Econometric Analysis of Linearized Singular Dynamic Stochastic General Equilibrium Models

Herman J. Bierens*
Pennsylvania State University

Abstract

In this paper I propose an alternative to calibration of linearized singular dynamic stochastic general equilibrium models. Given an a-theoretical econometric model as a representative of the data generating process, I will construct an information measure which compares the conditional distribution of the econometric model variables with the corresponding singular conditional distribution of the theoretical model variables. The singularity problem will be solved by using convolutions of both distributions with a non-singular distribution. This information measure will then be maximized to the deep parameters of the theoretical model, which links these parameters to the parameters of the econometric model and provides an alternative to calibration. This approach will be illustrated by an application to a linearized version of the stochastic growth model of King, Plosser and Rebelo.

Key words: Estimation, Dynamic general equilibrium, Singularity, Calibration

JEL classifications: C13, C32, C52, D90

*Correspondence address: Herman J. Bierens, Department of Economics, Pennsylvania State University, 608 Kern Graduate Building, University Park, PA 16802. E-mail: hbierens@psu.edu.

1 Introduction

In the forties through the sixties of the past century the development of macroeconometrics was inspired and directed by Keynesian macroeconomic theory, and vice versa, the construction and estimation of large Keynesian macroeconomic models was facilitated by econometrics, in particular simultaneous equations theory. With the rise of neoclassical Dynamic Stochastic General Equilibrium (DSGE) macroeconomics, however, econometrics and economic theory have grown apart, to the point where most macro-theorists now consider econometrics irrelevant for what they do. Admittedly, quite a few econometricians have the same attitude towards theoretical macroeconomics.

Most economic theories, including DSGE theory, are partial theories in the sense that only a few related economic phenomena are studied. The analysis of this "partial" theory is justified, explicitly or implicitly, by the ceteris paribus assumption (other things being equal or constant). See Bierens and Swanson (2000) and the references therein. However, when simple economic models of this type are estimated using data which are themselves generated from a much more complex real economy, it is not surprising that they often fit poorly. Thus, these models do not represent data generating processes, and are not designed to do. The purpose of these models is to gain insight in particular related economic phenomena rather than to describe an actual economy, and to conduct numerical experiments. Consequently, most macro-theorists do not bother to estimate their models, but instead calibrate the model parameters. See Hansen and Heckman (1996) for a review of calibration, and Sims (1996) and Kydland and Prescott (1996) for opposite views on calibration.

The literature on econometric analysis of DSGE models can be divided in two rather short strands. One strand of literature is concerned with model evaluation, i.e., the problem how to measure the fit of these models. The other strand of literature is concerned with finding alternatives to calibration.

Watson (1993) proposes to augment the variables in the theoretical model with just enough stochastic error so that the model can match the second moments of the actual data. Measures of fit for the model, called relative mean square approximation errors, are then constructed on the basis of the variance of this stochastic error relative to the variance of the actual series. An alternative approach is to compare the empirical VAR innovation response curves with those computed on the basis of artificial data generated by the calibrated theoretical model. See for example the papers in Pagan (1994), in particular Feve and Langot (1994) and Nason and Cogley (1994). Schorfheide (2000) compares two DSGE models with a benchmark model, using a Bayesian approach. Bierens and Swanson (2000) propose a new measure of fit, called the average conditional reality bound, which compares the non-singular part of a linearized DSGE model with a corresponding marginalized econometric model. Corradi and Swanson (2004) also compare DSGE models with a benchmark model, using squared differences of their distribution functions.

DeJong, Ingram, and Whiteman (1996, 2000) and Geweke (1999) propose a Bayesian approach. They assume prior distribution for the deep parameters centered around calibrated values. This is indeed a natural extension of calibration. However, there are two major limitations to the Bayesian approach. First, one has to assume that conditional on the parameters the theoretical model represents the data generating process, which is too far-fetched an assumption. Second, the Bayesian approach requires the existence of the conditional density of the model variables, whereas in most of these models the model variables are driven by only a few random shock. The latter renders the theoretical distribution involved singular. DeJong, Ingram, and Whiteman (1996, 2000) circumvent the singularity problem by focusing on a subset of model variables for which the conditional distribution is non-singular. Geweke (1999) applies the Bayesian approach to a one-dimensional equity premium model. Ireland (2003) proposes to add noise to a linearized DSGE model in order to estimate the resulting hybrid model by maximum likelihood. Also this approach suffers from the limitation that one has to assume that the hybrid model involved represents the data generating process.

The singularity problem also prevents direct estimation of a DSGE model by GMM, because due to the singularity some moment conditions will hold exactly for each time period, so that the number of moment condition will exceed the number of observations. Therefore, the application of GMM is only possible after (explicitly or implicitly) adding noise to the exact moment equations. See for example Ambler et al. (2003).

In this paper I will propose an alternative non-Bayesian approach to calibration of singular DSGE models, which takes into account that these models do not represent data generating processes and are singular. Given an a-theoretical econometric model as a representative of the data generating process, I will construct an information measure (called the multiplicative conditional reality bound) which compares the conditional distribution of the econometric model variables with the corresponding singular conditional distribution of the theoretical model variables, along the lines in Bierens and Swanson (2000). The singularity problem will be solved by using convolutions of **both** distributions with a nonsingular distribution. This information criterion can be interpreted as the probability that the distribution of the convoluted econometric model is generated by the distribution of the convoluted theoretical model, conditional on the data. The information criterion involved will then be maximized to the deep parameters of the theoretical model, which links these parameters to the parameters of the econometric model and provides an alternative to calibration.

This approach will be applied to a linearized version of the stochastic growth model of King, Plosser and Rebelo (KPR) (1988a,b). The linearization procedure is different from the one proposed by KPR, though. I will solve the model without using linearization to the point where the only control variable left is the consumption-output ratio. At that point I will only linearize the state variable process of the concentrated model around the deterministic steady state, and link the parameters of the linearized model to the deep parameters. On the other hand, KPR linearize the (deterministic) Lagrange multiplier solution of their model at an earlier stage. Although it is not impossible to link the parameters of their linearized model to the deep parameters[1], it is more complicated than in my approach. Consequently, KPR (1988a) do not provide this link, except for a deterministic version of their model with fixed labor. See KPR (1988a, Footnote 17).

A separate appendix to this paper containing the details of some tedious derivations is downloadable from web page

http://econ.la.psu.edu/~hbierens/SDSGEMAPP.PDF.

2 Singularity

Dynamic stochastic general equilibrium (DSGE) models, possibly after transforming the control and state variables, take the form of a dynamic stochastic optimization problem:

$$\max \ E_0\left[\sum_{t=0}^{\infty}\lambda^t g(C_t,S_t)\right] \tag{1}$$

[1]See for example Ireland (2003), who analyses Hansen's (1985) model, which is similar to the KPR model.

subject to

$S_t = f(C_t, S_{t-1}, U_t)$,

where g is a utility function, U_t is a vector of random shocks, $S_t \in \mathbb{R}^s$ is a vector of state variables, $C_t \in \mathbb{R}^c$ is a vector of control variables, $\lambda \in (0, 1)$ is a time preference parameter, and E_0 is the conditional expectation operator relative to the information up to time $t = 0$.

Often the dimension m of U_t is smaller than the dimension s of S_t, which renders the conditional distribution of S_t given the past singular. However, this is not the only source of singularity.

As is well-known [see Stokey, Lucas and Prescott (1989)], under some regularity conditions the solution to this dynamic programming problem takes the form of a contingency plan for C_t, which is a time-invariant Borel measurable mapping Ψ: $\mathbb{R}^s \to \mathbb{R}^c$ such that $C_t = \Psi(S_{t-1})$. Thus, the dynamic optimization problem (1) can then be reformulated as

$$\max_{\Psi:\mathbb{R}^s \to \mathbb{R}^c} E_0 \left[\sum_{t=0}^{\infty} \lambda^t g\left(\Psi(S_{t-1}), S_t\right) \right] \tag{2}$$

subject to

$$S_t = f\left(\Psi(S_{t-1}), S_{t-1}, U_t\right).$$

The solution $C_t = \Psi(S_{t-1})$ is another source of singularity of the joint conditional distribution of the model variables $Y_t = \binom{C_t}{S_t}$, because it holds without errors. Therefore, singularity cannot be cured by including more random shocks in the state variables process. Thus, singularity is an inherent property of dynamic stochastic general equilibrium models.

Given the optimal contingency plan $C_t = \Psi(S_{t-1})$, model (2) postulates that at time t,

$$Y_t = \begin{pmatrix} C_t \\ S_t \end{pmatrix} = \begin{pmatrix} \Psi(S_{t-1}) \\ f\left(\Psi(S_{t-1}), S_{t-1}, U_t\right) \end{pmatrix} = H(Y_{t-1}, U_t|\beta) \in \mathbb{R}^k,$$

say, where $k = s + c$, β is a vector of "deep" parameters, $U_t \in \mathbb{R}^m$ is a stochastic shock process with dimension $m < k$, and H is the functional specification of the model. In particular, linearized versions of these models take the form of a singular VAR(p) model:

$$Y_t = A(\beta)X_{t-1} + B(\beta)U_t, \tag{3}$$

where $X_{t-1} = (1, Y_{t-1}^T, ...Y_{t-p}^T)^T$, $A(\beta)$ is an $k \times k.p$ matrix, and $B(\beta)$ is a $k \times m$ matrix of functions of β. If it is assumed (as usual) that U_t is an m-variate Gaussian white noise process: $U_t \sim i.i.d.\ N_m[0, I_m]$, then the distribution of Y_t conditional on X_{t-1} is singular normal: $N_k \left[A(\beta)X_{t-1}, B(\beta)B(\beta)^T\right]$.

3 Linking distributions

In this paper I propose an information criterion, similarly to Bierens and Swanson (2000), which links a linearized singular DSGE model of the type (3) to an linear econometric model, for example a Gaussian VAR(q) model

$$Y_t = \Pi Z_{t-1} + V_t,\ V_t \sim i.i.d.\ N_k[0, \Theta],\ \det(\Theta) \neq 0, \tag{4}$$

where $Z_{t-1} = (1, Y_{t-1}^T, ...Y_{t-q}^T)^T$ and Π is a $k \times (k.q)$ matrix of coefficients. The econometric model (4) is assumed to represent the data-generating process. This procedure then links the vector β of the parameters of the theoretical model to the parameter matrices Π and Θ of the econometric model (4): $\beta = \Phi(\Pi, \Theta)$, say. Plugging in the maximum likelihood estimators $\widehat{\Pi}$ and $\widehat{\Theta}$ of Π and Θ then yields an estimator $\widehat{\beta} = \Phi(\widehat{\Pi}, \widehat{\Theta})$ of β. Moreover, using the well-known delta method it is then possible to derive confidence intervals of the estimates in $\widehat{\beta}$.

3.1 Embedding densities

Consider two densities, $f(y)$ and $f_0(y)$, with common support. It is always possible to squeeze $f_0(y)$ under $f(y)$, by multiplying $f_0(y)$ by a "squeeze" factor $p_0 \in [0, 1]$, such that $p_0 f_0(y) \leq f(y)$ for all y. The maximal p_0 involved is:

$$p_0 = \inf_y [f(y)/f_0(y)] \leq 1.$$

Of course, it is possible that $p_0 = 0$. This procedure is illustrated in Figure 1.

Now $f(y)$ can be written as a mixture:

$$f(y) = p_0 f_0(y) + (1 - p_0) f_1(y),$$

where $f_1(y) = (f(y) - p_0 f_0(y)) / (1 - p_0)$ is a density. Thus, if we draw with probability p_0 from the distribution with density $f_0(y)$ and with probability

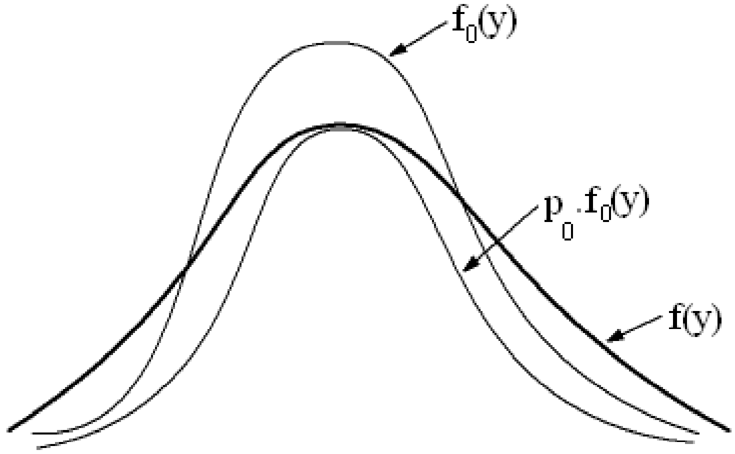

Figure 1: Embedding densities

$1-p_0$ from the distribution with density $f_1(y)$, the result is actually a random drawing from the distribution with density $f(y)$.

Note that the quantity $-\ln(p_0)$ is an information criterion which compares the closeness of the two densities $f_0(y)$ and $f(y)$, i.e., $-\ln(p_0) = 0$ if and only if $f_0(y) \equiv f(y)$, and $-\ln(p_0) > 0$ otherwise.

