9-1-2008

Estimating the Speed of Convergence in the Neoclassical Growth Model: An MLE Estimation of Structural Parameters Using the Stochastic Neoclassical Growth Model, Time-Series Data, and the Kalman Filter

Daniel G. Swaine
College of the Holy Cross, dswaine@holycross.edu

Estimating the Speed of Convergence in the Neoclassical Growth Model: An MLE Estimation of Structural Parameters Using the Stochastic Neoclassical Growth Model, Time-Series Data, and the Kalman Filter

By

Daniel G. Swaine

September 2008

Department of Economics
College of the Holy Cross
Box 45A
Worcester, Massachusetts 01610
(508) 793-3362 (phone)
(508) 793-3710 (fax)

http://www.holycross.edu/departments/economics/website

Estimating the Speed of Convergence in the Neoclassical Growth Model: An MLE Estimation of Structural Parameters Using the Stochastic Neoclassical Growth Model, Time-Series Data, and the Kalman Filter

By
Daniel G. Swaine[†]
College of the Holy Cross

September 2008

Abstract

An important question is whether underdeveloped countries will converge to the per-capita income level of developed countries. Economists have used the disequilibrium adjustment property of growth models to justify the view that convergence should occur. Unfortunately, the empirical literature does not obey the "Lucas" admonition of estimating the structural parameters of a growth model that has the conditional convergence property and then computing the speed of convergence implied by the estimated structural parameters. In this paper, we use U.S. time-series data to estimate the structural parameters of a stochastic neoclassical growth model and compute the speed of conditional convergence in the non-stochastic model from the structural parameter estimates. We follow an approach used to econometrically estimate business cycle models via maximum likelihood. We obtain a speed of conditional convergence of 12.8 percent per-year for logarithmic consumer preferences and find that the data rejects the hypothesis of the 2 percent per-year speed of conditional convergence obtained in the empirical literature.

JEL Classification Codes: C30, C32, E 13, E32, O10, O11, O40, O41

Keywords: Convergence, Transitional Dynamics, Economic Growth, Economic Development, Real Business Cycle Models, Stochastic Growth Models, Time-Series Analysis

[†]Department of Economics, College of the Holy Cross, Worcester, MA 01610-2395, 508-793-3047 (phone), 508-793-3710 (fax), dswaine@holycross.edu

I. Introduction

In this paper, we use time-series data from the United States to estimate the structural parameters of a stochastic neoclassical growth model and compute the dynamic speed of disequilibrium adjustment in the non-stochastic growth model that are implied by the structural parameter estimates. In the economic development literature a very important question is whether less-developed countries will achieve the per-capita income level of developed countries. Starting with the seminal work of Barro (1991), Barro & Sala-i-Martin (1991,1992), Mankiw, Romer, & Weil (1992), economists have used the dynamic disequilibrium adjustment property – *known as conditional convergence* – of certain growth models to justify the view that underdeveloped countries would, for a time, grow faster than developed countries. If underdeveloped countries do grow faster, then, eventually, they may catch-up to the per-capita income level of the developed world – *called absolute convergence* – which would negate the need for any policy intervention to alleviate poverty.

Four different strains exist in the empirical literature to explore the convergence properties of growth models. These four strains are: (1) cross-sectional growth regressions; (2) panel data growth regressions; (3) time-series cointegration tests; (4) panel data cointegration tests. In the seminal Barro & Sala-i-Martin (1992) paper, the speed of convergence is one of the parameters estimated from a cross-sectional growth regression; they obtain an estimate of the dynamic speed of adjustment of 2 percent per-year. In later work, Sala-i-Martin (1996, 2002) claimed the 2 percent rate of convergence is an empirical regularity that occurs in every data set used. However, note that dynamic adjustment is a time-series process so estimating its speed with cross-sectional data seems rather odd. Recognizing this, Islam (1995) used panel data with fixed effects in order to allow time-series to influence the estimated speed of convergence. But, consistent with the cross-sectional growth regression literature, Islam constrained the speed of adjustment to be the same across countries. Islam estimates a larger speed of adjustment of about 10 percent per-year, suggesting that time-series effects are important.

In the time-series literature, most follow Bernard and Durlauf (1995), who employ cointegration techniques to test for a single common trend – a common long-run growth path. Disequilibrium adjustment dynamics to a single common growth path implies absolute convergence. Most studies in this vein of the literature demonstrate that the data reject absolute convergence. Because the cointegration procedure focuses on detecting long-run absolute convergence rather than on disequilibrium dynamics, this literature does not provide

any actual estimates of the dynamic speed of adjustment. Finally, the fourth strain uses panel data (with shorter time-series) to conduct a panel cointegration test for a single common trend to detect absolute convergence.

For four reasons, it is not really clear that the correlation that is estimated in the cross-section and panel data growth regression literature, and then interpreted as the speed of convergence, is, in fact, consistent with a speed of convergence that would be computed from estimates of a growth model's structural parameters. First, the cross-section and panel data empirical literatures do not obey the "Lucas" admonition (Lucas and Sargent (1979)) of estimating all of the structural parameters of a particular growth model that has the conditional convergence property and then computing the speed of convergence directly implied by the structural parameter estimates. The reduced-form specification that is employed in the literature contains coefficients (including the coefficient used to estimate the speed of convergence), that are complex, nonlinear functions of the growth model's structural parameters. Therefore, the structural parameters of the growth model impose coefficient restrictions on the empirical specifications, restrictions that are not utilized in estimation.

Second, to infer the speed of convergence directly from one of these coefficients using the standard cross-sectional growth regression, it is necessary to assume absolute convergence – directly implying that each observation in the cross-section has a growth model with structural parameter values that are identical across observations. Yet the same literature, stresses a conditional convergence result, with the direct implication that the structural parameter values are *not identical* across observations – thus the assumption needed to estimate a single speed of convergence is invalid. Third, it is not clear that the conditioning variables that proxy for the unobserved steady-state equilibrium (also a complex function of a growth model's structural parameters) are, in fact, really good proxies – another reason for concern about the unbiasedness of the convergence speed estimator. Fourth, the assumption that conditioning variables can proxy for the variation in structural parameter values across observations that would lead to variation in the steady-states is inconsistent with the assumption that the structural parameter values are identical across observations, which is necessary to infer the speed of convergence from both the cross-sectional and panel data growth regressions. For these four reasons, the claim that the two percent speed of convergence is an empirical regularity along with any claim that they have validly estimated the speed of convergence should be called into question.

In this paper, we follow a different approach that takes the Lucas admonition seriously and is used to econometrically estimate business cycle models (Altug (1989), McGrattan (1994), McGrattan, Rogerson, & Wright (1997), and Ireland (2001, 2004)). We use the Kalman Filter to obtain the likelihood function of a log-linearized approximation to the stochastic neoclassical growth model, estimate via maximum likelihood the growth model's structural parameters, and then compute the dynamic speed of adjustment in the non-stochastic model that is implied by the structural parameter estimates. We obtain a dynamic speed of conditional convergence of 12.8 percent per-year for logarithmic consumer preferences. A confidence interval for this dynamic speed of adjustment is very tightly estimated (from 12.6 percent per-year to 13.0 percent per-year) and simple t-tests reject the 2-percent per-year speed of convergence that Sala-i-Martin claims is an empirical regularity. We find that United States macroeconomic time-series data rejects the hypothesis of a 2 percent per-year speed of conditional convergence.

