrated. Finally, Guo and Sturzenegger (1998) also show

¹¹The degree of returns to scale is calibrated at $0.24/0.77 + 0.7/0.77 = 1.224$ while the estimated

Figure 1: Prior and Posterior Distribution for λ



that if increasing returns to scale are large enough¹² then sunspot equilibria are possible. The evidence here, however, suggests that the increasing returns to scale, if present at all, are not large enough to allow for sunspot equilibria.

Therefore, we find that there is very little evidence in favor of using an increasing returns to scale technology in an IRBC model. Furthermore, there is evidence to suggest that the level of increasing returns to scale would not be large enough to allow for sunspot equilibria.

Finally, we are also able to compare the effect of allowing some parameters to vary. The log-Bayes factors in favor of allowing parameters to be estimated versus using the calibrated parameters can be found in Table 6 below.

degree of returns to scale is $0.24/0.7844 + 0.7/0.7844 = 1.20$.

¹²They use a value of $\lambda=0.58$ to generate the necessary level of increasing returns to scale.

Table 6: Log-Bayes Factors: Calibration vrs. Estimation

	ψ_z free	ψ_z and λ free
CRS	405.0 (0.07)	-
IRS	436.2 (0.12)	439.2 (0.60)

It is clear from the log-Bayes factors given in Table 6 that the models using the estimated parameters clearly out-perform, in terms of out-of-sample prediction performance, the models using the calibrated values. Therefore allowing for prior uncertainty in some of the parameters of the model leads to dramatic improvements in the performance of both models.

Table 7: Parameter Estimates: CRS Model

Parameter	CRS Model	IRS Model
Comparison I		
\hat{K}_1	0.0117 (0.1723)	0.0005 (0.1317)
Comparison II		
\hat{K}_1	0.0000 (0.0066)	-0.0061 (0.1101)
$\text{var}(Z_1)$	0.0070^2 (3.4×10^{-6})	0.0072^2 (3.5×10^{-6})
$\text{var}(Z_2)$	0.0071^2 (3.7×10^{-6})	0.0072^2 (3.5×10^{-6})
$\text{corr}(Z_1, Z_2)$	0.9972 (0.0005)	0.9998 (4.1×10^{-5})
Ω	$\begin{bmatrix} 0.9880 & 0.0078 \\ 0.0677 & 0.9551 \end{bmatrix}$ $\begin{bmatrix} (0.0065) & (0.0062) \\ (0.0185) & (0.0204) \end{bmatrix}$	$\begin{bmatrix} 0.9466 & 0.0541 \\ 0.0411 & 0.9577 \end{bmatrix}$ $\begin{bmatrix} (0.0220) & (0.0225) \\ (0.0190) & (0.0192) \end{bmatrix}$
Comparison III		
\hat{K}_1	0.0000 (0.0066)	-0.0158 (0.1282)
$\text{var}(Z_1)$	0.0070^2 (3.4×10^{-6})	0.0072^2 (3.4×10^{-6})
$\text{var}(Z_2)$	0.0071^2 (3.7×10^{-6})	0.0072^2 (3.4×10^{-6})
$\text{corr}(Z_1, Z_2)$	0.9972 (0.0005)	0.9999 (2.1×10^{-5})
Ω	$\begin{bmatrix} 0.9880 & 0.0078 \\ 0.0677 & 0.9551 \end{bmatrix}$ $\begin{bmatrix} (0.0065) & (0.0062) \\ (0.0185) & (0.0204) \end{bmatrix}$	$\begin{bmatrix} 0.9506 & 0.0502 \\ 0.0365 & 0.9625 \end{bmatrix}$ $\begin{bmatrix} (0.0181) & (0.0184) \\ (0.0156) & (0.0158) \end{bmatrix}$
λ	-	0.7844 (0.1094)

6 Conclusion

In this paper, we employed likelihood methods to compare two variants of a standard international real business cycle model. The two variants were based on models presented in Guo and Sturzenegger (1998), where they argued that a model with an increasing returns to scale technology had better properties than the standard constant returns to scale model. The aim of this paper was to formally compare these two models using the full dimension of the data, rather than comparing the models using a limited number of moments from the data.

The formal comparison used in this paper utilized Bayesian model comparison techniques, in particular, Bayes factors. The main reason for using Bayes factors in comparing models, is the fact that the Bayes factor can be interpreted as the relative cumulative out-of-sample prediction performance of one model over another. Apart from being a natural way to compare two or more models, the Bayes factor also satisfies the likelihood criteria, in that all inferences are based on information contained in the likelihood function. However, in order to use Bayes factors, a likelihood function needs to be computed for each model. The models compared in this paper are DSGE models that suffer from a dimensionality problem; that is, there are fewer stochastic terms than variables modelled. In order to overcome this problem, a linearized version is computed to represent each model.

An algorithm is proposed to calculate the value of the likelihood function of the two models for any values of the parameters by constructing as many independent linear combinations of the data as there are stochastic terms. Therefore, we are able to evaluate the likelihood function of the two models under comparison at any value of the parameter space. We are then able to use the random-walk Hastings-Metropolis algorithm to estimate the free parameters of the models and to calculate the marginal likelihoods and Bayes factors.