3.2 Convolutions

In order to use this information criterion to compare the conditional distributions of (3) and (4) we need to make their supports equal. This will be done by using convolutions. The idea is to add i.i.d. nonsingular k-variate normal noise R_t^*, R_t^{**} to Y_t in (3) and (4), respectively[2], so that (3) becomes

$$Y_t^{TM} = A(\beta)X_{t-1} + B(\beta)U_t + R_t^*, \tag{5}$$

and (4) becomes

$$Y_t^{EM} = \Pi Z_{t-1} + V_t + R_t^{**}. \tag{6}$$

[2] The reason for adding this noise to both models is to keep the conditional distributions involved comparable.

Then the conditional distribution of Y_t^{TM} in (5) is absolutely continuous with density $h_{_{TM}}(y|X_{t-1},\beta)$, say. Next, let

$$p_{t-1}(\Pi,\Theta,\beta) = \inf_y \frac{h_{_{EM}}(y|Z_{t-1},\Pi,\Theta)}{h_{_{TM}}(y|X_{t-1},\beta)}, \tag{7}$$

where $h_{_{EM}}(y|Z_{t-1},\Pi,\Theta)$ is the conditional density of Y_t^{EM} in (6).

The interpretation of $p_{t-1}(\Pi,\Theta,\beta)$ is similar as before: if we draw Y_t^{EM} and Y_t^{ET} randomly from the conditional distributions with densities $h_{_{TM}}(y| X_{t-1},\beta)$, and $h_{_{EM}}(y|Z_{t-1},\Pi,\Theta)$, respectively, then $p_{t-1}(\Pi,\Theta,\beta)$ may be interpreted as the probability that Y_t^{EM} and Y_t^{ET} have the same conditional distribution, given the data up to time $t-1$. Moreover,

$$\prod_{t=1}^{n} p_{t-1}(\Pi,\Theta,\beta)$$

may then be interpreted as the probability that **conditional on the data** the joint distribution of Y_1^{ET}, ..., Y_n^{ET} is the same as the joint distribution of Y_1^{EM}, ..., Y_n^{EM}. The latter suggests to link the parameter vector β of the linearized DSGE model (3) to the estimated parameters of the econometric model (4) by

$$\beta_n(\Pi,\Theta) = \arg\max_\beta \sum_{t=1}^{n} \ln\left[p_{t-1}(\Pi,\Theta,\beta)\right], \tag{8}$$

as an alternative to calibration.

Note that similar to Bierens and Swanson (2000) the statistic

$$\max_\beta \left(\prod_{t=1}^{n} p_{t-1}(\Pi,\Theta,\beta)\right)^{1/n}$$

may be used as a reality check on the theoretical model.

This approach is somewhat related to the approach taken by Watson (1993), where the variables in the theoretical model are augmented with just enough stochastic error so that the model can match the second moments of the actual data. One of the crucial differences with the Watson approach is that I propose to augment also the actual data with the same stochastic error, in order to penalize the singularity of the theoretical model. Another fundamental difference is that Watson uses calibrated parameters, whereas I will estimate them.

Of course, there are alternative ways to link theoretical models to econometric models. For example, one could use

$$\begin{aligned}\beta_n(\Pi, \Theta) &= \arg\min_{\beta} \sum_{t=1}^{n} \int \ln\left(\frac{h_{EM}(y|Z_{t-1}, \Pi, \Theta)}{h_{TM}(y|X_{t-1}, \beta)}\right) \\ &\quad \times h_{EM}(y|Z_{t-1}, \Pi, \Theta) dy\end{aligned} \tag{9}$$

instead of (8). The integral in (9) is the well-known Kullback-Leibler (1951) information criterion, which measures the closeness of the two densities involved. However, recall that $-\ln[p_{t-1}(\Pi, \Theta, \beta)]$ is also an information criterion which measures the closeness of the two densities involved. In particular, $-\ln[p_{t-1}(\Pi, \Theta, \beta)] = 0$ if and only if $h_{EM}(y|Z_{t-1}, \Pi, \Theta) = h_{TM}(y|Y_{t-1}, \beta)$ for all y in the common support of the two densities involved, and $-\ln[p_{t-1}(\Pi, \Theta, \beta)] > 0$ otherwise. The main reason for using (8) is the neat interpretation of $p_{t-1}(\Pi, \Theta, \beta)$ as the probability that a random drawing from $h_{TM}(y|X_{t-1}, \beta)$ generates a random drawing from $h_{EM}(y|Z_{t-1}, \Pi, \Theta)$.

3.3 Implementation

In order to make this approach operational, let us rewrite the conditional distribution of the variables in the linearized DSGE model (3) as

$$Y_t^* \sim N_k\left[\mu_{t-1}(\beta), \Sigma(\beta)\right], \ \det(\Sigma(\beta)) = 0, \tag{10}$$

where $\mu_{t-1}(\beta) = A(\beta)X_{t-1}$ [3] and $\Sigma(\beta) = B(\beta)B(\beta)^T$, and the conditional distribution of the variables in the econometric model (4) as

$$Y_t^{**} \sim N_k\left[\eta_{t-1}, \Theta\right], \ \det(\Theta) > 0, \tag{11}$$

where $\eta_{t-1} = \Pi Z_{t-1}$. Now add independent equally distributed noise R_t^*, R_t^{**} to the dependent variables in (10) and (11), respectively. A natural choice is

$$R_t^*, R_t^{**} \sim N_k\left[0, \tau\Theta\right] \text{ for some } \tau > 0.$$

Then

$$Y_t^{TM} = Y_t^* + R_t^* \sim N_k\left[\mu_{t-1}(\beta), \Sigma(\beta) + \tau\Theta\right]$$

[3] In our application below $\mu_{t-1}(\beta)$ is actually a function of infinitely many lagged Y_t's because the linearized DSGE model involved is partly an ARMA(1,1) process.

and

$$Y_t^{EM} = Y_t^{**} + R_t^{**} \sim N_k \left[\eta_{t-1}, (1+\tau)\Theta\right],$$

with corresponding densities

$$h_{t-1}^{TM}(y|\beta,\tau) = \frac{\exp\left[-\frac{1}{2}\left(y-\mu_{t-1}(\beta)\right)^T \left(\Sigma(\beta)+\tau\Theta\right)^{-1}\left(y-\mu_{t-1}(\beta)\right)\right]}{\left(\sqrt{2\pi}\right)^k \sqrt{\det\left(\Sigma(\beta)+\tau\Theta\right)}}$$

and

$$h_{t-1}^{EM}(y|\tau) = \frac{\exp\left[-\frac{1}{2(1+\tau)}\left(y-\eta_{t-1}\right)^T \Theta^{-1}\left(y-\eta_{t-1}\right)\right]}{\left(\sqrt{2\pi}\right)^k \sqrt{(1+\tau)^k \det\left(\Theta\right)}},$$

respectively. Hence

$$\begin{aligned}
&\frac{h_{t-1}^{EM}(y|\tau)}{h_{t-1}^{TM}(y|\beta,\tau)} = \sqrt{\frac{\det(\Sigma(\beta)+\tau\Theta)}{(1+\tau)^k \det\left(\Theta\right)}} \qquad (12)\\
&\times \exp\left[\frac{1}{2}y^T\Psi(\beta,\tau)y - y^T\Psi(\beta,\tau)\mu_{t-1}(\beta)\right.\\
&+\frac{1}{1+\tau}y^T\Theta^{-1}\left(\eta_{t-1}-\mu_{t-1}(\beta)\right) + \frac{1}{2}\mu_{t-1}(\beta)^T\left(\Sigma(\beta)+\tau\Theta\right)^{-1}\mu_{t-1}(\beta)\\
&\left.-\frac{1}{2}\left(1+\tau\right)^{-1}\eta_{t-1}^T\Theta^{-1}\eta_{t-1}\right],
\end{aligned}$$

where

$$\begin{aligned}
\Psi(\beta,\tau) &= \left(\Sigma(\beta)+\tau\Theta\right)^{-1} - \frac{1}{1+\tau}\Theta^{-1} \qquad (13)\\
&= \frac{1}{1+\tau}\Theta^{-1/2}\left(\left(\frac{1}{1+\tau}\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2} + \frac{\tau}{1+\tau}I\right)^{-1} - I\right)\Theta^{-1/2}.
\end{aligned}$$

The matrix $\Psi(\beta,\tau)$ is positive definite if all the eigenvalues of $\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2}$ are less than 1, or equivalently if all the solutions of the generalized eigenvalue problem

$$\det\left(\Sigma(\beta)-\lambda\Theta\right) = 0 \qquad (14)$$

are less than 1. If so, (12) is minimal for

$$y = \mu_{t-1}(\beta) - \frac{1}{1+\tau}\Psi(\beta,\tau)^{-1}\Theta^{-1}\left(\eta_{t-1} - \mu_{t-1}(\beta)\right).$$

Substituting this solution in (12) yields, after some straightforward derivations,

$$\begin{aligned} \inf_y \frac{h_{t-1}^{EM}(y|\tau)}{h_{t-1}^{TM}(y|\beta,\tau)} &= \sqrt{\frac{\det(\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2}+\tau I)}{(1+\tau)^k}} \\ &\times \exp\left(\frac{-1}{2(1+\tau)}\left(\eta_{t-1}-\mu_{t-1}(\beta)\right)^T \Phi(\beta,\tau)\left(\eta_{t-1}-\mu_{t-1}(\beta)\right)\right), \end{aligned}$$

where

$$\Phi(\beta,\tau) = \frac{1}{1+\tau}\left[\Theta^{-1}\Psi(\beta,\tau)^{-1}\Theta^{-1} + (1+\tau)\Theta^{-1}\right] = \tag{15}$$

$$= \Theta^{-1/2}\left[\left(\left(\frac{1}{1+\tau}\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2} + \frac{\tau}{1+\tau}I_k\right)^{-1} - I_k\right)^{-1} + I_k\right]\Theta^{-1/2},$$

provided that the maximum eigenvalue of $\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2}$ is less than 1:

$$\lambda_{\max}\left[\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2}\right] < 1.$$

Next, let $\lambda_1(\beta) \geq \lambda_2(\beta) \geq ... \geq \lambda_m(\beta)$ be the m positive eigenvalues of $\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2}$. We can write

$$\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2} = Q(\beta)\Lambda(\beta)Q(\beta)^T = Q_1(\beta)\Lambda_1(\beta)Q_1(\beta)^T,$$

where $\Lambda_1(\beta) = \mathrm{diag}(\lambda_1(\beta), \lambda_2(\beta), ..., \lambda_m(\beta))$, $Q_1(\beta)$ is the $k \times m$ matrix of corresponding orthonormal eigenvectors, and $Q_2(\beta)$ is the $k \times (k-m)$ matrix of orthonormal eigenvectors corresponding to the zero eigenvalues. Thus,

$$\begin{aligned} \det(\Theta^{-1/2}\Sigma(\beta)\Theta^{-1/2} + \tau I_k) &= \tau^{k-m}\prod_{j=1}^{m}(\lambda_j(\beta)+\tau) \\ &= \tau^{k-m}\det\left(\Lambda_m(\beta) + \tau I_m\right) \end{aligned}$$

and

$$\begin{aligned} \Phi(\beta,\tau) &= \Theta^{-1/2}Q(\beta)\left[\left((1+\tau)\left(\Lambda(\beta)+\tau I_k\right)^{-1} - I_k\right)^{-1} + I_k\right]Q(\beta)^T\Theta^{-1/2} \\ &= (1+\tau)\,\Theta^{-1/2}Q(\beta)\left(I_k - \Lambda(\beta)\right)^{-1}Q(\beta)^T\Theta^{-1/2}. \end{aligned}$$

Hence

$$p_{t-1}(\Pi,\Theta,\beta|\tau)=\inf_y \frac{h_{t-1}^{EM}(y|\tau)}{h_{t-1}^{TM}(y|\beta,\tau)}=\sqrt{\frac{\tau^{k-m}\det\left(\Lambda_1(\beta)+\tau I_m\right)}{(1+\tau)^k}} \tag{16}$$
$$\times\exp\left[-\tfrac{1}{2}\left(\eta_{t-1}-\mu_{t-1}(\beta)\right)^T\Theta^{-1/2}Q(\beta)\left(I_k-\Lambda(\beta)\right)^{-1}Q(\beta)^T\Theta^{-1/2}\right.$$
$$\left.\times\left(\eta_{t-1}-\mu_{t-1}(\beta)\right)\right].$$