The plan for the remainder of this paper follows. Because the empirical convergence literature does not estimate structural parameters and the real business cycle literature is interested in a different set of questions, we do not provide a review of the literature on convergence, but refer the reader to the excellent survey of the empirical convergence literature by Islam (2003). In section 2, we present the stochastic neoclassical growth model and derive the cross-sectional and panel-data growth regressions that are used in the literature to demonstrate the problems with this literature that we pointed to above, and as a contrast to the approach that we employ in this paper. In section 3, we discuss the estimation method that we employ and the data used in the estimation. In section 4, we present and discuss the empirical results. Finally, in section 5, we discuss our conclusions and present ideas for additional research.

II. The Stochastic Neoclassical Growth Model

The stochastic neoclassical growth model is a dynamic stochastic general equilibrium (DSGE) model. Because of the general equilibrium nature of the model, we briefly present both the consumer and producer's optimization problems in Appendix A. In this section, we lay out the social planning problem that is typically used to solve competitive equilibrium business cycle models.

2.1 The Government Planner's Problem

The dynamic optimization problem for the government planner is to:

(1) $\underset{\widetilde{c}_t, \widetilde{k}_{t+1}}{\text{Max}} \quad E_0\left[\sum_{t=0}^{\infty} \overline{B}^t u(\widetilde{c}_t)\right] \quad$ s.t. $\quad 0 \le E_t\left(\phi(\widetilde{k}_t, z_t) - (\eta + \gamma + \delta)\widetilde{k}_t - \Delta\widetilde{k}_t - \widetilde{c}_t\right) \quad \forall \quad t = 0, \ldots, \infty$

where: $\widetilde{k}_t$: The stock of adjusted per-capita capital in period t. $\quad \forall \quad t = 1, \ldots, \infty$.

$\widetilde{c}_t$: The adjusted per-capita level of consumption in period t. $\quad \forall \quad t = 0, \ldots, \infty$.

$u(\widetilde{c}_t) = \dfrac{\widetilde{c}_t^{1-\theta} - 1}{1 - \theta}$: a CIES (Constant Intertemporal Elasticity of Substitution) function.

when $\theta = 1$, $u(\widetilde{c}_t) = \ln\widetilde{c}_t$

$\overline{B}$: is the discount rate, with $\overline{B} = \dfrac{1}{(1 + (\rho - \eta - (1-\theta)\gamma))}$,

where: $(\rho - \eta - (1-\theta)\gamma) > 0$.

$\phi(\widetilde{k}_t, z_t) \equiv A\widetilde{k}_t^{\alpha} z_t$: is a Cobb-Douglas production function

$\rho, \eta, \theta, A, \gamma, \alpha, \delta$ are the structural parameters of the problem, with ρ being the consumer's rate of time preference; η being the rate of growth of the population; γ being the rate of growth of Harrod-neutral technical change; θ being the coefficient of relative risk aversion, which is the inverse of the Intertemporal substitution elasticity; α being the output elasticity of capital; δ being the rate of depreciation on physical capital; A is a parameter that sets the scale of output in the Cobb-Douglas production function.

z_t : is a stochastic technology shock that obeys the following assumptions. The informational assumption is that the technology shock, z_t, is revealed at the beginning of period t.

(i) $\ln\hat{z}_t = \omega\ln\ln\hat{z}_{t-1} + \varepsilon_t$, where: $\quad 0 < \omega < 1$

where: $\ln\hat{z}_t \equiv \ln z_t - \ln\overline{z}$

(ii) $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$

2.2 Model Solution for the Speed of Convergence

The Lagrangian of the social planning problem is given by:

(2) $L = E_0\left[\sum_{t=0}^{\infty} \overline{B}^t\left\{\dfrac{\widetilde{c}_t^{1-\theta} - 1}{1 - \theta} + \lambda_t\left(A\widetilde{k}_t^{\alpha} z_t + [1 - (\eta + \gamma + \delta)]\widetilde{k}_t - \widetilde{k}_{t+1} - \widetilde{c}_t\right)\right\}\right]$

The three main first-order conditions (ignoring terminal and initial conditions) are:

(3) $E_t\left[\dfrac{\partial L}{\partial\widetilde{c}_t}\right] = E_t\left[\beta^t\left(\widetilde{c}_t^{-\theta} - \lambda_t\right)\right] = 0 \quad \forall \quad t = 0, \ldots, \infty$

(4) $E_t\left[\dfrac{\partial L}{\partial\widetilde{k}_{t+1}}\right] = E_t\left\{\beta\lambda_{t+1}\left[\alpha A\widetilde{k}_{t+1}^{-(1-\alpha)} z_{t+1} + [1 - (\eta + \gamma + \delta)]\right] - \lambda_t\right\} = 0 \quad \forall \quad t = 0, \ldots, \infty$

(5) $E_t\left[\dfrac{\partial L}{\partial\lambda_t}\right] = E_t\left[\beta^t\left(A\widetilde{k}_t^{\alpha} z_t + [1 - (\eta + \gamma + \delta)]\widetilde{k}_t - \widetilde{k}_{t+1} - \widetilde{c}_t\right)\right] = 0 \quad \forall \quad t = 0, \ldots, \infty$

This system of first-order conditions is log-linearized and then reduced to a two dimensional system of first-order log-linear stochastic difference equations, given by:

$$(6) \quad \begin{bmatrix} E_t \ln \hat{\tilde{k}}_{t+1} \\ E_t \ln \hat{\lambda}_{t+1} \end{bmatrix} = \mathbf{M} \begin{bmatrix} \ln \hat{\tilde{k}}_t \\ \ln \hat{\lambda}_t \end{bmatrix} + \mathbf{Q} [\ln \hat{z}_t]$$

where: $\ln \hat{\tilde{k}}_t$: is given by: $\ln \hat{\tilde{k}}_t = \ln \tilde{k}_t - \ln \overline{\tilde{k}}$, where: $\overline{\tilde{k}}$ is the non-stochastic steady-state, or intertemporal equilibrium value of the adjusted per-capita capital stock.

$\ln \hat{\lambda}_t$: is given by: $\ln \hat{\lambda}_t = \ln \lambda_t - \ln \overline{\lambda}$, where: $\overline{\lambda}$ is the non-stochastic steady-state, or intertemporal equilibrium value of the shadow price of adjusted per-capita consumption – the marginal utility of adjusted per-capita consumption.

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \ \mathbf{Q} = \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix}$$

$$M_{11} = [1 + (\rho - \eta - (1-\theta)\gamma)]$$

$$M_{12} = \frac{(\rho + \theta\gamma + \delta) - \alpha(\eta + \gamma + \delta)}{\alpha\theta}$$

$$M_{21} = (1-\alpha)(\rho + \theta\gamma + \delta)$$

$$M_{22} = 1 + \frac{(1-\alpha)(\rho + \theta\gamma + \delta)[(\rho + \theta\gamma + \delta) - \alpha(\eta + \gamma + \delta)]}{\alpha\theta[1 + (\rho - \eta - (1-\theta)\gamma)]}$$

$$Q_{11} = \frac{(\rho + \theta\gamma + \delta)}{\alpha}$$

$$Q_{21} = \frac{(\rho + \theta\gamma + \delta)((1-\alpha)(\rho + \theta\gamma + \delta) - \alpha\omega)}{\alpha[1 + (\rho - \eta - (1-\theta)\gamma)]}$$

To solve this system of log-linear stochastic difference equations, we employ the Blanchard and Kahn (1980) procedure, which yields the following decision rules for the state and co-state variables, respectively:

$$(7) \quad \ln \hat{\tilde{k}}_{t+1} = \Lambda_1 \ln \hat{\tilde{k}}_t + \frac{[(\Lambda_2 - \omega) + (\Lambda_1 - M_{11})]Q_{11} - M_{12}Q_{21}}{(\Lambda_2 - \omega)} \ln \hat{z}_t$$

$$(8) \quad \ln \hat{\lambda}_t = \frac{(\Lambda_1 - M_{11})}{M_{12}} \ln \hat{\tilde{k}}_t + \frac{(\Lambda_1 - M_{11})Q_{11} - M_{12}Q_{21}}{M_{12}(\Lambda_2 - \omega)} \ln \hat{z}_t$$

where: Λ_1 : is the stable eigenvalue of the matrix $\mathbf{M}$.