There are three comparisons undertaken in this paper. The first comparison replicates the comparison undertaken in Guo and Sturzenegger (1998) in that the values of the parameters are set to their calibrated values and are not allowed to be freely estimated. Then we allow the parameters that describe the law of motion of the stochastic terms to be freely estimated in order to check the robustness of the results in the first comparison to the persistence and variability of the two shocks in the model. Finally, we allow the parameter that determines the degree of increasing returns to scale to be freely estimated.

We find strong evidence in favor of having a constant returns to scale technology in all three comparisons. That is, the result that the IRBC model with constant returns to scale technology is preferred to the IRBC model with increasing returns to scale technology is robust to changes in the degree of increasing returns to scale and to the persistence and variability of the stochastic components of the models. Moreover, the degree of increasing returns to scale that is estimated is smaller than that calibrated in Guo and Sturzenegger (1998). Hence there seems to be little evidence of the degree of increasing returns to scale necessary for the sunspot model that Guo and Sturzenegger support. We also show that the estimated values that govern the law of motion of the shock terms and the degree of increasing returns strongly improve the out-of-sample prediction performance of all models used in this paper. In conclusion, we find no evidence that increasing returns to scale improves the performance of the model used in this paper.

It should be noted, however that these results are conditional on the models used and the calibrated values of the structural parameters that we did not allow to be freely estimated. Checks on the robustness of our results to the form of the model and to the values of the structural parameters has been left for further research.

References

- Anderson, E.W., L.P. Hansen, E.R. McGratten, and T.J. Sargent (1996) ‘Mechanics of forming and estimating dynamic linear economies.’ In *Handbook of Computational Economics*, ed. H.M. Amman, D.A. Kendrick, and J. Rust, vol. I (Elsevier Science B.V.)
- Backus, D., P. Kehoe, and F. Kydland (1992) ‘International real business cycles.’ *Journal of Political Economy* 100(4), 745–75
- Baxter, M., and R. King (1991) ‘Productive externalities and the business cycles.’ Working Paper 53, Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis
- Canova, F, and A. Ubide (1998) ‘International business cycles, financial markets and household production.’ *Journal of Economic Dynamics and Control* 22(4), 545–72
- Canova, F., and E. Ortega (1995) ‘Testing calibrated general equilibrium models.’ unpublished manuscript
- Chib, S., and E. Greenberg (1995) ‘Understanding the metropolis-hastings algorithm.’ *The American Statistician*
- DeJong, D., B. Ingram, and C. Whiteman (2000) ‘A Bayesian approach to dynamic macroeconomics.’ *Journal of Econometrics* 98(2), 203–223
- Devereux, M., A. Gregory, and G. Smith (1991) ‘Realistic cross-country consumption correlations in a two-country, equilibrium, business cycle model.’ *Journal of International Money and Finance* 11, 3–16
- Farmer, R.E.A. (1993) *The Macroeconomics of Self-Fulfilling Prophecies* (Cambridge, Massachusetts: M.I.T. Press)

- Farmer, R.E.A., and J-T. Guo (1994) ‘Real business cycles and the animal spirits hypothesis.’ *Journal of Economic Theory* 63, 42–72
- Gelfand, A.E., and D.K. Dey (1994) ‘Bayesian model choice: Asymptotics and exact calculations.’ *Journal of the Royal Statistical Society Series B*
- Geweke, J.F. (1995) ‘Bayesian comparison of econometric models.’ University of Minnesota Working Paper.
- (1999) ‘Using simulation methods for Bayesian econometric models: Inference, development and communication.’ *Econometric Reviews* 18(1), 1–73
- Guo, J-T., and F. Sturzenegger (1998) ‘Crazy explanations of international business cycles.’ *International Economic Review* 39(1), 111–133
- Hodrick, R.J., and E.C. Prescott (1997) ‘Post-war U.S. business cycles: An empirical investigation.’ *Journal of Money, Credit and Banking* 29(1), 1–16
- Ireland, P.N. (1999) ‘A method for taking models to data.’ Department of Economics Working Paper 421, Boston College.
- Johnson, R. A., and D. W. Wichern (1988) *Applied Multivariate Statistical Analysis*, second ed. (Prentice-Hall of Australia Pty. Limited, Sydney: Prentice-Hall International)
- Kim, K, and A.R. Pagan (1995) ‘The econometric analysis of calibrated macroeconomic models.’ In *Handbook of Applied Econometrics*, ed. H.H. Pesaran and Mike Wickens (Blackwell Press)
- Kydland, F.E., and E.C. Prescott (1982) ‘Time to build and aggregate fluctuations.’ *Econometrica* 50(6), 1345–70
- (1996) ‘The computational experiment: An econometric tool.’ *Journal of Economic Perspectives* 10(1), 69–85

- Landon-Lane, J.S. (1998) 'Bayesian comparison of dynamic macroeconomic models.' PhD dissertation, University of Minnesota
- (2002) 'Evaluating dynamic stochastic general equilibrium models using likelihood methods.' Department of Economics Working Paper 2002-11, Rutgers University
- Sargent, T.J. (1989) 'Two models of measurements and the investment accelerator.' *Journal of Political Economy* 97, 251–287
- Sims, C.A. (1996) 'Macroeconomics and methodology.' *The Journal of Economic Perspectives* 10(1), 105–20
- Tierney, L. (1994) 'Markov chains for exploring posterior distributions.' *Annals of Statistics* 22(4), 1701–1762