Now suppose that the parameter vector β can be partitioned as $\beta=\binom{\beta_1}{\beta_2}$ such that $\mu_{t-1}(\beta)=\mu_{t-1}(\beta_1)$ and β_2 is the vector of stacked diagonal and upper-diagonal elements of $\Sigma(\beta)$. The latter implies that $\Sigma=\Sigma(\beta)$ is unconstrained, except for the condition that Σ is positive semi-definite with rank $m<k$, and that $\lambda_{\max}\left[\Theta^{-1/2}\Sigma\Theta^{-1/2}\right]<1$. Then $Q=(Q_1,Q_2)=Q(\beta)$ and $\Lambda_1=\Lambda_1(\beta)$ are unconstrained too, except of course for the conditions that Q is orthogonal and that the diagonal elements $\lambda_1\geq....\geq\lambda_m$ of Λ_1 are confined to the unit interval $(0,1)$. Equation (16) can now be rewritten as

$$p_{t-1}(\Pi,\Theta,\beta_1,Q,\Lambda_1|\tau)=\sqrt{\frac{\tau^{k-m}\det\left(\Lambda_1+\tau I_m\right)}{(1+\tau)^k}} \tag{17}$$
$$\times\exp\left[-\tfrac{1}{2}\left(\eta_{t-1}-\mu_{t-1}(\beta_1)\right)^T\Theta^{-1/2}Q\left(I_k-\Lambda\right)^{-1}Q^T\Theta^{-1/2}\right.$$
$$\left.\times\left(\eta_{t-1}-\mu_{t-1}(\beta_1)\right)\right],$$

hence

$$\frac{1}{n}\sum_{t=1}^{n}\ln\left[p_{t-1}(\Pi,\Theta,\beta_1,Q,\Lambda_1|\tau)\right] \tag{18}$$
$$=\frac{k-m}{2}\ln(\tau)-\frac{k}{2}\ln(1+\tau)+\frac{1}{2}\sum_{j=1}^{m}\ln\left(\lambda_j+\tau\right)$$
$$-\frac{1}{2}\sum_{j=1}^{m}\frac{q_j^T\Gamma_n(\beta_1)q_j}{1-\lambda_j}-\frac{1}{2}\sum_{j=m+1}^{k}q_j^T\Gamma_n(\beta_1)q_j,$$

say, where $Q=(q_1,...,q_k)$ and

$$\Gamma_n(\beta_1)=\Theta^{-1/2}\left[\frac{1}{n}\sum_{t=1}^{n}\left(\eta_{t-1}-\mu_{t-1}(\beta_1)\right)\left(\eta_{t-1}-\mu_{t-1}(\beta_1)\right)^T\right]\Theta^{-1/2}. \tag{19}$$

Because $(1-\lambda_1)^{-1} \geq ... \geq (1-\lambda_m)^{-1} \geq 1$, it is easy to verify that the last two terms in (18) are maximal if we choose the q_j's equal to the orthonormal eigenvectors of $\Gamma_n(\beta_1)$ corresponding to its increasingly ordered eigenvalues. Thus, let

$$\Xi_n(\beta_1) = \text{diag}\left(\xi_{n,1}(\beta_1), ..., \xi_{n,k}(\beta_1)\right),$$

where $\xi_{n,1}(\beta_1) \leq .... \leq \xi_{n,k}(\beta_1)$ are the eigenvalues of $\Gamma_n(\beta_1)$, with corresponding orthogonal matrix $Q_n(\beta_1)$ of eigenvectors. Then

$$\begin{aligned} & \min_{Q^T Q = I_k} \text{trace}\left[(I_k - \Lambda)^{-1} Q^T \Gamma_n(\beta_1) Q\right] \\ = \ & \text{trace}\left[(I_k - \Lambda)^{-1} \Xi_n(\beta_1)\right] = \sum_{j=1}^{m} \frac{\xi_{n,j}(\beta_1)}{1-\lambda_j} + \sum_{j=m+1}^{k} \xi_{n,j}(\beta_1), \end{aligned}$$

hence

$$\begin{aligned} & \max_{Q^T Q = I_k} \frac{1}{n} \sum_{t=1}^{n} \ln\left[p_{t-1}(\Pi, \Theta, \beta_1, Q, \Lambda_1 | \tau)\right] \\ = \ & \frac{k-m}{2} \ln(\tau) - \frac{k}{2} \ln(1+\tau) + \frac{1}{2} \sum_{j=1}^{m} \ln(\lambda_j + \tau) \\ & -\frac{1}{2} \sum_{j=1}^{m} \frac{\xi_{n,j}(\beta_1)}{1-\lambda_j} - \frac{1}{2} \sum_{j=m+1}^{k} \xi_{n,j}(\beta_1). \end{aligned}$$

Next, maximize $\varphi_{n,j}(\lambda_j) = \ln(\lambda_j + \tau) - \xi_{n,j}(\beta_1)/(1-\lambda_j)$ to $\lambda_j \in [0,1]$. The optimal value of λ_j is

$$\begin{aligned} \lambda_j(\beta_1, \tau) &= 1 + \frac{\xi_{n,j}(\beta_1)}{2} \\ & \quad -\frac{1}{2} \sqrt{\xi_{n,j}(\beta_1)^2 + 4(1+\tau)\xi_{n,j}(\beta_1)} \text{ if } \xi_{n,j}(\beta_1)^{-1} > \tau \\ \lambda_j(\beta_1, \tau) &= 0 \text{ if } \xi_{n,j}(\beta_1)^{-1} \leq \tau, \end{aligned} \tag{20}$$

which satisfies $\lambda_j(\beta_1, \tau) < 1$, and $\lambda_j(\beta_1, \tau) > 0$ if and only if

$$\tau < \xi_{n,m}(\beta_1)^{-1}. \tag{21}$$

Then

$$\sup_{Q,\Lambda_1,\lambda_{\max}[\Lambda_1]<1} \frac{1}{n}\sum_{t=1}^{n} \ln\left[p_{t-1}(\Pi,\Theta,\beta_1,Q,\Lambda_1|\tau)\right] \tag{22}$$

$$= \frac{k-m}{2}\ln(\tau) - \frac{k}{2}\ln(1+\tau) + \frac{1}{2}\sum_{j=1}^{m}\ln\left(\lambda_j(\beta_1,\tau)+\tau\right)$$
$$-\frac{1}{2}\sum_{j=1}^{m}\frac{\xi_{n,j}(\beta_1)}{1-\lambda_j(\beta_1,\tau)} - \frac{1}{2}\sum_{j=m+1}^{k}\xi_{n,j}(\beta_1).$$

Moreover, denoting $\Lambda_{1,n}(\beta_1,\tau) = \mathrm{diag}(\lambda_1(\beta_1,\tau),....,\lambda_m(\beta_1,\tau))$, the optimal solution for Σ given β_1 is

$$\Sigma_n(\beta_1,\tau) = \Theta^{1/2}Q_n(\beta_1)\begin{pmatrix}\Lambda_{1,n}(\beta_1,\tau) & O \\ O & O\end{pmatrix}Q_n(\beta_1)^T\Theta^{1/2}. \tag{23}$$

Maximizing (22) to β_1 now yields a solution $\beta_{1,n}(\Pi,\Theta|\tau)$ with corresponding solution for Σ:

$$\Sigma_n(\Pi,\Theta|\tau) = \Sigma_n\left(\beta_{1,n}(\Pi,\Theta|\tau),\tau\right).$$

3.4 Preserving the structure of the theoretical variance matrix

If the variables in model (10) are arranged such that singular variance matrix Σ is block-diagonal, i.e.,

$$\Sigma = \begin{pmatrix}\Sigma_1 & O \\ O & O\end{pmatrix},$$

say, where Σ_1 is an unconstrained positive definite $m\times m$ matrix, one should impose the same structure on the estimate of Σ, as follows. Partition $\Theta^{-1/2}$ conformably to Σ as

$$\Theta^{-1/2} = \begin{pmatrix}\Theta_{11}^* & \Theta_{12}^* \\ \Theta_{21}^* & \Theta_{22}^*\end{pmatrix}.$$

Then (15) becomes

$$\Phi(\tau) = \tag{24}$$
$$\Theta^{-1/2}\left[\begin{pmatrix}\left[(1+\tau)\left(\Theta_{11}^*\Sigma_1\Theta_{11}^*+\tau I_m\right)^{-1} - I_m\right]^{-1} + I_m & O \\ O & (1+\tau)I_{k-m}\end{pmatrix}\right]\Theta^{-1/2}.$$

Next, let

$$\Theta_{11}^{*}\Sigma_1\Theta_{11}^{*} = Q_{11}\Lambda_1 Q_{11}^T,$$

where $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$, with $\lambda_1 \geq \ldots \geq \lambda_m > 0$ the eigenvalues of $\Theta_{11}^{*}\Sigma_1\Theta_{11}^{*}$, and Q_{11} is the $m \times m$ orthogonal matrix of corresponding eigenvectors. Then

$$\Phi(\tau) = (1+\tau)\,\Theta^{-1/2}Q\,(I_k - \Lambda)^{-1}\,Q^T\Theta^{-1/2}, \tag{25}$$

where Λ is the same as before:

$$\Lambda = \begin{pmatrix} \Lambda_1 & O \\ O & O \end{pmatrix},$$

but now

$$Q = \begin{pmatrix} Q_{11} & O \\ O & I_{k-m} \end{pmatrix}.$$

Therefore, the main difference with the previous case is the latter constraint on the matrix Q. Partitioning the matrix (19) according to Q,

$$\Gamma_n(\beta_1) = \begin{pmatrix} \Gamma_{1,1,n}(\beta_1) & \Gamma_{1,2,n}(\beta_1) \\ \Gamma_{2,1,n}(\beta_1) & \Gamma_{2,2,n}(\beta_1) \end{pmatrix},$$

it follows easily that

$$\begin{aligned} & \mathrm{trace}\left[(I_k - \Lambda)^{-1}\,Q^T\Gamma_n(\beta_1)Q\right] \\ = \; & \mathrm{trace}\left[(I_m - \Lambda_1)^{-1}\,Q_{11}^T\Gamma_{1,1,n}(\beta_1)Q_{11}\right] + \mathrm{trace}\left[\Gamma_{2,2,n}(\beta_1)\right]. \end{aligned}$$

Hence, (18) becomes

$$\begin{aligned} & \frac{1}{n}\sum_{t=1}^{n}\ln\left[p_{t-1}(\Pi,\Theta,\beta_1,Q,\Lambda_1|\tau)\right] \\ = \; & \frac{k-m}{2}\ln(\tau) - \frac{k}{2}\ln(1+\tau) + \frac{1}{2}\sum_{j=1}^{m}\ln\left(\lambda_j+\tau\right) \\ & -\frac{1}{2}\sum_{j=1}^{m}\frac{q_j^T\Gamma_{1,1n}(\beta_1)q_j}{1-\lambda_j} - \frac{1}{2}\mathrm{trace}\left[\Gamma_{2,2,n}(\beta_1)\right], \end{aligned}$$

where now $Q_{11} = (q_1, ..., q_m)$.

Similarly as before, the optimal q_j's are equal to the orthonormal eigenvectors of $\Gamma_{1,1,n}(\beta_1)$ corresponding to its increasingly ordered eigenvalues. Thus, let now $\xi_{n,1}(\beta_1) \leq .... \leq \xi_{n,m}(\beta_1)$ be the eigenvalues of $\Gamma_{1,1,n}(\beta_1)$, and let $Q_{1,1,n}(\beta_1)$ be the corresponding orthogonal matrix of eigenvectors. Then

$$\begin{aligned} & \sup_{Q_{1,1},\Lambda_1,\lambda_{\max}[\Lambda_1]<1} \frac{1}{n}\sum_{t=1}^{n} \ln\left[p_{t-1}(\Pi,\Theta,\beta_1,Q,\Lambda_1|\tau)\right] \qquad (26) \\ = & \frac{k-m}{2}\ln(\tau) - \frac{k}{2}\ln(1+\tau) + \frac{1}{2}\sum_{j=1}^{m}\ln\left(\lambda_j(\beta_1,\tau)+\tau\right) \\ & -\frac{1}{2}\sum_{j=1}^{m}\frac{\xi_{n,j}(\beta_1)}{1-\lambda_j(\beta_1,\tau)} - \frac{1}{2}\text{trace}\left[\Gamma_{2,2,n}(\beta_1)\right] \end{aligned}$$

for $Q_{1,1} = Q_{1,1,n}(\beta_1)$ and $\Lambda_1 = \Lambda_{1,n}(\beta_1,\tau) = \text{diag}(\lambda_1(\beta_1,\tau),....,\lambda_m(\beta_1,\tau))$, where the $\lambda_j(\beta_1,\tau)$'s are the same as in (20). The optimal solution for Σ_1 given β_1 is now

$$\Sigma_{1,n}(\beta_1,\tau) = \left(\Theta_{11}^*\right)^{-1} Q_{1,1,n}(\beta_1)\Lambda_{1,n}(\beta_1,\tau)Q_{1,1,n}(\beta_1)^T \left(\Theta_{11}^*\right)^{-1}. \qquad (27)$$

4 The KPR model

I will now apply the above approach to a partially linearized version of the stochastic growth model of King, Plosser and Rebelo (KPR) (1988a,b), which is derived from Kydland and Prescott (1982) and Hansen (1985). The reasons for using this model is that it is a relatively simple real business cycle textbook model, and that I have used it before in Bierens and Swanson (2000).