Λ_2 : is the unstable eigenvalue of the matrix $\mathbf{M}$.

The dynamic speed of disequilibrium adjustment, (the speed of convergence), called beta in the literature, is a linear function of the stable eigenvalue, Λ_1, of the matrix $\mathbf{M}$. The stable eigenvalue is:

$$(9) \ \Lambda_1 = \frac{(M_{11} + M_{22})}{2} - \frac{\left[(M_{11} + M_{22})^2 - 4(M_{11}M_{22} - M_{12}M_{21})\right]^{\frac{1}{2}}}{2}$$

while the equation determining the speed of convergence is given by:

$$(10) \quad \beta = (1 - \Lambda_1).$$

In this paper, our task is to obtain a good time-series estimate of the parameter beta, which is computed after obtaining the estimates of the "deep" structural parameters of the stochastic neoclassical growth model.

2.3 Empirical Implementations for Estimating the Speed of Convergence

2.3.1 Empirical Implementations in the Convergence Literature

The specification that is estimated in the empirical convergence literature is derived from two equations above: (equation (7) and a log-linearized version of the production function), which when combined and reduced yields the following result:

$$(11) \quad \ln \hat{\tilde{y}}_t = \Lambda_1 \ln \hat{\tilde{y}}_{t-1} + v_t$$

$$\text{where: } v_t = (1 + \alpha\vartheta)\ln \hat{z}_t - \Lambda_1 \ln \hat{z}_{t-1}$$
$$\vartheta = \frac{\left[(\Lambda_2 - \omega) + (\Lambda_1 - M_{11})\right]Q_{11} - M_{12}Q_{21}}{(\Lambda_2 - \omega)}$$

From equation (11) we can derive two reduced-form equations that might be estimated directly from the data. The first equation is a time-series equation and is identical to a specification of an equation typically used in a Dickey-Fuller test for a unit root (a small estimate of beta would likely imply a unit-root in per-capita income).

$$(12) \quad \Delta \ln y_t = a + \beta t - \beta \ln y_{t-1} + v_t$$

$$\text{where: } a = (1 - \beta)\gamma + \beta \ln \bar{\bar{y}}$$
$$\Delta \ln y_t \equiv \ln y_t - \ln y_{t-1}$$

The second equation is the derived theoretical specification used in the cross-sectional and panel data literature:

$$(13) \quad \frac{\ln y_T - \ln y_0}{T} = \gamma - \frac{(1 - \Lambda_1^T)}{T}\ln y_0 + \frac{(1 - \Lambda_1^T)}{T}\ln \bar{\bar{y}} + u_T$$

$$\text{where: } \frac{\ln y_T - \ln y_0}{T} \quad \text{is the average growth rate of per-capita income.}$$
$$\ln y_0 \quad \text{is the log of initial per-capita income.}$$

$\ln \overline{\overline{y}}$ is the log of the steady-state level of adjusted per-capita income, which is unobserved and is replaced with a set of control variables. Note that the steady-state level of adjusted per-capita income is given by:

$$\overline{\overline{y}} = A^{\frac{1}{(1-\alpha)}} \left[\frac{\alpha}{(\rho + \theta\gamma + \delta)} \right]^{\frac{\alpha}{(1-\alpha)}}$$

$\dfrac{\left(1 - \Lambda_1^T\right)}{T}$ is the coefficient from which beta is estimated.

$$u_T = \frac{1}{T} \sum_{j=0}^{T} \Lambda_1^j v_{t-j}$$

There are a number of variants of equation (13) that are used in the empirical work in the cross-section and panel data literature on convergence, which we present below.

(13a) $\left(\dfrac{\ln y_T - \ln y_0}{T} \right)_{it} = a_i - b_i (\ln y_0)_{it} + b_i \ln \overline{\overline{y}}_{it} + (u_T)_{it}$

where:
$$a_i = \gamma_i$$
$$b_i = \frac{\left(1 - \Lambda_1^T\right)_i}{T}$$

(13b) $\left(\dfrac{\ln y_T - \ln y_0}{T} \right)_i = a - b(\ln y_0)_i + (u_T)_i$

where:
$$\widetilde{a} = \gamma + \frac{\left(1 - \Lambda_1^T\right)}{T} \ln \overline{\overline{y}}$$
$$b = \frac{\left(1 - \Lambda_1^T\right)}{T}$$

Equation (13a) is the broadest possible variant of equation (13) and is a conditional convergence equation specified for panel data. Equation (13b) assumes (without testing) absolute convergence, an assumption that may be invalid. Under this assumption, which implies that the structural parameter values of a neoclassical growth model are identical across observations, cross sectional data would be used in estimation and the only reason for variation across observations on the right-hand-side is due to the different starting points. To obtain equation (13b) from equation (13a), it is necessary to impose the following coefficient restrictions (embedding the assumption of absolute convergence): (1) $\ln \overline{\overline{y}}_{it} = \ln \overline{\overline{y}} \quad \forall \quad i = 1,\ldots,N; t = 1,\ldots,T$; (2) $a_i = a \quad \forall \quad i = 1,\ldots,N$; (3) $b_i = b \quad \forall \quad i = 1,\ldots,N$. From this discussion, it should be apparent that the cross-sectional approach to convergence was designed mainly to analyze conditions of absolute convergence.

To analyze conditional convergence using equation (13), variant equation (13a) must be used. However, as specified, equation (13a) is not estimable for two reasons: (i) the steady-state equilibrium $\ln \overline{\overline{y}}_{it} = \ln \overline{\overline{y}}_i \quad \forall \quad t = 1,\ldots,T$ is not observable without knowledge of the values of the structural parameters of the growth model for each cross-sectional observation – under conditional convergence these parameters vary over the cross-section, but not necessarily across time; (ii) all of the coefficients in equation (13a) vary over the cross-section, so that only the time-series information from each cross-sectional observation actually would be used in estimation. Yet, the approach that derives equation (13) was not designed for time-series analysis, thus the time-series in the panel is likely to be short, so that degrees of freedom will be problematic.

To address these two estimation problems, two other variants of equation (13) are of note, and are the predominant specifications that are used in the empirical cross-section and panel data literature. We present these two variants as follows.

$$(13a.1) \quad \left(\frac{\ln y_T - \ln y_0}{T} \right)_{it} = a_i - b(\ln y_0)_{it} + b\mathbf{d}'\mathbf{x}_{it} + (u_T)_{it}$$

$$(13a.2) \quad \left(\frac{\ln y_T - \ln y_0}{T} \right)_i = a - b(\ln y_0)_i + b\mathbf{d}'\mathbf{x}_i + (u_T)_i$$

where: $\mathbf{x}_{it}$, and $\mathbf{x}_i$ is a vector of conditioning variables.

Equation (13a.1) is the panel data implementation of equation (13) that is used in the literature, while equation (13.a.2) is the cross-sectional implementation of equation (13) that is used in the literature. Equation (13a.1) imposes the following coefficient restrictions on equation (13a): (a) $b_i = b \quad \forall \quad i = 1,\ldots,N$; (b) $\ln \overline{\overline{y}}_{it} \equiv \mathbf{d}'\mathbf{x}_{it}$. Equation (13a.2) adds one additional restriction: (c) $a_i = a \quad \forall \quad i = 1,\ldots,N$. The main problem with equations (13a.1) and equation (13a.2) is that they are logically inconsistent with respect to its theoretical counterpart of equation (13). First, restriction (b) says that the reason for the steady-state equilibrium to vary across cross-sectional observations is only because of the conditioning variables. The real reason that the steady-state equilibrium varies across the observations is that the structural parameter values of the growth model vary across the cross-section. The same structural parameters that are contained in the expression for the steady-state equilibrium also are contained in the expression for the convergence coefficient (the stable eigenvalue of the matrix $\mathbf{M}$). The convergence coefficients are assumed to be the same across the cross-section, but the steady-

states are assumed to vary – thus the logical inconsistency between the empirical implementations of equation (13a.1) and (13a.2) and its theoretical counterpart of equation (13). Additionally, the conditioning variables used to proxy for variation in the steady-state may be very imperfect proxies (also untested), thereby biasing the estimate of the speed of convergence due to measurement error.

Finally, but even more importantly, it should be noted that in equations (12), (13), (13a), (13a.1), (13a.2), (13b), the structural parameters impose restrictions on the coefficients of all of the variables in these equation specifications. These theoretically important coefficient restrictions are not imposed on the specifications estimated in the empirical literature, because the empirical literature does not attempt to estimate the structural parameters. With respect to equations (13a.1) and (13a.2), these specifications are estimated, as is, in unrestricted form, and then the coefficient on initial per-capita income is used to back-out the speed of convergence rather than using the coefficient estimates of the structural parameters to compute the implied speed of convergence. Therefore, in our opinion, not estimating the structural parameters from the data creates problems in the ability to interpret the empirical results that are obtained relative to what would occur if the structural parameters had been estimated and these restrictions imposed on the specifications estimated.

2.3.2 *Empirical Implementation used in this paper*

The argument for a different approach is made by Sargent et al (2005, 2007) in the ABC's and D's of understanding VARs. VARs are similarly estimated in unrestricted form. Sargent et al argue convincingly that economic theory imposes restrictions on the coefficients of empirical specifications and that the likelihood function contains all of the information that the data contains about any set of structural parameters of an economic model. They further state that an analyst should estimate the structural parameters of a trusted model via likelihood methods rather than turning their back on this approach, as the empirical convergence literature seems to do. The new dynamic macro has built the foundations of macroeconomics on the back of the stochastic neoclassical growth model, suggesting that the profession certainly trusts this model. All linearized, or log-linearized DSGE models (which the neoclassical growth model certainly is) yield the following state-space system of equations:

$$(14) \qquad \mathbf{y}_t = \mathbf{A}\mathbf{s}_t + \mathbf{B}\mathbf{x}_t + \mathbf{\mu}_t$$

$$(15) \qquad \mathbf{s}_{t+1} = \mathbf{C}\mathbf{s}_t + \mathbf{D}\mathbf{\varepsilon}_t$$

where: $\mathbf{y}_t$ is an (M x1) vector of the observed endogenous variables.

          $\mathbf{x}_t$ is a (K x 1) vector of the observed exogenous variables.

          $\mathbf{s}_t$ is an (n x 1) vector of unobserved "state" variables.

          $\boldsymbol{\varepsilon}_t$ is an (n x 1) vector of "driving" or shock variables .

          $\boldsymbol{\mu}_t$ is an (M x 1) vector of "measurement" errors in the observed endogenous variables.

          A is an (M x n) matrix of estimable parameters.

          B is an (M x K) matrix of estimable parameters.

          C is an (n x n) matrix of estimable parameters.

          D is an (n x n) selection matrix.

          R is the (M x M) contemporaneous variance-covariance matrix of the measurement errors, such that: $\mathbf{R} = E(\boldsymbol{\mu}_t \boldsymbol{\mu}_t')$.

          Q is the (n x n) contemporaneous variance-covariance matrix of the shocks, such that: $\mathbf{Q} = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$.

Equations (14) are known as the measurement system of equations. Equations (15) are known as the state-system of equations. There are two sets of stochastic shocks in a state-space system. The meaningful economic shocks of the model are incorporated into, $\boldsymbol{\varepsilon}_t$. Most DSGE's result in the second set of shocks, $\boldsymbol{\mu}_t$, being zero. However, it is these shocks, $\boldsymbol{\mu}_t$, that are called the measurement errors (representing measurement error in the basic economic data that are used) that are needed for econometric estimation of the model.

In our application, we set-up the following state-space vectors with these variables:

$\mathbf{y}_t \equiv \begin{bmatrix} \ln c_t & \ln i_t & \ln y_t & \ln k_t & r_t & \ln w_t \end{bmatrix}'$; $\mathbf{x}_t = \begin{bmatrix} 1 & t \end{bmatrix}'$; $\mathbf{s}_t = \begin{bmatrix} \ln \tilde{k}_t & \ln z_t \end{bmatrix}'$; $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_t \end{bmatrix}$. The elements of the matrices **A**, **B**, **C, Q** are all complex, nonlinear functions of the nine "structural" parameters of the neoclassical growth model, which are contained in the parameter vector: $\boldsymbol{\Theta}_e = \begin{bmatrix} \rho & \eta & \theta & A & \alpha & \gamma & \delta & \omega & \sigma_\varepsilon \end{bmatrix}'$. The entire parameter vector contains fifteen elements: $\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{\Theta}_e & \boldsymbol{\Theta}_v \end{bmatrix}$; the vector, $\boldsymbol{\Theta}_v = \operatorname{diag}(\mathbf{R})$, contains the six measurement error variances.

III. Estimation Method and Data

3.1 Estimation Method

The system of equations in equations (14) and (15) cannot be estimated because of the unobservable state vector. However, the Kalman Filter provides a means to estimate the state vector and provides an equivalent but

observable state-space system called the innovations representation. The innovations representation allows one to

obtain the likelihood function that is used to estimate DSGE models via the method of maximum likelihood. The

innovations representation is given by:

(16) $\mathbf{y}_t = \mathbf{A}\hat{\mathbf{s}}_{t|t-1} + \mathbf{B}\mathbf{x}_t + \hat{\mathbf{u}}_t$

(17) $\hat{\mathbf{s}}_{t+1|t} = \mathbf{C}\hat{\mathbf{s}}_{t|t-1} + \hat{\mathbf{K}}_t\hat{\mathbf{u}}_t$

where: $\hat{\mathbf{s}}_{t|t-1}$ is an estimate of the (n x 1) "state" vector using information on the observed endogenous variables, $\mathbf{y}$, and the exogenous variables, $\mathbf{x}$, through time-period (t-1).

$\hat{\mathbf{u}}_t$ is an (M x 1) vector of innovations. If we add and subtract $\mathbf{A}\hat{\mathbf{s}}_{t|t-1}$ to equation (14), the innovations are given by: $\hat{\mathbf{u}}_t = \mathbf{A}(\mathbf{s}_t - \hat{\mathbf{s}}_{t|t-1}) + \boldsymbol{\mu}_t$. From equation (16), the innovations are the errors of the equation and are given by: $\hat{\mathbf{u}}_t = \mathbf{y}_t - \mathbf{A}\hat{\mathbf{s}}_{t|t-1} - \mathbf{B}\mathbf{x}_t = \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}$.

$\hat{\mathbf{y}}_{t|t-1}$: is the predicted value of the observable variables, which is given by: $\hat{\mathbf{y}}_{t|t-1} = \mathbf{A}\hat{\mathbf{s}}_{t|t-1} + \mathbf{B}\mathbf{x}_t$.

$\mathbf{s}_t - \hat{\mathbf{s}}_{t|t-1}$ is an (n x n) vector of "forecast" errors for the "state" estimate.