4.1 The initial model

The stochastic version of the KPR model that I will consider takes the form:

KPR model 1:

$$\max\ E_0\left[\sum_{t=0}^{\infty}\lambda^t\left(\ln(C_t)+\theta\ln(1-N_t)\right)\right],\ \lambda<1, \qquad (28)$$

subject to

$$Q_t = C_t + I_t = A_t^{\alpha} K_t^{1-\alpha} N_t^{\alpha}, \tag{29}$$

$$K_t = (1-\delta)K_{t-1} + I_t, \tag{30}$$

$$\ln(A_t) = \gamma - v_t + \ln(A_{t-1}), \quad v_t \textit{ is i.i.d. } N(0, \sigma^2), \tag{31}$$

where C_t denotes consumption, Q_t is output, K_t is capital, N_t is employment, I_t is investment, and A_t is a technology index. The negative sign of v_t in (31) is harmless, of course, but has a notational advantage. Note that the log of the technology index A_t follows a Gaussian random walk, with drift equal to γ. Except for the technology index A_t, the variables in the above theoretical model may be interpreted as per capita aggregates.

Note that KPR model 1 is a special case of one of the models proposed in KPR (1988a). In particular, KPR use the Cobb-Douglas production function $Q_t = A_t K_t^{1-\alpha} (X_t N_t)^{\alpha}$, where X_t is the labor productivity index and A_t represents temporary changes in factor productivity. In (29) labor productivity changes are included in the technology index A_t. Moreover, the original KPR (1988a) model is a deterministic model: Their objective function is $\sum_{t=0}^{\infty} \lambda^t (\ln(C_t) + \theta \ln(1 - N_t))$ rather than $E_0 \left[\sum_{t=0}^{\infty} \lambda^t (\ln(C_t) + \theta \ln(1 - N_t))\right]$, so that they assume perfect foresight with respect to the technology index A_t. Therefore, KPR solve the problem

$$\begin{aligned} &\max \sum_{t=0}^{\infty} \lambda^t (\ln(C_t) + \theta \ln(1 - N_t)) \\ &\text{subject to} \\ &A_t^{\alpha} K_t^{1-\alpha} N_t^{\alpha} - C_t - K_t + (1-\delta)K_{t-1} = 0 \end{aligned} \tag{32}$$

via the first-order conditions of the Lagrange function

$$\begin{aligned} \mathcal{L} &= \sum_{t=0}^{\infty} \lambda^t (\ln(C_t) + \theta \ln(1 - N_t)) \\ &\quad + \sum_{t=0}^{\infty} \Lambda_t \left(A_t^{\alpha} K_t^{1-\alpha} N_t^{\alpha} - C_t - K_t + (1-\delta)K_{t-1}\right), \end{aligned} \tag{33}$$

where the Λ_t's are the Lagrange multipliers, together with the transversality condition $\lim_{t\to\infty} \Lambda_t K_t = 0$. In particular, KPR linearize the first-order conditions of the Lagrangian (33) around the steady state solution.

In this paper I will solve KPR model 1 analytically as far as I can, by reformulating the model and concentrating as much control and state variables out as possible, up to the point where the last equation, for the consumption-output ratio C_t/Q_t, can only be derived from the Bellman equation. In the process of concentrating variables out I will derive a deterministic nonlinear relation between N_t and C_t/Q_t. Upon arriving at the Bellman equation for C_t/Q_t I will partially linearized the concentrated model around the deterministic steady state solution. Finally, the solution for Q_t/A_t will be converted into a nonlinear ARMA(1,1) model for $\ln(Q_t/Q_{t-1})$, which then can be linearized as a stationary linear ARMA(1,1) model.

The resulting system of equations is quite different from the linearized version of the model in KPR (1988a,b). Admittedly, this linearization approach is specific to the KPR model. However, the econometric approach in this paper is applicable to any DSGE model that can be linearized such that the link between the parameters of the linear model and the deep parameters is preserved.

4.2 Reformulation of the KPR model

I will assume that once output is used as capital it is no longer fit for consumption. This assumption implies that consumption cannot exceed output, $0 \leq C_t \leq Q_t$, and neither can investment. Therefore, without loss of generality we may now replace the control variable C_t with $x_t Q_t$, where the consumption-output ratio $x_t = C_t/Q_t$ is the new control variable. The advantage is that x_t has non-stochastic bounds: $0 \leq x_t \leq 1$. Thus, denote

$$x_t = \frac{C_t}{Q_t},\ q_t = \frac{Q_t}{A_t},\ k_t = \frac{K_t}{A_t}.$$

Then (29) can be rewritten as

$$q_t = k_t^{1-\alpha} N_t^{\alpha}, \tag{34}$$

C_t is now equal to

$$C_t = x_t k_t^{1-\alpha} N_t^{\alpha} A_t,$$

and (30) can be rewritten as

$$\begin{aligned} k_t &= \frac{K_t}{A_t} = (1-\delta)\frac{K_{t-1}}{A_t} + \frac{I_t}{A_t} \\ &= (1-\delta)\frac{A_{t-1}}{A_t} k_{t-1} + (1-x_t)\, q_t \\ &= \exp\left[\ln(1-\delta) - \gamma + v_t\right] k_{t-1} + (1-x_t)\, q_t \\ &= \exp\left[\ln(1-\delta) - \gamma + v_t + \ln k_{t-1}\right] + (1-x_t)\, k_t^{1-\alpha} N_t^{\alpha}. \end{aligned} \tag{35}$$

KPR model 1 is now equivalent to:[4]

KPR model 2:

$$\max\ E_0\left[\sum_{t=0}^{\infty} \lambda^t \left((1-\alpha)\ln(k_t) + \ln(x_t) + \alpha\ln(N_t) + \theta\ln(1-N_t)\right)\right], \tag{36}$$

subject to

$$0 \leq x_t \leq 1,\ 0 \leq N_t \leq 1, \tag{37}$$

(34) *and* (35).

It follows from (35) that

$$\begin{aligned} k_t &= \exp\left(\ln(1-x_t) + \alpha\ln N_t\right) k_t^{1-\alpha} \\ &\quad + \exp\left(\ln(1-\delta) - \gamma + v_t + \ln k_{t-1}\right), \end{aligned} \tag{38}$$

which has a unique solution of the form

$$\begin{aligned} \ln k_t &= \ln g_\alpha\left(-\ln(1-x_t) - \alpha\ln N_t, \ln(1-\delta) - \gamma + v_t + \ln k_{t-1}\right) \\ &= \ln g_\alpha(z_t, y_{t-1}), \end{aligned} \tag{39}$$

say, where

$$z_t = -\ln(1-x_t) - \alpha\ln N_t \geq 0, \tag{40}$$

and

$$y_{t-1} = \ln(1-\delta) - \gamma + v_t + \ln k_{t-1}. \tag{41}$$

[4] Without loss of generality we may ignore the term $\sum_{t=0}^{\infty} \lambda^t \ln A_t$ in the objective function.

Thus, $g_\alpha(z_t, y_{t-1})$ is the fixed point solution of the equation

$$k_t = \exp\left(-z_t\right) k_t^{1-\alpha} + \exp(y_{t-1}), \tag{42}$$

i.e., $g_\alpha(z, y)$ corresponds to the intersection of the curve

$$f_1(k) = \exp\left(-z\right) k^{1-\alpha} + \exp(y),\ \ k \geq 0,$$

with the 45% line $f_2(k) = k$. This intersection is unique, and shifts up with $\exp\left(-z\right)$ and $\exp(y)$, hence $g_\alpha(z, y)$ is decreasing in z and increasing in y. In particular, it is not hard to verify that

$$\begin{aligned} \frac{\partial \ln g_\alpha(z,y)}{\partial y} &= \frac{\exp(y)}{\alpha g_\alpha(z,y) + (1-\alpha)\exp(y)} \in (0,1), \\ \frac{\partial \ln g_\alpha(z,y)}{\partial z} &= \frac{1}{\alpha}\left(\frac{\partial \ln g_\alpha(z,y)}{\partial y} - 1\right) < 0. \end{aligned} \tag{43}$$

Because by (40),

$$\ln x_t = \ln\left(N_t^\alpha - \exp\left(-z_t\right)\right) - \alpha \ln N_t, \tag{44}$$

KPR model 2 now becomes

KPR model 3:

$$\max\ E_0\left[\sum_{t=0}^{\infty} \lambda^t \left((1-\alpha)\ln g_\alpha(z_t, y_{t-1}) + \ln\left(N_t^\alpha - \exp\left(-z_t\right)\right) + \theta \ln(1-N_t)\right)\right] \tag{45}$$

subject to $z_t \geq 0$, $0 \leq N_t \leq 1$,

$$\begin{aligned} y_t &= \ln g_\alpha(z_t, y_{t-1}) + \ln(1-\delta) - \gamma + v_{t+1}, \\ t &\geq 0, \end{aligned} \tag{46}$$

and the initial condition

$$y_{-1} = \ln(1-\delta) - \gamma + v_0 + \ln k_{-1}. \tag{47}$$

4.3 The relation between employment and the consumption-output ratio

Given the optimal z_t's, the optimal solutions for the N_t's is now given by

$$\max_{N_t} \left\{ \ln\left(N_t^\alpha - \exp\left(-z_t\right)\right) + \theta \ln(1 - N_t) \right\},$$

with first order conditions

$$\frac{\alpha N_t^{\alpha-1}}{N_t^\alpha - \exp\left(-z_t\right)} = \frac{\theta}{1 - N_t}. \tag{48}$$

Substituting $\exp(-z_t) = (1 - x_t)N_t^\alpha$ [c.f. (40)] in (48) yields

$$N_t = 1/\left(1 + (\theta/\alpha)x_t\right). \tag{49}$$

Note that by $0 \le x_t \le 1$,

$$\frac{\alpha}{\alpha + \theta} \le N_t \le 1. \tag{50}$$

4.4 The law of motion of output

Due to (49), we can now write output as a function of x_t and lagged output, as follows. It follows from (40) and (49) that now

$$\begin{aligned} z_t &= \alpha \ln\left(1 + (\theta/\alpha)x_t\right) - \ln(1 - x_t) && (51) \\ &= h_{\alpha,\theta}(x_t), && (52) \end{aligned}$$

say, which is a monotonic increasing function of x_t:

$$h'_{\alpha,\theta}(x_t) = \frac{\alpha(1+\theta) + \theta(1-\alpha)x_t}{\left(1 - x_t\right)\left(\alpha + \theta x_t\right)} > 0. \tag{53}$$

Next, it follows from (29), and (49) that

$$Q_t = A_t^\alpha K_t^{1-\alpha} \left(1 + (\theta/\alpha)x_t\right)^{-\alpha}, \tag{54}$$

so that

$$K_t = Q_t^{1/(1-\alpha)} \left(1 + (\theta/\alpha)x_t\right)^{\alpha/(1-\alpha)} A_t^{-\alpha/(1-\alpha)}, \tag{55}$$

or in log form

$$\ln k_t = \ln(K_t/A_t) = \frac{\alpha}{1-\alpha} \ln\left(1 + (\theta/\alpha)x_t\right) + \frac{1}{1-\alpha} \ln(Q_t/A_t). \tag{56}$$

Furthermore, it follows from (30) and (55) that

$$\begin{aligned} K_t &= Q_t^{1/(1-\alpha)} \left(1+(\theta/\alpha)x_t\right)^{\alpha/(1-\alpha)} A_t^{-\alpha/(1-\alpha)} \\ &= (1-\delta)K_{t-1} + (1-x_t)Q_t \\ &= (1-\delta)Q_{t-1}^{1/(1-\alpha)} \left(1+(\theta/\alpha)x_{t-1}\right)^{\alpha/(1-\alpha)} A_{t-1}^{-\alpha/(1-\alpha)} \\ &\quad +(1-x_t)Q_t, \end{aligned}$$

hence

$$\begin{aligned} & (Q_t/A_t)^{1/(1-\alpha)} \left(1+(\theta/\alpha)x_t\right)^{\alpha/(1-\alpha)} - (1-x_t)Q_t/A_t \\ = \ & (1-\delta)(Q_{t-1}/A_{t-1})^{1/(1-\alpha)} \left(1+(\theta/\alpha)x_{t-1}\right)^{\alpha/(1-\alpha)} A_{t-1}A_t^{-1}. \end{aligned} \tag{57}$$

Given the right-hand side of (57) and x_t, we can solve $q_t = Q_t/A_t$ from (57). Hence the law of motion for output is given by (31), together with

$$\begin{aligned} \ln(Q_t/A_t) &= \eta_{\alpha,\theta}\left(x_t,\ \ln(1-\delta) - \gamma + \frac{\alpha}{1-\alpha}\ln\left(1+(\theta/\alpha)x_{t-1}\right)\right. \\ &\quad \left. +\frac{1}{1-\alpha}\ln(Q_{t-1}/A_{t-1}) + v_t\right), \end{aligned} \tag{58}$$

where $\exp(\eta_{\alpha,\theta}(x,y)) = q$ is the solution of the equation

$$q^{1/(1-\alpha)} \left(1+(\theta/\alpha)x\right)^{\alpha/(1-\alpha)} - (1-x)q = \exp(y), \tag{59}$$

i.e., given x and y, $q = \exp(\eta_{\alpha,\theta}(x,y))$ corresponds to the point of intersection of the curve $q^{1/(1-\alpha)} \left(1+(\theta/\alpha)x\right)^{\alpha/(1-\alpha)}$ with the line $(1-x)q + \exp(y)$.