$\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{R},\mathbf{Q}$: are matrices of estimable parameters as in equations (14) and (15).

$\hat{\mathbf{K}}_t$ is an (n x M) matrix that is called the Kalman gain matrix.

$\hat{\boldsymbol{\Omega}}_t$ is the (M x M) contemporaneous variance-covariance matrix of the innovations, with: $\hat{\boldsymbol{\Omega}}_t = \mathrm{E}(\hat{\mathbf{u}}_t\hat{\mathbf{u}}_t') = \mathrm{E}\left[(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})'\right]$.

$\hat{\boldsymbol{\Sigma}}_{t|t-1}$ is the (n x n) contemporaneous variance-covariance matrix of the "state" forecast errors, such that: $\hat{\boldsymbol{\Sigma}}_{t|t-1} = \mathrm{E}\left[(\mathbf{s}_t - \hat{\mathbf{s}}_{t|t-1})(\mathbf{s}_t - \hat{\mathbf{s}}_{t|t-1})'\right]$.

$\hat{\mathbf{s}}_{t|t-1}$, $\hat{\mathbf{K}}_t$, $\hat{\boldsymbol{\Omega}}_t$, $\hat{\boldsymbol{\Sigma}}_{t|t-1}$: are computed from the observable data using the Kalman Filter, which is a well-known set of recursions (See: Hamilton, Chapter 13, (1994))

The log-likelihood function is computed from the above innovations representation of the model.

Maximum likelihood estimation uses numerical methods (a Quasi-Newton hill-climbing algorithm for

unconstrained optimization) to optimize the log of the likelihood function, which is given by:

(18) $\ln \mathrm{L}(\boldsymbol{\Theta}) = -\dfrac{\mathrm{MT}}{2}\ln(2\pi) - \dfrac{1}{2}\left[\sum_{t=1}^{T}\ln\left|\hat{\boldsymbol{\Omega}}_t\right| + \sum_{t=1}^{T}\hat{\mathbf{u}}_t'\hat{\boldsymbol{\Omega}}_t^{-1}\hat{\mathbf{u}}_t\right]$

We implement both the Kalman Filter, used to construct the innovations representation from the data, and the

Quasi-Newton hill-climbing algorithm in a FORTRAN 95 program specially written for this task by the author.

Quasi-Newton methods depend upon finding a good point near the optimum, from which to start the hill-climb. To obtain a good starting point, we employ the Christiano & Eichenbaum (1992) procedure of specifying a sufficient set of first moments from which to just identify and estimate all of the structural parameters of the model via Generalized Method of Moments estimation (GMM). (see Appendix B).

In our application, to obtain the state-space system of equations (14) and (15), the three main variables in the model ($\widetilde{c}_t$, $\widetilde{i}_t$, $\widetilde{y}_t$) provide the basic measurement equations – we log-linearize the consumption first-order condition of the government planning problem (equation (3)) to obtain the measurement equation for consumption and insert the decision rule for the shadow price (equation (8)); we log-linearize the economy's resource constraint (equation (5)) to obtain the measurement equation for investment and insert the decision rule for capital; and we log-linearize the production function to obtain the measurement equation for output. In our application, which includes more parameters than are typically estimated in the one sector real business cycle model, we add two additional measurement equations – we log-linearize the producer's first-order condition (equation (A2.2)) to obtain the measurement equation for the real interest rate; and finally, we log-linearize the producer's zero profit condition (equation (A2.3)) to obtain the measurement equation for the log real wage. The decision rule for capital (equation (7)) provides the first state equation and the law of motion for the shock (stochastic assumption (i)) provides the second state equation.

3.2 Estimation and Data

The observable variables that we employ are defined as:

c_t : is per-capita consumption, with: $c_t = \dfrac{C_t}{L_t}$ $\qquad$ i_t : is per-capita investment, with: $i_t = \dfrac{I_t}{L_t}$

y_t : is per-capita output, with: $y_t = \dfrac{Y_t}{L_t}$ $\qquad$ k_t : is per-capita capital stock, with: $k_t = \dfrac{K_t}{L_t}$

r_t : is the real interest rate. $\qquad$ w_t : is the real wage rate.

C_t : is aggregate real consumption. $\qquad$ I_t : is aggregate real investment.

Y_t : is aggregate real output, measure by real GDP. $\qquad$ K_t : is aggregate real capital stock.

L_t : is aggregate employment in persons.

The connection between these observable variables and the model variables is given by the following

trend-stationary definitions: $\widetilde{c}_t = \dfrac{c_t}{\psi^t}$; $\quad \widetilde{i}_t = \dfrac{i_t}{\psi^t}$; $\quad \widetilde{y}_t = \dfrac{y_t}{\psi^t}$; $\quad \widetilde{k}_t = \dfrac{k_t}{\psi^t}$; and $\quad \psi = (1+\gamma)$. Because the

underlying per-capita variables cannot reject unit roots, we assume the variables are difference-stationary. After

log-linearly de-trending the per-capita data, we pre-filter the data with the Christiano-Fitzgerald Band-Pass filter,

which keeps only the business cycle and very high frequencies in the filtered data, and then add back the log-

linear trend to the data for the MLE estimation procedure.

We obtain our quarterly data from the Bureau of Economic Analysis's National Income and Product

Accounts (NIPA) for the time-span of 1948:1 to 2005:4. We re-classify durable goods purchases from the NIPA

as investment and estimate durable good consumption as the flow from the stock of durable goods (from the

tangible wealth accounts in NIPA). We also classify government investment as investment and include

government consumption as consumption to match the model definitions of consumption and investment. Net

exports is included in consumption. Output is measured by GDP. The capital stock is taken from the net tangible

wealth accounts of the NIPA. The real gross interest rate $(r_t + \delta)$ is estimated from real capital income in the

national income accounts and then dividing by the capital stock. The net real interest rate is obtained after

eliminating real depreciation from the capital income series and then dividing by the capital stock. The real wage

rate is obtained from the real wage income series, and then dividing by employment, which is drawn from the

monthly establishment series of the BLS.

In an earlier paper, Swaine (2003), we noted that much of the data was plagued by structural breaks,

which yielded estimated parametric instability in the stochastic neoclassical growth model using the GMM

method, with the first moment GMM estimating equations being non-stationary. Because the Hansen (1982)

GMM estimation procedure that is used to produce a good starting point for MLE estimation requires stationary

first-moment equations, we adjust all data used in both the GMM estimation and the MLE estimation for

structural breaks prior to estimation (see Swaine (2003) for information on these type of adjustments).

IV. Empirical Results

4.1 Choice of Functional Form for the Utility Function

The choice between the Constant Intertemporal Elasticity of Substitution (CIES) utility function and the logarithmic utility function is a test of whether or not the coefficient of relative risk aversion, θ, which is the inverse of the elasticity of intertemporal substitution in consumption, is equal to one. We estimate both CIES and logarithmic specifications in order to perform this test. Because the production function scale parameter, A, is not very important to our analysis, we set the value for the coefficient, A, equal to its GMM estimate and conduct the test conditional upon the value of this coefficient. Like most of the literature that has attempted to estimate the intertemporal substitution elasticity (Hall (1988), McGrattan et al (1997), Imai & Keane (2004), Lee (2001), Yogo (2004)) we find that the data is uninformative about the value of this parameter – the coefficient of relative risk aversion is imprecisely estimated. A likelihood ratio test of the hypothesis that the coefficient of relative risk aversion is equal to one is such that one is unable to reject this hypothesis. Table 1 summarizes the results of the LR test. The Likelihood ratio is 0.4766 and it's chi-square probability value is 0.48994. Because of this test result, we employ logarithmic preferences in our estimation of the speed of convergence, which implicitly sets both the coefficient of relative risk aversion, θ, and the intertemporal substitution elasticity equal to one. Note that logarithmic preferences are used in much of the business cycle literature that we follow in this paper (McGrattan, Rogerson & Wright (1997), Ireland (2001, 2004), and Ireland & Schuh (2006)), so we are not alone in this choice.

4.2 GMM and MLE estimates of the Structural Parameters of the Neoclassical Growth Model

Table 2 presents the GMM estimates (the starting point for the MLE estimation procedure, see Appendix B for the moment equations used to obtain the starting point estimates) for the structural parameters of the neoclassical growth model, while Tables 3 & 4 present the MLE estimates for the structural parameters – Table 3 presents the MLE estimates conditional upon the GMM estimate of the structural parameter A, while Table 4 presents the full-set of parameter estimates, including A. In Tables 3 and 4, the reported standard errors are the square roots of the diagonal elements of the inverse of the Hessian matrix of the likelihood function. Table 5 presents an LR test of the restriction to the GMM estimate of the scale parameter, A, that we employed in the test for functional form above. All of the estimated parameters make sense, have the correct sign, and are statistically significant. Most of the parameters are not substantially different from the GMM estimates, with the exception of the stochastic shock persistence and the shock variance. This should not be surprising as the two approaches

estimate these two parameters by very different methods -- the MLE procedure assumes that capital and output are measured with error, while the GMM estimation procedure uses the Solow residual, which does not assume measurement error.

4.3 Computed Estimates of the Speed of Convergence

Table 6 presents our estimate of the speed of convergence computed from the structural parameter estimates listed in Tables 3 & 4. To compute this parameter, we use equations 9 and 10. A point estimate of the speed of convergence is 12.7 percent per-year computed from the structural parameter estimates in Table 3, and 12.8 percent per-year computed from the structural parameter estimates in Table 4. In Table 6, we report the constructed interval estimates of the speed of convergence and present t-tests of the hypothesis that the speed of convergence is equal to 2 percent. In order to construct standard errors for the computed speed of convergence, we employ the following formula: $\sigma_\beta = \sqrt{\mathbf{g}'\mathbf{H}^{-1}\mathbf{g}}$, where: $\mathbf{g}$ is the gradient vector (with respect to the five parameters that determine it) of the speed of convergence (computed with central finite differences), and $\mathbf{H}$ is the Hessian matrix of the likelihood function (also computed with central finite differences).

In Table 6, a 95 percent confidence interval estimate of the speed of convergence computed from the structural parameter estimates listed in Table 3 (conditional on the GMM estimate of the scale parameter A, which does not enter into the calculation of the speed of convergence) provide an interval for beta from 12.55 percent per-year to 12.84 percent per-year, which is very tightly estimated. A 95 percent confidence interval estimate of the speed of convergence computed from the structural parameter estimates listed in Table 4 provide a slightly larger interval of beta (because of the additional uncertainty in estimating the scale coefficient, A) from 12.6 percent per-year to 13.0 percent per year, which is still very tightly estimated. As one can see, the 2-percent claim of Sala-i-Martin is not within this interval, nor is the 10 percent rate estimated by Islam. A simple t-statistic test of the 2-percent speed of convergence hypothesis is rejected at the 5 percent level (t-stats of 145.34 for the Table 3 estimate, and 100.83 for the Table 4 estimate) and contains a probability value of 0.0000 in both cases (meaning that there isn't a probability level of the test that could support the 2-percent hypothesis. We feel quite safe in concluding that the macroeconomic time-series data of the United States does not support a 2 percent speed of convergence – a result very different from that obtained in cross-sectional estimation, and also different from

Islam's panel data procedure. This vast difference simply underscores the need to estimate structural parameters in the convergence literature.

Conducting a Likelihood Ratio (LR) test of the 2-percent speed of convergence hypothesis is far more difficult for two reasons. First, the test imposes a non-linear constraint on the MLE estimation procedure (which is a non-linear programming problem). Second, the Quasi-Newton hill-climbing method requires a good starting point from which to start the MLE estimation. From Table 6, it can be seen that the 2-percent hypothesis is quite far from the speed we estimated, and this creates a problem in finding a good starting point from which to test this hypothesis. We have written FORTRAN code for a sequential quadratic programming problem to implement the non-linear programming problem of imposing the non-linear restriction of a particular speed of convergence on the MLE estimation procedure. However, because of the problem of finding a good starting point, we have tested a much less ambitious hypothesis. Table 7 contains the LR non-linear test of hypothesis that the speed of convergence is 12 percent. The log-likelihood for a constrained estimation that restricts the speed of convergence to 12 percent is 3779.9821305. The LR test from the Table 4 estimate of beta gives a log-likelihood ratio statistic of 352.4827144. The hypothesis is rejected at a 5 percent level, and the likelihood ratio statistic has a probability of zero. Thus, the LR test must necessarily reject the hypothesis for every speed of convergence less than 12 percent (including the 2 percent and 10 percent hypotheses).

One implication of this much faster estimate of the speed of convergence is that the half-life of a difference in per-capita income from its intertemporal equilibrium value will last only 5.1 years (using the Table 3 estimate) or 5.06 years (using the Table 4 estimate). This is much faster than the 35 years estimated by Barro & Sala-i-Martin on the basis of a 2-percent speed of convergence. The policy implication of the faster speed is that if underdeveloped countries have conditional convergence speeds equally fast, then they must be converging to much different intertemporal equilibrium positions and this cries out for the need for policy intervention in order to alleviate the poverty that exists in underdeveloped societies. In contrast, the Barro & Sala-i-Martin estimate suggests that we do not need a policy intervention and if all economies are converging to the same equilibrium, then we can just wait for convergence to occur rather than to implement what some might consider a messy policy intervention.

V. Conclusion

In this paper we estimated the structural parameters of the stochastic neoclassical growth model in order to compute an estimate of the speed of convergence. The literature takes a far different approach of estimating unrestricted reduced-form equations for which the structural parameters impose restrictions on the empirical specification, but where the analyst does not impose the coefficient restrictions in the process of estimation. Sargent et al (2005, 2007) argue that this approach, similarly used in estimating VARs, is equivalent to turning one's back on the likelihood principle, by which the likelihood function contains all of the information that the data has on any particular economic model. Sargent et al argue that an analyst should use the likelihood principle to estimate the structural parameters for a trusted economic model. Because the dynamic macro theory that has supplanted the earlier static Keynesian theory (due to the Lucas critique) is built on the foundation of the stochastic neoclassical growth model, one must conclude that the profession believes that the neoclassical growth model is a trusted model.

Because of the persuasiveness of the Sargent et al argument, we decided to take a different approach with this paper, where we estimate the structural parameters of the profession's trusted model – the stochastic neoclassical growth model. Using logarithmic preferences, as does most of the business cycle literature because the data is uninformative about the intertemporal substitution elasticity, we estimate a speed of convergence which is more than six times faster than the 2 percent speed estimated in the cross-sectional literature and 2.8 percentage points higher than the 10 percent speed estimated in the panel data literature. We construct a 95 percent confidence interval estimate of the speed of convergence which is very tightly estimated, and for which neither the 2-percent speed of convergence from the cross-sectional literature, nor the 10-percent speed of convergence estimated in the panel data literature is contained within this interval. Simple t-tests reject the 2-percent speed of convergence at the 5 percent level, with a probability value that is zero. This empirical result underscores the need for analysts to estimate structural parameters in the empirical convergence literature.

References

S. Altug, "Time-to-Build and Aggregate Fluctuations: Some New Evidence," **International Economic Review**, Volume 30, 1989, pp. 889-920.