Note that the partial derivatives of $\eta_{\alpha,\theta}(x,y)$ are:

$$\begin{aligned} \eta_{\alpha,\theta}^{(1)}(x,y) &= \frac{\partial\eta_{\alpha,\theta}(x,y)}{\partial x} = \frac{-1}{(1+(\theta/\alpha)x)} \\ &\quad - \frac{\left(\theta(1-x)+1\right)(1-\alpha)\exp(\eta_{\alpha,\theta}(x,y))}{\left(\exp(y) + \alpha(1-x)\exp(\eta_{\alpha,\theta}(x,y))\right)\left(1+(\theta/\alpha)x\right)} \\ &< -\alpha/\left(\alpha+\theta\right), \\ \eta_{\alpha,\theta}^{(2)}(x,y) &= \frac{\partial\eta_{\alpha,\theta}(x,y)}{\partial y} \\ &= \frac{(1-\alpha)\exp(y)}{\exp(y) + \alpha(1-x)\exp(\eta_{\alpha,\theta}(x,y))} \in (0, 1-\alpha)\,. \end{aligned} \tag{60}$$

The latter implies that

$$\partial\ln(Q_t/A_t)/\partial\ln(Q_{t-1}/A_{t-1}) \in (0,1). \tag{61}$$

4.5 Concentrating employment out

Plugging in $N_t = \alpha\left(\alpha+\theta x_t\right)^{-1}$ and (51) in $\ln\left(N_t^{\alpha}-\exp\left(-z_t\right)\right)+\theta\ln(1-N_t)$ yields

$$\begin{aligned} & \ln\left(N_t^{\alpha}-\exp\left(-z_t\right)\right)+\theta\ln(1-N_t) \\ = & \ \alpha\ln\alpha+\theta\ln\theta+(1+\theta)\ln\left(x_t\right)-(\alpha+\theta)\ln\left(\alpha+\theta x_t\right) \\ = & \ \alpha\ln\alpha+\theta\ln\theta+(1-\alpha)f_{\alpha,\theta}\left(x_t\right), \end{aligned}$$

say, where

$$f_{\alpha,\theta}\left(x_t\right)=\frac{(1+\theta)\ln\left(x_t\right)-(\alpha+\theta)\ln\left(\alpha+\theta x_t\right)}{1-\alpha}. \tag{62}$$

Note that $f_{\alpha,\theta}\left(x\right)$ is monotonic increasing:

$$f'_{\alpha,\theta}\left(x\right)=\frac{\alpha(1+\theta)+\theta(1-\alpha)x}{x\left(\alpha+\theta x\right)\left(1-\alpha\right)}>0. \tag{63}$$

Thus KPR model 3 now becomes:

KPR model 4:

$$\max\ E_0\sum_{t=0}^{\infty}\lambda^t\left[\ln g_{\alpha}\left(h_{\alpha,\theta}(x_t),y_{t-1}\right)+f_{\alpha,\theta}\left(x_t\right)\right] \tag{64}$$

subject to $0\le x_t\le 1$ *and*

$$\begin{aligned} y_t &= \ln g_{\alpha}\left(h_{\alpha,\theta}(x_t),\ y_{t-1}\right)+\ln(1-\delta)-\gamma+v_{t+1} \qquad (65)\\ t &\ge 0, \\ y_{-1} &= \frac{\alpha}{1-\alpha}\ln\left(1+(\theta/\alpha)x_{-1}\right)+\frac{1}{1-\alpha}\ln\left(Q_{-1}/A_{-1}\right) \\ &\quad +\ln(1-\delta)-\gamma+v_0 \\ v_t &\sim \textit{i.i.d. } N(0,\sigma^2) \end{aligned}$$

Note that it follows from (31), (56) and (66) that

$$\begin{aligned} y_t &= \ln(1-\delta)-\gamma+\ln(K_t/A_t)+v_{t+1} \\ &= \ln(1-\delta)-\gamma+\frac{\alpha}{1-\alpha}\ln\left(1+(\theta/\alpha)x_t\right) \\ &\quad +\frac{1}{1-\alpha}\ln\left(Q_t/A_t\right)+v_{t+1}. \end{aligned}$$

4.6 The Bellman equation of the concentrated model

As is well-known, KPR model 4 reads as: Find a *contingency plan* $x_t = \Psi(y_{t-1}) \in [0,1]$, where Ψ is a nonrandom Borel measurable function, such that (64) is maximized subject to (65). See Stokey, Lucas and Prescott (1989). Then $x_0 = \Psi(y_{-1})$ is the actual decision made by the economic agent at time $t = 0$, which applies to all t:

$$\begin{aligned} x_t &= \Psi(y_{t-1}) = \Psi\left(\ln(1-\delta) - \gamma + \ln(K_{t-1}/A_{t-1}) + v_t\right) \\ &= \Psi\left(\ln(1-\delta) - \gamma + \frac{\alpha}{1-\alpha}\ln\left(1 + (\theta/\alpha)x_{t-1}\right)\right. \\ &\quad \left. + \frac{1}{1-\alpha}\ln\left(Q_{t-1}/A_{t-1}\right) + v_t\right). \end{aligned} \tag{66}$$

The optimal contingency plan can in principle be derived from the value function, i.e., $x_0 = \Psi(y_{-1})$ maximizes the value function

$$\begin{aligned} & V(x_0|y_{-1}) \\ = {} & \ln g_\alpha\left(h_{\alpha,\theta}(x_0), y_{-1}\right) + f_{\alpha,\theta}(x_0) + \lambda E_0\left[\max_{0\le x_1\le 1} V(x_1|y_0)\right] \\ = {} & \ln g_\alpha\left(h_{\alpha,\theta}(x_0), y_{-1}\right) + f_{\alpha,\theta}(x_0) + \lambda \int_{-\infty}^{\infty} V\left(\Psi(z)|z\right) \\ & \times \frac{\exp\left[-\frac{1}{2}\left(z - \ln(1-\delta) + \gamma - \ln g_\alpha\left(h_{\alpha,\theta}(x_0), y_{-1}\right)\right)^2/\sigma^2\right]}{\sigma\sqrt{2\pi}} dz \end{aligned} \tag{67}$$

In practice this equality is too difficult to solve analytically, because $V(x_0|y_{-1})$ depends on Ψ, whereas Ψ has to be determined by maximizing $V(x_0|y_{-1})$. It is for this very reason that Kydland and Prescott (1982) propose an iterative procedure starting from an initial quadratic approximation of $V(x_0|y_{-1})$.

5 Linearization

I will now solve (64) subject to a linearized version of the state process (65). In particular, I will linearize (65) around the deterministic steady state.

Set v_{t+1} in (65) equal to zero for all t, and denote the corresponding variables in KPR model 4 by $\overline{x}_t$ and $\overline{y}_t$. Then the deterministic version of KPR model 4 is:

Deterministic KPR model 4:

$$\max \sum_{t=0}^{\infty} \lambda^t \left[\ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}_t), \overline{y}_{t-1}\right) + f_{\alpha,\theta}\left(\overline{x}_t\right)\right],$$

subject to $0 \leq \overline{x}_t \leq 1$ *and*

$$\overline{y}_t = \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}_t),\ \overline{y}_{t-1}\right) + \ln(1-\delta) - \gamma. \tag{68}$$

In the deterministic steady state, $\overline{x}_t \to \overline{x}$ as $t \to \infty$, because then C_t and Q_t will grow at the same exponential rate, and $\overline{y}_t \to \overline{y}$, where

$$\overline{y} = \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}),\ \overline{y}\right) + \ln(1-\delta) - \gamma. \tag{69}$$

In order to determine $\overline{x}_t$ and its limit $\overline{x}$ I will now linearize (68) around $(\overline{x}, \overline{y})$. Observe from (43) and (69) that

$$\begin{aligned}
\frac{\partial \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}),\ \overline{y}\right)}{\partial \overline{y}} &= \frac{\exp(\overline{y})}{\alpha g_\alpha \left(h_{\alpha,\theta}(\overline{x}),\ \overline{y}\right) + (1-\alpha)\exp(\overline{y})} \\
&= \frac{1-\delta}{\alpha \exp(\gamma) + (1-\alpha)(1-\delta)}, \\
\frac{\partial \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}),\ \overline{y}\right)}{\partial \overline{x}} &= \frac{h'_{\alpha,\theta}(\overline{x})}{\alpha} \left(\frac{\partial \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}),\ \overline{y}\right)}{\partial \overline{y}} - 1\right) \\
&= -h'_{\alpha,\theta}(\overline{x}) \left(\frac{\exp(\gamma) - (1-\delta)}{\alpha \exp(\gamma) + (1-\alpha)(1-\delta)}\right),
\end{aligned}$$

hence

$$\begin{aligned}
\overline{y}_t - \overline{y} &= \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}_t),\ \overline{y}_{t-1}\right) - \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}),\ \overline{y}\right) \\
&\approx \frac{\partial \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}),\ \overline{y}\right)}{\partial \overline{x}} (\overline{x}_t - \overline{x}) \\
&\quad + \frac{\partial \ln g_\alpha \left(h_{\alpha,\theta}(\overline{x}),\ \overline{y}\right)}{\partial \overline{y}} (\overline{y}_{t-1} - \overline{y})
\end{aligned}$$

$$\begin{aligned} &= \frac{1-\delta}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)}(\overline{y}_{t-1}-\overline{y}) \\ &\quad -h'_{\alpha,\theta}(\overline{x})\left(\frac{\exp(\gamma)-(1-\delta)}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)}\right)(\overline{x}_t-\overline{x}). \end{aligned}$$

Thus

$$\overline{y}_t \approx \beta_0+\beta_1\overline{y}_{t-1}-\beta_2\overline{x}_t, \tag{70}$$

where

$$\begin{aligned} \beta_0 &= \overline{y}+\gamma-\ln(1-\delta)-\frac{1-\delta}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)}\overline{y} \\ &\quad +\overline{x}.h'_{\alpha,\theta}(\overline{x})\left(\frac{\exp(\gamma)-(1-\delta)}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)}\right)\overline{x} \end{aligned}$$

$$\beta_1 = \frac{1-\delta}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)} \in (0,1-\delta) \tag{71}$$

$$\beta_2 = h'_{\alpha,\theta}(\overline{x})\left(\frac{\exp(\gamma)-(1-\delta)}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)}\right) > 0 \tag{72}$$

Replacing (68) by (70) in deterministic KPR model 4 now yields:

Linearized deterministic KPR model 4:

$$\max \sum_{t=0}^{\infty}\lambda^t\left[\overline{y}_t+f_{\alpha,\theta}(\overline{x}_t)\right] \tag{73}$$

subject to $0\leq\overline{x}_t\leq 1$ *and* $\overline{y}_t=\beta_0+\beta_1\overline{y}_{t-1}-\beta_2\overline{x}_t$.

Because

$$\overline{y}_t = \frac{1-\beta_1^{t+1}}{1-\beta_1}+\beta_1^{t+1}\overline{y}_{-1}-\beta_2\left(\sum_{j=0}^{t}\beta_1^j\overline{x}_{t-j}\right),$$

and

$$\sum_{t=0}^{\infty}\lambda^t\left[\sum_{j=0}^{t}\beta_1^j\overline{x}_{t-j}\right] = \frac{1}{1-\lambda\beta_1}\sum_{t=0}^{\infty}\lambda^t\overline{x}_t,$$

it follows that (73) is equivalent to

$$\max_{0\le\overline{x}_t\le1}\sum_{t=0}^{\infty}\lambda^t\left[f_{\alpha,\theta}\left(\overline{x}_t\right)-\frac{\beta_2}{1-\lambda\beta_1}\overline{x}_t\right].$$

Thus, each $\overline{x}_t$ is the solution of

$$\max_x\left[f_{\alpha,\theta}\left(x\right)-\frac{\beta_2}{1-\lambda\beta_1}x\right]. \tag{74}$$

and so is its limit $\overline{x}$.

Substituting (71) and (72) in (74) it follows that the first-order condition for (74) is

$$f'_{\alpha,\theta}\left(\overline{x}\right)=h'_{\alpha,\theta}(\overline{x})\left(\frac{\exp(\gamma)-(1-\delta)}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)-\lambda\left(1-\delta\right)}\right).$$

Hence it follows from (53) and (63) that

$$\frac{\overline{x}}{1-\overline{x}}=\frac{\alpha}{1-\alpha}+\frac{\left(1-\lambda\right)\left(1-\delta\right)}{\left(1-\alpha\right)\left(\exp(\gamma)-(1-\delta)\right)}. \tag{75}$$

Along the same lines we can linearize (65) around $(\overline{x},\overline{y})$, which yields:

Linearized KPR model 4:

$$\max\ E_0\sum_{t=0}^{\infty}\lambda^t\left[y_t+f_{\alpha,\theta}\left(x_t\right)\right]$$

subject to $0\le x_t\le 1$ *and*

$$y_t=\beta_0+\beta_1 y_{t-1}-\beta_2 x_t+v_{t+1}.$$

Also in this case x_t is the solution of (74), hence the solution involved is the same as (75):

$$\frac{C_t}{I_t}=\frac{x_t}{1-x_t}=\frac{\overline{x}}{1-\overline{x}}=\frac{\alpha}{1-\alpha}+\frac{\left(1-\lambda\right)\left(1-\delta\right)}{\left(1-\alpha\right)\left(\exp(\gamma)-(1-\delta)\right)}. \tag{76}$$

This solution, together with (31), (49) and (58), now completes the solution of the KPR model. Given (76), it follows from (58) and (61) that $\ln(Q_t/A_t)$ is a stationary nonlinear AR(1) process, hence $\ln(Q_t)$ and $\ln(A_t)$ are cointegrated, with cointegrating vector $(1,-1)^T$. However, $\ln(A_t)$ is not observable. Therefore, I will convert to law of motion of output into a law of motion for the output growth $\ln(Q_t/Q_{t-1})$.

5.1 A nonlinear ARMA(1,1) model for output growth

Assuming that the economic agent restricts his choice of contingency plans for x_t to a constant $\overline{x}$ given by (76), the complete model consists of the following four equations:

$$C_t/Q_t = \overline{x}, \tag{77}$$

$$N_t \equiv 1/\left(1+(\theta/\alpha)\overline{x}\right) = \nu, \tag{78}$$

say, and

$$\ln(Q_t/A_t) = \eta_{\alpha,\theta}\left(\overline{x},\ \mu + \frac{1}{1-\alpha}\ln(Q_{t-1}/A_{t-1}) + v_t\right), \tag{79}$$

where $\mu = \ln(1-\delta) - \gamma - (\alpha/(1-\alpha))\ln(\nu)$, together with equation (31), written as

$$\ln(Q_t/Q_{t-1}) = \ln(Q_t/A_t) - \ln(Q_{t-1}/A_{t-1}) - v_t + \gamma. \tag{80}$$

Recall from (58) that (79) now reads as

$$\begin{aligned} & (Q_t/A_t)^{1/(1-\alpha)}\nu^{-\alpha/(1-\alpha)} - (1-\overline{x})(Q_t/A_t) \\ = \ & (1-\delta)(Q_{t-1}/A_{t-1})^{1/(1-\alpha)}\nu^{-\alpha/(1-\alpha)}\exp(v_t-\gamma). \end{aligned} \tag{81}$$

The equations (80) and (81) can be combined into an equation for Q_t/Q_{t-1} only, as follows. Substitute

$$(Q_{t-1}/A_{t-1}) = (Q_t/A_t)(Q_t/Q_{t-1})^{-1}\exp(-v_t+\gamma) \tag{82}$$

in (81), and solve for Q_t/A_t, i.e.,

$$\begin{aligned} & Q_t/A_t = \nu\exp(v_t-\gamma)(Q_t/Q_{t-1})^{1/\alpha} \\ & \times\left[\frac{1-\overline{x}}{(Q_t/Q_{t-1})^{1/(1-\alpha)}(\exp(v_t-\gamma))^{\alpha/(1-\alpha)} - 1 + \delta}\right]^{(1-\alpha)/\alpha}. \end{aligned} \tag{83}$$

This result implies that

$$P[(Q_t/Q_{t-1})^{1/(1-\alpha)}(\exp(v_t-\gamma))^{\alpha/(1-\alpha)}-1+\delta>0]=1,$$

hence

$$P\left[\ln(Q_t/Q_{t-1})>(1-\alpha)\ln(1-\delta)+\alpha\gamma-\alpha v_t\right]=1, \tag{84}$$

and consequently

$$\ln(Q_t/Q_{t-1})=(1-\alpha)\ln(1-\delta)+\alpha\gamma-\alpha v_t+w_t, \tag{85}$$

where

$$P[w_t>0]=1.$$

Substituting (85) in (83) now yields

$$\begin{aligned}\ln(Q_t/A_t)&=\frac{1-\alpha}{\alpha}\left[\ln(1-\overline{x})\right]+\ln(\nu)\\ &+w_t/\alpha-\frac{1-\alpha}{\alpha}\ln\left[\exp(w_t/(1-\alpha))-1\right],\end{aligned} \tag{86}$$

and combining (80), (85) and (86) yields

$$\begin{aligned}&\frac{\exp(w_t)}{\exp(w_t/(1-\alpha))-1}\\ =\ &\frac{\exp(w_{t-1}/(1-\alpha))}{\exp(w_{t-1}/(1-\alpha))-1}\times(1-\delta)^{\alpha}\exp\left[\alpha(v_t-\gamma)\right].\end{aligned} \tag{87}$$

This equation can be solved as

$$w_t=\varphi_{\alpha,\delta,\gamma}(w_{t-1},v_t), \tag{88}$$

say. Hence

$$\begin{aligned}&\ln(Q_t/Q_{t-1})\\ =\ &(1-\alpha)\ln(1-\delta)+\alpha\gamma-\alpha v_t\\ &+\varphi_{\alpha,\delta,\gamma}(\ln(Q_{t-1}/Q_{t-2})-(1-\alpha)\ln(1-\delta)-\alpha\gamma+\alpha v_{t-1},v_t)\\ =\ &\Phi_{\alpha,\delta,\gamma}\left(\ln(Q_{t-1}/Q_{t-2}),v_t,v_{t-1}\right),\end{aligned} \tag{89}$$

say, which is a nonlinear ARMA(1,1) process.

Note that I have implicitly assumed that $\delta < 1$, because if $\delta \uparrow 1$ then it follows from (85) that $w_t \to \infty$. In the case $\delta = 1$ equation (81) becomes

$$(Q_t/A_t)^{1/(1-\alpha)}\nu^{-\alpha/(1-\alpha)} = (1-\overline{x})(Q_t/A_t) \tag{90}$$

which can be rewritten as

$$\ln(Q_t/A_t) = \frac{1-\alpha}{\alpha}\ln(1-\overline{x}) + \ln(\nu), \tag{91}$$

which is the limit of (86) for $w_t \to \infty$. It follows now from (91) and (80) that in the case $\delta = 1$,

$$\ln(Q_t/Q_{t-1}) = \gamma - v_t.$$

5.2 A linearized ARMA(1,1) model for output growth.

The nonlinear ARMA(1,1) model for $\ln(Q_t/Q_{t-1})$ is pretty intractable. Therefore, I will now linearize this model around the deterministic steady state for w_t, as follows. Replace v_t in (87) by 0, and let the resulting steady state solution be $\overline{w}$. It it easy to verify that

$$\overline{w} = (1-\alpha)(\gamma - \ln(1-\delta)). \tag{92}$$

After some tedious calculations it follows that

$$\begin{aligned} & \left.\frac{\partial\varphi_{\alpha,\delta,\gamma}(w_{t-1},v_t)}{\partial w_{t-1}}\right|_{w_{t-1}=\overline{w},v_t=0} \\ = & \frac{1-\delta}{\alpha\exp\gamma + (1-\delta)(1-\alpha)} \end{aligned} \tag{93}$$

and

$$\begin{aligned} & \left.\frac{\partial\varphi_{\alpha,\delta,\gamma}(w_{t-1},v_t)}{\partial v_t}\right|_{w_{t-1}=\overline{w},v_t=0} \\ = & \frac{-\alpha(1-\alpha)(\exp\gamma - (1-\delta))}{\alpha\exp\gamma + (1-\delta)(1-\alpha)}. \end{aligned} \tag{94}$$

Therefore, we have approximately,

$$\begin{aligned} w_t \approx & \left(1 - \frac{1-\delta}{\alpha\exp(\gamma) + (1-\delta)(1-\alpha)}\right)(1-\alpha)(\gamma - \ln(1-\delta)) \\ & + \frac{1-\delta}{\alpha\exp\gamma + (1-\delta)(1-\alpha)}w_{t-1} - \frac{\alpha(1-\alpha)(\exp(\gamma) - (1-\delta))}{\alpha\exp(\gamma) + (1-\alpha)(1-\delta)}v_t. \end{aligned} \tag{95}$$

Substituting (85) in (95) it follows, after some tedious but straightforward manipulations, that approximately

$$\begin{aligned}
\Delta \ln(Q_t) &= \ln(Q_t/Q_{t-1}) \\
&= \alpha\gamma\left(\frac{\exp\gamma-(1-\delta)}{\alpha\exp\gamma+(1-\delta)(1-\alpha)}\right) \\
&\quad+\left(\frac{1-\delta}{\alpha\exp\gamma+(1-\delta)(1-\alpha)}\right)\ln(Q_{t-1}/Q_{t-2}) \\
&\quad+\left(\frac{-\alpha\exp(\gamma)}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)}\right)v_t \\
&\quad-\left(\frac{1-\delta}{\exp(\gamma)}\right)\left(\frac{-\alpha\exp(\gamma)}{\alpha\exp\gamma+(1-\delta)(1-\alpha)}\right)v_{t-1}
\end{aligned} \tag{96}$$

It is convenient to reparametrize this ARMA(1,1) model as

$$\Delta\ln(Q_t)=\kappa\Delta\ln(Q_{t-1})+\xi+\varepsilon_t-\varsigma\varepsilon_{t-1}, \tag{97}$$

where

$$\kappa=\frac{1-\delta}{\alpha\exp\gamma+(1-\alpha)(1-\delta)}, \tag{98}$$

$$\xi=\alpha\gamma\left[\frac{\exp\gamma-(1-\delta)}{\alpha\exp\gamma+(1-\delta)(1-\alpha)}\right]=\gamma(1-\kappa), \tag{99}$$

$$\varsigma=\frac{1-\delta}{\exp(\gamma)}<\kappa \tag{100}$$

and

$$\varepsilon_t=-\left(\frac{\alpha\exp(\gamma)}{\alpha\exp(\gamma)+(1-\alpha)(1-\delta)}\right)v_t. \tag{101}$$

Note that if $\gamma>0$ then $\exp\gamma-(1-\delta)>0$, hence $\kappa<1$. Moreover, note that the value of x_t does no longer play a role in (97), because the linearization involved employs the assumption that x_t is constant.

6 The data

The data for C_t and I_t is the same as used by Watson (1993), although the sample period has been updated through 1994:4, so that we have 188 quarterly US observations, from 1948:1 to 1994:4, on C_t = per capita total consumption expenditures in1987 dollars, and I_t = per capita total fixed investment in 1987 dollars. As output I will use $Q_t = C_t + I_t$. KPR (1988a,b), Watson (1993) and Bierens and Swanson (2000) use a different Q_t, namely per capita GNP less government purchases of goods and services in1987 dollars, but in that case C_t and I_t do not add up exactly to Q_t. These authors use as employment N_t the per capita total labor hours in private nonagricultural establishments. However, the KPR model assumes $0 < N_t < 1$. Because there is no clear upper bound to the per capita total labor hours, it will not be possible to identify the parameter θ in (78) if we use this variable. Therefore, in the empirical application I will focus on output growth and the consumption-output ratio only.

Table 1 provides some descriptive data statistics of the model variables.

Table 1: Data statistics

Variable	*Minimum*	*Maximum*	*Mean*	*Median*
$x_t = C_t/Q_t$	0.8117	0.8523	0.8315	0.8319
$\Delta\ln(Q_t)$	-0.0399	0.0518	0.0049	0.0055

We see from Table 1 that there is little variation in $x_t = C_t/Q_t$. Therefore, the linearized solution (76) for $x_t = C_t/Q_t$ reflects the stylized fact that in reality $x_t = C_t/Q_t$ is approximately constant.

I have subjected the time series $\ln(C_t)$, $\ln(I_t)$, and $\ln(x_t/(1-x_t)) = \ln(C_t) - \ln(I_t)$ [5] to a variety of unit root and stationarity tests[6], in particular the ADF test [c.f. Fuller. (1996)] with lag length determined by the maximum of the Hannan-Quinn and Schwarz information criteria, the Phillips-Perron (1988) test, the Breitung (2002) test, the Bierens (1993) unit root test on the basis of higher order autocorrelations, the Bierens-Guo (1993) stationarity test, and the KPSS (1992) stationarity test. It appears that $\ln(C_t)$ and $\ln(I_t)$ are unit root with drift processes and that $\ln(C_t) - \ln(I_t)$ is stationary, hence $\ln(C_t)$ and $\ln(I_t)$ are cointegrated.

[5] The reason for the transformation $\ln(x_t/(1-x_t))$ is to make the variable involved unbounded, because a bounded time series cannot be a unit root process.

[6] These tests have been conducted using *EasyReg International* [Bierens (2003)] using the *EasyReg* default setting for the truncation lags and other test parameters.