R.J. Barro, "Economic Growth in a Cross-Section of Countries," **Quarterly Journal of Economics**, May 1991, pp. 407-443.

R.J. Barro, "Determinants of Economic Growth in a Panel of Countries," **Annals of Economics & Finance**, Volume 4, 2003, pp.231-274

R.J. Barro & X. Sala-i-Martin, "Convergence Across States & Regions," **Brookings Papers on Economic Activity**, Part I, 1991, pp. 107-182.

R.J. Barro & X. Sala-i-Martin, "Convergence", **Journal of Political Economy**, April 1992, pp. 223-251.

A. Bernard & S. Durlauf, "Convergence in International Output," **Journal of Applied Econometrics**, 1995, pp. 97-108.

O. Blanchard and C. Kahn, "The Solution of Linear Difference Models Under Rational Expectations," **Econometrica**, Volume 48, No. 5, 1980, pp. 1305-1311.

L. Christiano and M. Eichenbaum, "Current Real Business Cycle Theories and Aggregate Labor-Market Fluctuations," **American Economic Review**, Volume 82, No. 3, June 1992, pp. 430-450.

L. Christiano and T. Fitzgerald, "The Band-Pass Filter," **International Economic Review**, May 2003, Volume 44, Issue 2, pp. 435-465.

J. Fernandez-Villaverde, J. Rubio-Ramirez, & T. Sargent, "A,B,C's (and D's) of Understanding VARS," **Working Paper**, 2005.

J. Fernandez-Villaverde, J. Rubio-Ramirez, T. Sargent, & M. Watson, "A,B,C's (and D's) of Understanding VARS," **American Economic Review**, June 2007, pp. 1021-1026.

R.E. Hall, "Intertemporal Substitution in Consumption," **Journal of Political Economy**, 1988, pp. 339-357.

J.D. Hamilton, **Time-Series Analysis**, Princeton University Press, 1994

L. Hansen, "Large Sample Properties of Generalized Method of Moments Estimators," **Econometrica**, Volume 5, July 1982, pp. 1029-1054.

S. Imai & M. Keane, "Intertemporal Labor Supply and Human Capital Accumulation," **International EconomicReview**, May 2004, pp. 601-641.

P. Ireland, "Technology Shocks and the Business Cycle: An Empirical Investigation," **Journal of Economic Dynamics and Control**, Volume 25, 2001, pp. 703-719.

P. Ireland, "A Method of Taking the Models to the Data," **Journal of Economic Dynamics and Control**, March 2004, pp. 1205-1226.

P. Ireland & S. Schuh, "Productivity and Macroeconomic Performance: Interpreting the Past and Predicting the Future with a Two-Sector RBC Model," **FRB Working Paper 06-10**, April 2006.

N. Islam, "Growth Empirics: A Panel Data Approach," **Quarterly Journal of Economics**, November 1995, pp. 1127-1170.

N. Islam, "What have We Learned From the Convergence Debate," **Journal of Economic Surveys**, 2003, No. 3, pp. 309-362.

C. Lee, "Finite Sample Bias in IV Estimation of Intertemporal Labor Supply Models: Is the Intertemporal Substitution Elasticity Really Small, **Review of Economics & Statistics**, November 2001, pp. 638-646.

R.E. Lucas & T. Sargent, "After Keynesian Macroeconomics," in Preston Miller (ed), **The Rational Expectations Revolution: Readings From the Front Line**, 1995.

N. G. Mankiw, D. Romer, & D. Weil, "A Contribution to the Empirics of Growth," **Quarterly Journal of Economics**, May 1992, pp. 407-437.

E. McGrattan, "The Macroeconomic Effects of Distortionary Taxation," **Journal of Monetary Economics**, Volume 33, 1994, pp. 573-601.

E. McGrattan, R. Rogerson, & R. Wright, "An Equilibrium Model of the Business Cycle with Household Production and Fiscal Policy," **International Economic Review**, Volume 38, 1997, pp. 267-290.

M. Ogaki and C. Reinhart, "Measuring Intertemporal Substitution: The Role of Durable Goods," **Journal of Political Economy**, 1998, Volume 106, No. 5, pp. 1078-1098.

X. Sala-i-Martin, "Regional Cohesion: Evidence & Theories of Regional Growth and Convergence," **European Economic Review**, June 1996, pp. 1325-1352.

X. Sala-i-Martin, "Fifteen Years of New Growth Economics: What have We Learned?," **Working Paper**, 2002.

X. Sala-i-Martin, et al, "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates Approach," **American Economic Review**, September 2004, pp. 813-835.

D. Swaine, "Are Taste and Technology Parameters Stable?: A Test of Deep Parameter Stability in RBC Models of the US Economy," **FRB Working Paper 01-5**, November 2001, Revised January 2003.

M. Yogo, "Estimating the Elasticity of Intertemporal Substitution When Instruments Are Weak," **Review of Economics and Statistics**, August 2004, Volume 86, pp. 797-810.

Table 1
Test for functional form of Utility Function

CIES Specification

Coefficient	Coefficient Estimate	Standard Error	Log Likelihood
θ	1.405733	0.759238	3955.479610

Logarithmic Specification

Coefficient	Coefficient Estimate	Standard Error	Log Likelihood
θ	1.000000	0.000000	3955.241279

Likelihood Ratio Test

Log LR	Chi-Square Critical	Probability
0.476662755	3.841459	0.489937854

Table 2
GMM Estimates of the Structural Parameter
(Logarithmic Preferences, theta=1)

Coefficient	Vbl	Coefficient Estimate	Standard Error	t-statistic (Significance)
ρ		0.076292	0.001463	52.130925
η		0.019080	0.001035	18.437830
α		0.339537	0.001035	328.109295
γ		0.016948	0.001035	16.377586
δ		0.043397	0.001035	41.936399
A		1229.779000	1.272611	966.343004
ω		0.649266	0.120959	5.367654
σ	ε	0.006577	NA	

Table 3
MLE Estimates of the Structural Parameters
(Logarithmic Preferences, theta=1, A=1229.779)

Coefficient	Vbl	Coefficient Estimate	Standard Error	t-statistic (Significance)
ρ		0.07402828	0.00093639	79.05707753
η		0.01735792	0.00060131	28.86684179
α		0.33985938	0.00025234	1346.82349234
γ		0.01693309	0.00002221	762.30766325
δ		0.04349558	0.00024570	177.02372557
ω		0.95077301	0.00553460	171.78714622
σ	ε	0.00575823	0.00046977	12.25745580
σ	lnc	0.03279583	0.00155105	21.14428077
σ	lni	0.14210606	0.00660598	21.51171494
σ	lny	0.00283251	0.00049226	5.75414613
σ	lnk	0.01132460	0.00052875	21.41764762
σ	r	0.00280249	0.00013741	20.39461849
σ	lnw	0.00720250	0.00040727	17.68486539

Log-Likelihood   3955.24127905

Table 4
MLE Estimates of the Structural Parameters
(Logarithmic Preferences, theta=1)

Coefficient	Vbl	Coefficient Estimate	Standard Error	t-statistic (Significance)
ρ		0.07538237	0.00136737	55.12964391
η		0.01786953	0.00069704	25.63619373
α		0.33954286	0.00033513	1013.17043588
γ		0.01693014	0.00002272	745.06405118
δ		0.04338134	0.00025628	169.27441138
A		1240.24863756	7.52407589	164.83733758
ω		0.94995459	0.00649150	146.33826273
σ	ε	0.00571032	0.00046669	12.23574766
σ	lnc	0.03278369	0.00155175	21.12692815
σ	lni	0.14210777	0.00660542	21.51381875
σ	lny	0.00285642	0.00048816	5.85143576
σ	lnk	0.01132671	0.00052968	21.38422630
σ	r	0.00280720	0.00013790	20.35701995
σ	lnw	0.00718239	0.00040690	17.65166579