7 The linearized theoretical model

Recall that (76) predicts that

$$\ln(C_t) - \ln(I_t) = \omega, \tag{102}$$

say, where

$$\omega = \ln\left(\frac{\alpha}{1-\alpha} + \frac{(1-\lambda)(1-\delta)}{(1-\alpha)(\exp(\gamma)-(1-\delta))}\right), \tag{103}$$

hence $\ln(C_t)$ and $\ln(I_t)$ are cointegrated without error, and the same holds of course for $\ln(C_t)$ and $\ln(Q_t)$. Consequently, model (97) also applies to $\Delta\ln(C_t)$ and $\Delta\ln(I_t)$ with the same parameters and error term. Therefore my linearized version of the KPR model is determined by the following two equations:

$$\begin{aligned} \Delta\ln(C_t) &= \kappa\Delta\ln(C_{t-1}) + \xi + \varepsilon_t - \varsigma\varepsilon_{t-1}, & (104)\\ \ln(C_t) - \ln(I_t) &= \omega. & (105) \end{aligned}$$

Note that the deep parameter $\alpha, \gamma, \delta, \lambda$ are identified from $\kappa, \xi, \varsigma, \omega$. Thus, the vector time series process I will work with is:

$$Y_t = \begin{pmatrix} \Delta\ln(C_t) \\ \ln(C_t) - \ln(I_t) \end{pmatrix},$$

which is assumed to be observable for $t = 0, 1, 2, ..., n$. The linearized DSGE model (10) now corresponds to

$$\begin{aligned} \mu_{t-1}(\beta_1) &= \begin{pmatrix} \kappa\Delta\ln(C_{t-1}) + \xi - \varsigma r_{t-1} \\ \omega \end{pmatrix} & (106)\\ \Sigma &= \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

where

$$\beta_1 = (\alpha, \delta, \lambda, \gamma)^T,$$

with κ, ξ, ς defined by (98), (99) and (100), respectively, ω by (103), $\sigma_\varepsilon^2 = \mathrm{var}(\varepsilon_t)$, and r_t is defined recursively by

$$\begin{aligned} r_t &= \varsigma r_{t-1} + \Delta\ln(C_t) - \kappa\Delta\ln(C_{t-1}) - \xi \text{ for } t \geq 1, \\ r_t &= 0 \text{ for } t \leq 0. \end{aligned}$$

If one would assume that the ARMA(1,1) model for $\Delta \ln(C_t)$ in (104) represents the data generating process then it is justified to estimate the parameters κ, ς and σ_ε^2 by maximum likelihood, and the ML estimates of κ and ς can then be used to estimate γ, α and δ:

Table 2: ML estimation results for $\Delta \ln(C_t)$

parameters	*estimates*	*t-value*
κ	0.075000	0.027
ς	0.049000	0.017
$\gamma = \xi/(1-\kappa)$	0.004936	7.193
$\alpha = \left(\frac{1}{\kappa} - 1\right)\frac{\varsigma}{1-\varsigma}$	0.635471	
$\delta = 1 - \varsigma \exp(\gamma)$	0.950758	
$\sigma_\varepsilon^2 =$ 0.0000833569	$R^2 =$	0.0007

Although the estimated value of α is close to the usual calibrated value (0.65), the estimated value of the capital depreciation rate δ is extremely high. Moreover, replacing α, δ and γ in (103) by their estimated values, and ω by the sample mean 1.597920 of $\ln(C_t) - \ln(I_t)$ yield $\lambda = -21.63578$, which is an impossible value.

Of course, the estimation approach proposed in this paper does not require that the theoretical model represents the data generating process, and due to the singularity of model (105) it is unrealistic to assume that it does. Therefore, the estimation results in Table 2 are only preliminary, and merely serve as an illustration of what happens if the model is taken directly to the data.

8 The econometric model

The starting point for the specification of an econometric model is a vector error correction model (VECM) for

$$Y_t^{**} = \begin{pmatrix} \ln(C_t) \\ \ln(I_t) \end{pmatrix}.$$

The Johansen cointegration analyses [Johansen (1988, 1991, 1994), Johansen and Juselius (1990)] indicates that Y_t^{**} is cointegrated, with cointegrating vector $(1,-1)^T$ and VECM order 3:

$$\Delta Y_t^{**} = \pi_0 + \beta(1,-1)Y_{t-1}^{**} + \Pi_1 \Delta Y_{t-1}^{**} + \Pi_2 \Delta Y_{t-2}^{**} + V_t^*.$$

It is easy to verify that this model can be written as a VAR(2) model for

$$Y_t = \begin{pmatrix} \Delta \ln C_t \\ \ln(C_t/I_t) \end{pmatrix}.$$

Thus, the conditional expectation η_{t-1} of the econometric model (11) takes the form

$$\eta_{t-1} = \begin{pmatrix} \varphi_1 \Delta \ln C_{t-1} + \varphi_2 \ln(C_{t-1}/I_{t-1}) + \varphi_3 \Delta \ln C_{t-2} + \varphi_4 \ln(C_{t-2}/I_{t-2}) + \varphi_5 \\ \psi_1 \Delta \ln C_{t-1} + \psi_2 \ln(C_{t-1}/I_{t-1}) + \psi_3 \Delta \ln C_{t-2} + \psi_4 \ln(C_{t-2}/I_{t-2}) + \psi_5 \end{pmatrix}.$$

The maximum likelihood estimation results involved are given in Table 3:

Table 3: Estimation result for the econometric model

	$\Delta \ln(C_t)$		$\ln(C_t/I_t)$	
	estimates	*t-values*	*estimates*	*t-values*
$\Delta \ln(C_{t-1})$	-0.049888	-0.681	-0.681538	-4.081
$\ln(C_{t-1}/I_{t-1})$	-0.040625	-1.293	1.283050	17.919
$\Delta \ln(C_{t-2})$	0.256873	3.425	-0.119271	-0.698
$\ln(C_{t-2}/I_{t-2})$	0.067755	2.165	-0.375238	-5.262
1	-0.039430	-2.517	0.151406	4.242
R^2	0.140		0.911	

$$\Theta = \begin{pmatrix} 7.10750955\text{E-}5 & -4.34092710\text{E-}5 \\ -4.34092710\text{E-}5 & 3.69134238\text{E-}4 \end{pmatrix}$$

Note that the parameters of $\Delta \ln(C_{t-1})$ and $\ln(C_{t-1}/I_{t-1})$ in the model for $\Delta \ln(C_t)$ are not significant, and neither are they jointly at any conventional significance level, as appears from the Wald test involved.

In Figures 2-5 the responses of $\Delta \ln(C_t)$ and $\ln(C_t/I_t)$ to unit shocks in the innovations of these variables are presented, based on the upper-triangular Cholesky decomposition of Θ, so that the shocks are imposed in the order $\ln(C_t/I_t)$, $\Delta \ln(C_t)$. The dots represent the asymptotic one- and two-times standard error bands, computed on the basis of Baillie's (1987) approach.

Figures 2 and 3 confirm that the lagged $\ln(C_t/I_t)$ and $\Delta \ln(C_t)$ do not have much effect on $\Delta \ln(C_t)$, as suggested by the low R^2. Moreover, Figures 4 and 5 confirm once more that the theoretical model solution $\ln(C_t/I_t) \equiv \omega$ is quite unrealistic.

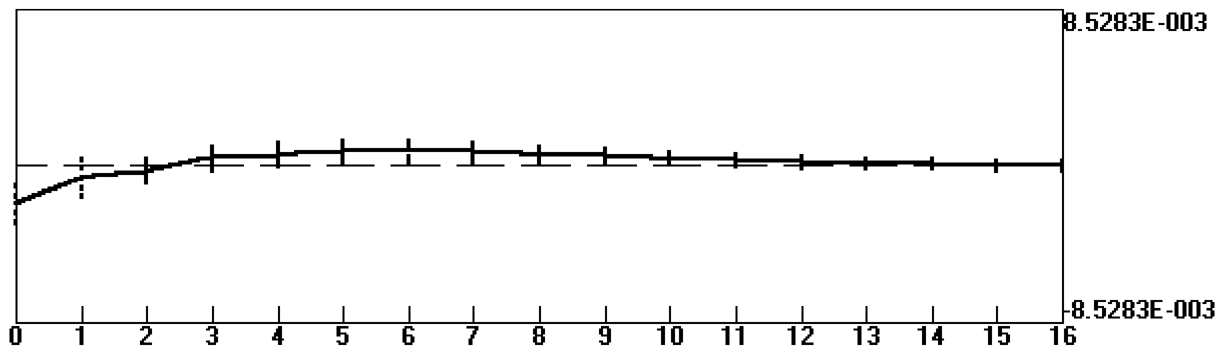

Figure 2: Innovation response of $\Delta \ln(C_t)$ to a unit shock in $\ln(C_t/I_t)$.

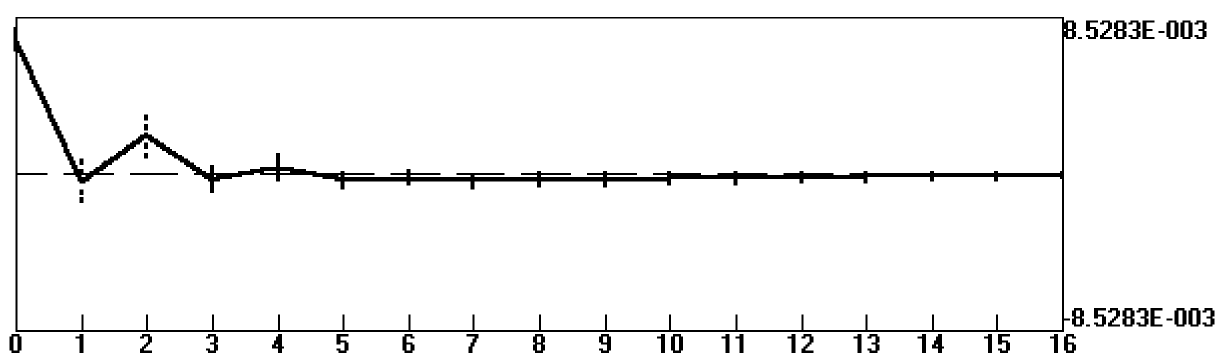

Figure 3: Innovation response of $\Delta \ln(C_t)$ to a unit shock in $\Delta \ln(C_t)$.

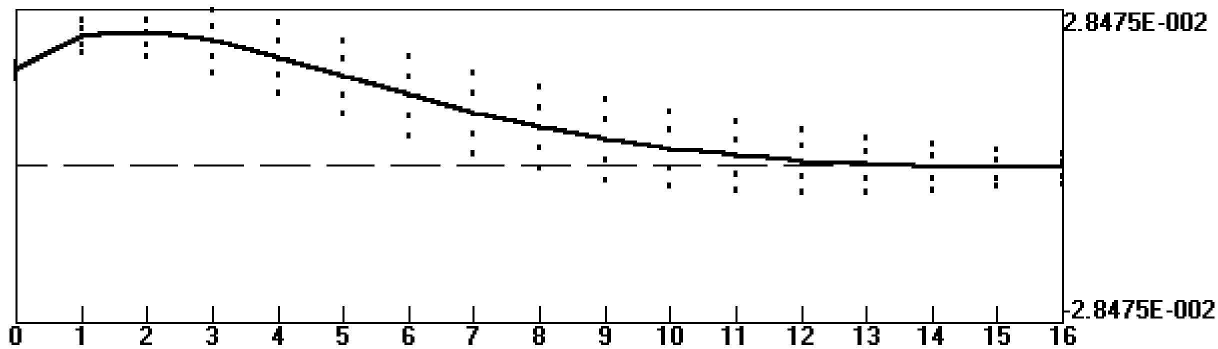

Figure 4: Innovation response of $\ln(C_t/I_t)$ to a unit shock in $\ln(C_t/I_t)$.

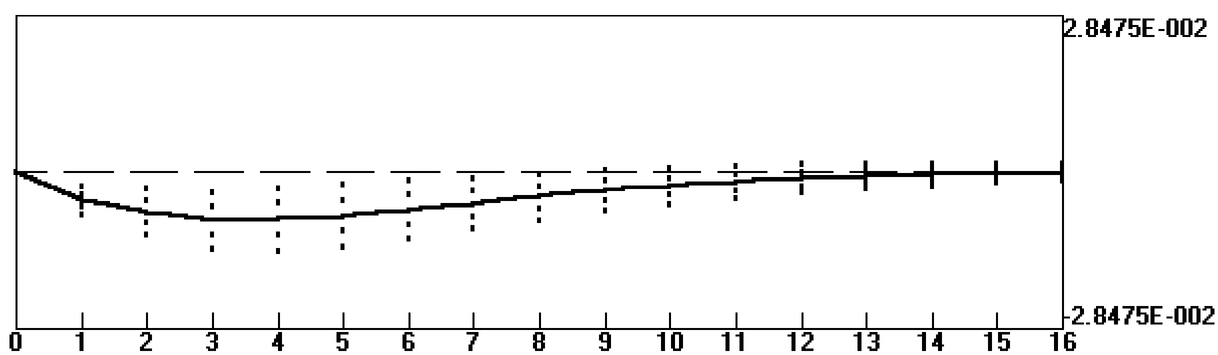

Figure 5: Innovation response of $\ln(C_t/I_t)$ to a unit shock in $\Delta\ln(C_t)$.

9 Re-estimation of the KPR model

Given the estimates in Table 3, I have maximized (26) to $\beta_1 = (\alpha, \delta, \lambda, \gamma)^T$ over the parameter space $\mathcal{B} = [0,1] \times [0,1] \times [0,1] \times [0, 0.1]$, using the simplex method of Nelder and Mead (1965), for $\tau = 0.1$. Although one may limit the parameter space to an area of "acceptable" values, acceptable in the sense that the parameter space corresponds to values that theorists expect and find acceptable as calibrated values, I have left the deep parameters free in order to let the data speak. Only the deterministic growth rate γ has been restricted to the interval $[0, 0.1]$, but this is done for numerical reasons. The choice of $\tau = 0.1$ is arbitrary, of course. Due to the restriction (21) we cannot choose τ too large, but what is too large depends on the optimal β_1. Anyhow, I have experimented with smaller values of τ, and the estimation results appears to be about the same as for $\tau = 0.1$.

Because (26) appears to have quite a few local maxima on $\mathcal{B}$, the simplex iteration has been restarted ten time from the last solution, and then this procedure has been repeated starting from random drawing from the uniform distribution on $\mathcal{B}$. After running for a few hours the solution with the largest value of (26) has been chosen. The restriction of the solution to $\mathcal{B}$ has been enforced by assigning a the value -10^{29} to the objective function for values outside $\mathcal{B}$, and the restriction (21) has been enforced in the same way. The results are presented in Table 4, for the deep parameters as well as for the corresponding parameters of model (104), together with the 95% confidence intervals of the deep parameter. The standard errors on which the confidence intervals are based have been derived from the (10×10) variance matrix of the estimated parameters of the VAR(2) model in Table 3, and the (numerically computed) 4×10 matrix of derivatives of the deep parameters $\alpha, \delta, \lambda, \gamma$ to the parameters of the VAR(2) model, using the well-known delta method. The confidence intervals have been modified by taking intersections with the interval [0,1]. The estimate of σ_ε^2 (the variance of ε_t in (97)) is based on (27), with $\tau = 0.1$, and the estimate of the deep parameter σ^2 (the variance of v_t in (31)) has been derived from (101).

Table 4: Estimation results for $\tau = 0.1$

Deep parameters [95% confidence intervals]			Model (104)-(105)	
α	0.831885	$[0.831062, 0.832707]$	κ	0.0000030
δ	0.999997	$[0.999397, 1]$	ς	0.0000025
λ	0.039859	$[0, 1]$	ξ	0.0049482
γ	0.004948	$[0.002823, 0.007073]$	ω	1.5990481
σ^2	0.000027		σ_ε^2	0.0000273

10 Discussion

The most striking result in Table 4 is the value of the capital depreciation rate δ, which is almost equal to its upper bound 1, with very narrow 95% confidence interval. Therefore, the time preference parameter λ is in practice no longer identified from (103), which is reflected by the corresponding 95% confidence interval. Also, the estimate of α is higher than the usually calibrated value (0.65), but its 95% confidence interval is very narrow. On the other hand, observe from Table 1 that the estimated value of α is very close to the sample mean 0.8315 of the consumption-output ratio C_t/Q_t. If one would interpret C_t as the cost of the production factor labor and $Q_t - C_t = I_t$ as profit, this result corresponds to short-run profit maximization.

The estimated value of γ implies an annual deterministic growth rate of about 2%, which is not unreasonable. Note that the implied value of ω is very close to the sample mean 1.597920 of $\ln(C_t/I_t)$, as should be. Because the AR and MA parameters κ and ς are very close to zero[7], the theoretical process for $\Delta \ln C_t$ is approximately white noise. This was already apparent from the preliminary estimation results in Table 2, because the ML estimates of κ and ς were insignificant, and from the innovation responses in Figures 3 and 4. Moreover, observe from Table 3 that also the fit of the econometric model for $\Delta \ln C_t$ is very low. Furthermore, note that in the case $\delta = 1$ the linearized solution (104) of the KPR model is the exact solution, i.e.,

$$\Delta \ln(C_t) = \gamma - v_t,\ C_t/Q_t = \alpha.$$

The KPR model describes a Robinson Crusoe[8]-type economy (before the

[7] Therefore, κ and ς do no longer contribute to the identification of α, but instead α is now almost entirely indentified by (103).

[8] Daniel Defoe (1719).

arrival of Friday), where Robinson has to decide how much of his harvest of say potatoes to eat, how much to plant for the next harvest, and how much to work in the fields, in order to maximize his lifetime utility. Because the potatoes he plants yield only one harvest, in this economy the depreciation rate of capital (potatoes) is total: $\delta = 1$.

The results in Table 4 indicate that, given the structure of the KPR model and its linearization, this Robinson Crusoe economy is the best fit for the US economy. Theorists may find this conclusion a strong argument in favor of calibration. However, the fact that the estimated deep parameters are incredible indicates that the KPR model in its present form is of limited use in explaining economic growth and business cycles in a real economy, and using calibrated "credible" parameters moves the model even farther away from reality.

Note that my linearization procedure does not hinge on the assumption that the random variables v_t in (31) are i.i.d. $N(0, \sigma^2)$. The latter assumption is only used in the Bellman equation (67), but this equation itself has not been used. Therefore, it is possible to generalize the theoretical ARMA(1,1) model for $\Delta\ln(C_t)$ by assuming a more general model for the v_t's. It is even possible to specify the distribution of v_t such that the theoretical model (105) can almost match the corresponding econometric model, namely if $v_t = (\varsigma - \kappa)v_{t-1} + \kappa\varsigma v_{t-2} + u_t$, where u_t is i.i.d. $N(0, \sigma^2)$. Then $\varepsilon_t - \varsigma\varepsilon_{t-1} + \kappa(\varepsilon_{t-1} - \varsigma\varepsilon_{t-2}) = e_t$, where e_t is i.i.d. $N(0, \sigma_e^2)$, hence $\Delta\ln(C_t) = \kappa^2\Delta\ln(C_{t-2}) + (1 + \kappa)\xi + e_t$. Moreover, using (102) the latter model can be written as

$$\Delta\ln(C_t) = \kappa^2\Delta\ln(C_{t-2}) + \varphi\ln(C_{t-2}/I_{t-2}) + (1+\kappa)\xi - \varphi\omega + e_t, \qquad (107)$$

which is close to the corresponding econometric model because the parameters of $\Delta\ln(C_{t-1})$ and $\ln(C_{t-1}/I_{t-1})$ in the econometric model for $\Delta\ln(C_t)$ in Table 3 were not jointly significant. However, only the growth rate γ can be identified from (107) and (105), via (99), because given γ only the equations (98) and (103) are available to determine the parameters α, δ and λ.

11 Conclusion

In the current literature on econometric analysis of DSGE models it is implicitly or explicitly assumed, eventually after adding noise to eliminate singularity, that the model represents the data generating process, and that then the model parameters can be estimated by linking the model to the data via maximum likelihood, GMM or other estimation procedures. In this paper I adopt the theorist's view that these models are misspecified as representatives of data generating processes. Instead of linking a DSGE model directly to the data, I propose to link it indirectly to the data via an econometric model which is assumed to represent the data generating process. In doing so I can estimate the deep parameters as function of the parameters of the econometric model, without worrying about misspecification of the DSGE model. Moreover, via the delta method the estimated deep parameters inherit the asymptotic normality of the estimated parameters of the econometric model, so that the former estimates can be endowed with confidence intervals.

The estimation approach in this paper is applicable to any linearized DSGE model for which the link between the parameters of the linearized model and the deep parameters is preserved. For example, the DSGE models considered by Corradi and Swanson (2004) are linearized using iterated quadratic approximations of the value function, which preserves the link with the deep parameters, and can therefore be estimated by my approach. The same applies to the linearized DSGE model considered by Ireland (2003), which is derived from the model restrictions and an Euler equation. Of course, alternative linearization procedures may yield different estimates of the deep parameters.

As indicated before, by calibrating DSGE models theorists will limit their ability to detect model failure. In particular the extent of deviation of the estimated deep parameters from the usual calibrated values, as in the KPR case, provides useful information about possible model failure, and could (or should!) lead to a quest for more realistic models. This paper provides new econometric tools to assist in this endeavor.

Acknowledgments

Previous versions of this paper have been presented at Tilburg University, Penn State University, the University of Rochester, Pittsburgh University, the Universidade Federal do Ceara, Brazil, the Econometric Society European Meeting 2003, Stockholm, the University of Kansas, and the University of Guelph. The comments of William Barnett, Peter Ireland, Gustavo Ventura and a referee are gratefully acknowledged.

References

Ambler, S., A. Guay and L. Phaneuf, 2003, Labor Market Friction and Endogenous Business Cycle Propagation (paper presented at ESEM 2003, Stockholm).

Baillie, R. T., 1987, Inference in Dynamic Models Containing 'Surprise' Variables, Journal of Econometrics 35, 101-117.

Bierens, H. J., 1993, Higher Order Autocorrelations and the Unit Root Hypothesis, Journal of Econometrics 57, 137-160.

Bierens, H. J., 2003, EasyReg International, Pennsylvania State University (http://econ.la.psu.edu/~hbierens/EASYREG.HTM).

Bierens, H. J. and S. Guo, 1993, Testing Stationarity and Trend Stationarity Against the Unit Root Hypothesis, Econometric Reviews 12,1-32.

Bierens, H. J., and N. R. Swanson, 2000, The Econometric Consequences of The Ceteris Paribus Condition in Economic Theory, Journal of Econometrics 95, 223-253.

Breitung, J., 2002, Nonparametric Tests for Unit Roots and Cointegration., Journal of Econometrics 108, 343-364.

Corradi, V., and N. R. Swanson, 2004, Evaluation of Dynamic Stochastic General Equilibrium Models Based on Distributional Comparison of Simulated and Historical Data, Journal of Econometrics (forthcoming).

Defoe, D., 1719, The Life and Strange and Surprising Adventures of Robinson Crusoe.

DeJong, D. N., B. F. Ingram, and C. H. Whiteman, 1996, A Bayesian Approach to Calibration, Journal of Business and Economic Statistics 14, 1-9.

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

Feve, P., and F. Langot, 1994, The RBC Models Through Statistical Inference: An Application with French Data, Journal of Applied Econometrics 9, 11-35.

Fuller, W. A., 1996, Introduction to Statistical Time Series (John Wiley, New York, NY)

Geweke, J., 1999, Computational Experiments and Reality, Working paper (http://www.biz.uiowa.edu/faculty/jgeweke/papers/paperE/cer.pdf).

Hansen, G. D., 1985, Indivisible Labor and the Business Cycle, Journal of Monetary Economics 16, 309-327.

Hansen, L. P., and J. J. Heckman, 1996, The Empirical Foundations of Calibration, Journal of Economic Perspectives 10, 87-104.

Ireland, P. N., 2003, A Method for Taking Models to the Data, Journal of Economic Dynamics and Control (forthcoming).

Johansen, S., 1988, Statistical Analysis of Cointegrated Vectors, Journal of Economic Dynamics and Control 12, 231-254.

Johansen, S., 1991, Estimation and Hypothesis Testing of Cointegrated Vectors in Gaussian Vector Autoregressive Models, Econometrica 59, 1551-1580.

Johansen, S., 1994, The Role of the Constant and Linear Terms in Cointegration Analysis of Nonstationary Variables, Econometric Reviews 13, 205-229.

Johansen, S. and K. Juselius, 1990, Maximum Likelihood Estimation and Inference on Cointegration: With Applications to the Demand for Money, Oxford Bulletin of Economics and Statistics 52, 169-210.

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

King, R. G., C. I. Plosser, and S. T. Rebelo, 2001, Production, Growth and Business Cycles: Technical Appendix (http://www.kellogg.nwu.edu/faculty/rebelo/htm/kprnew.pdf).

Kullback, L., and R. A. Leibler, 1951, On Information and Sufficiency, Annals of Mathematical Statistics 22, 79-86.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (KPSS), 1992, Testing the Null of Stationarity Against the Alternative of a Unit Root, Journal of Econometrics 54, 159-178.

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

Kydland, F. E, and E. C. Prescott, 1996, The Computational Experiment: An Econometric Tool, Journal of Economic Perspectives 10, 69-85.

Nason, J. M., and T. Cogley, 1994, Testing the Implications of Long-Run Neutrality for Monetary Business Cycle Models, Journal of Applied Econometrics 9, 37-70.

Nelder, J. A., and R. Mead, 1965, A Simplex Method for Function Minimization, Computer Journal 7, 308-313.

Pagan, A. (editor), 1994, Calibration Techniques and Econometrics, Journal of Applied Econometrics 9.

Phillips, P. C. B. and P. Perron, 1988, Testing for a Unit Root in Time Series Regression, Biometrica 75, 335-346.

Schorfheide, F., 2000, Loss Function-Based Evaluation of DSGE Models, Journal of Applied Econometrics 15, 645-670.

Sims, C. A. (1996), Macroeconomics and Methodology, Journal of Economic Perspectives 10, 105-120.

Stokey, N. L., R. E. Lucas, and E. Prescott, 1989, Recursive Methods in Economic Dynamics (Harvard University Press, Cambridge, MA).

Watson, M. W., 1993, Measures of Fit for Calibrated Models, Journal of Political Economy 101, 1011-1041.