Log-Likelihood   3956.22348772

Table 5
Test for difference in GMM and MLE Estimate of A

GMM Estimate

Coefficient	Coefficient Estimate	Standard Error	Log Likelihood
A	1229.77900000	1.27261127	3955.24127905

MLE Estimate

Coefficient	Coefficient Estimate	Standard Error	Log Likelihood
A	1240.24863756	7.52407589	3956.22348772

Likelihood Ratio Test

Log LR	Chi-Square Critical	Probability
1.964417337	3.841459	0.161041747

Table 6
Estimate of the Speed of Convergence in the Neoclassical Growth Model

Computed from the Structural Parameter Estimates Listed in Table 3

Coefficient	Coefficient Estimate	Standard Error	t-statistic (Significance)
β	0.12694562	0.00073583	172.52072419

95 Percent Confidence Interval Estimate

Coefficient	Lower Bound	Upper Bound
β	0.1255	0.1284

t-test (test of 2 percent hypothesis)

Coefficient	t-statistic	t-critical value (5 percent)	Probability
β	145.3404677	1.651841183	0.00000000

Computed from the Structural Parameter Estimates Listed in Table 4

Coefficient	Coefficient Estimate	Standard Error	t-statistic (Significance)
β	0.12802778	0.00107135	119.50153410

95 Percent Confidence Interval Estimate

Coefficient	Lower Bound	Upper Bound
β	0.1259	0.1301

t-test (test of 2 percent hypothesis)

Coefficient	t-statistic	t-critical value (5 percent)	Probability
β	100.8334705	1.651841183	0.00000000

Table 7
Likelihood Ratio Test of Hypothesis (beta <= 0.12)

Coefficient	Coefficient Estimate	Log Likelihood
β	0.12802778	3956.223488
β	0.12000000	3779.982131

Likelihood Ratio Test

Log LR	Chi-Square Critical	Probability
352.4827144	3.841459	0.00000000

t-test (test of 12 percent hypothesis)

Coefficient	t-statistic	t-critical value (5 percent)	Probability
β	7.493152797	1.651841183	0.00000000

Appendix A

A.1 Consumer Behavior

The dynamic optimization problem of the consumer is to:

$$(A1.1) \quad \underset{\widetilde{c}_t,\ldots,\widetilde{a}_1,\ldots}{\text{Max}} \quad E_0\left(\sum_{t=0}^{\infty} \overline{B}^t u(\widetilde{c}_t)\right) \quad \text{s.t.} \quad 0 \le E_t\left(w_t + (r_t + \delta - (\eta + \gamma))\widetilde{a}_t - \widetilde{c} - \Delta\widetilde{a}_t\right) \quad \forall \quad t = 0,\ldots,\infty$$

where: $\widetilde{a}_t$: The stock of adjusted per-capita assets in period t. $\forall$ $t = 1,\ldots,\infty$.

$\widetilde{c}_t$: The adjusted per-capita level of consumption in period t. $\forall$ $t = 0,\ldots,\infty$.

w_t : the real wage rate in period t. $\forall$ $t = 1,\ldots,\infty$

r_t : the real interest rate in period t. $\forall$ $t = 1,\ldots,\infty$

$u(\widetilde{c}_t) = \dfrac{\widetilde{c}_t^{1-\theta} - 1}{1-\theta}$: a CIES (Constant Intertemporal Elasticity of Substitution) function.

when $\theta = 1$, $u(\widetilde{c}_t) = \ln\widetilde{c}_t$

$\overline{B}$: is the discount rate, with $\overline{B} = \dfrac{1}{(1 + (\rho - \eta - (1-\theta)\gamma))}$,

where: $(\rho - \eta - (1-\theta)\gamma) > 0$.

$\rho, \eta, \gamma, \theta, \delta$ are the structural parameters of the problem, with ρ being the consumer's rate of time preference; η being the rate of growth of the population; γ being the rate of growth of Harrod-neutral technical change; θ being the coefficient of relative risk aversion, which is the inverse of the intertemporal substitution elasticity; δ being the rate of depreciation on physical capital.

A.2 Producer Behavior

The static optimization problem of the producer is to:

$$(A2.1) \quad \underset{\dot{k}_t}{\text{Max}} \quad \widetilde{\pi}_t = A\widetilde{k}_t^{\alpha}z_t - w_t - (r_t + \delta)\widetilde{k}_t = 0$$

where: $\widetilde{k}_t$: The stock of adjusted per-capita capital in period t. $\forall$ $t = 1,\ldots,\infty$.

$A\widetilde{k}_t^{\alpha}z_t$: is a stochastic Cobb-Douglas production function. The informational assumption is that the technology shock, z_t , is revealed at the beginning of period t.

A, α: are structural parameters. A is a parameter that sets of scale of output in the Cobb-Douglas production function; α being the output elasticity of capital.

r_t, w_t : is as defined above in the consumer's problem.

Taking the first-order condition for the producer yields the condition for the real interest rate:

(A2.2) $\dfrac{\partial \widetilde{\pi}_t}{\partial \widetilde{k}_t} = \alpha A \widetilde{k}_t^{-(1-\alpha)} z_t - (r_t + \delta) = 0$, which implies: $\alpha A \widetilde{k}_t^{-(1-\alpha)} z_t = (r_t + \delta)$

The zero profit condition gives the value for the real wage rate:

(A2.3) $w_t = A \widetilde{k}_t^{\alpha} z_t - \alpha A \widetilde{k}_t^{-(1-\alpha)} z_t \widetilde{k}_t$

Substituting four equilibrium conditions (the producer's first-order condition, the zero profit condition for the real wage rate, the demand and supply condition for capital/assets, and the savings/investment equilibrium condition) into the consumer's budget constraint provides the resource constraint for the economy. The connection between the general equilibrium problem above and the social planner's problem presented in Section 2 is the second welfare theorem of economics, which states that any Pareto efficient allocation that is obtained by the optimization of a government planning problem is equivalent to a competitive equilibrium. Combining the consumer's utility function (the social welfare function for the government planner) and the economy's resource constraint, which constrains the government planner's allocation choice, we obtain the dynamic optimization problem for the government planner that was presented in Section 2.

Appendix B

GMM Moment Estimating Conditions

In this Appendix, we present the moment equations utilized in the GMM estimation procedure to obtain a good starting point for the MLE estimation procedure. We have to estimate nine structural parameters, and we use eight moment equations to obtain a just-identified GMM estimation of the nine parameters. The moment equations are given by:

(1) To estimate the growth rate of the Labor Force, η, we use the following moment equation:

(B.1) $E\left[4\Delta \ln L_t - \eta\right] = 0$

(2) To estimate the growth rate of Harrod-neutral technology, γ, we use the following moment equation:

(B.2) $E\left[4\Delta\ln y_t - \gamma\right] = 0$

(3) To estimate the depreciation rate, δ, we use the following moment equation:

(B.3) $E\left[d_t - \dfrac{\delta}{4}\right] = 0$

where: $d_t = \dfrac{i_t^g}{k_{t-1}} - \dfrac{k_t - k_{t-1}}{k_{t-1}}$

(4) To estimate the output elasticity of capital in the Cobb-Douglas production function, α, we use the following moment equation:

(B.4) $E\left[\dfrac{(r_t + \delta_t)k_t}{y_t} - \alpha\right] = 0$

(5) To estimate the rate of time-preference, ρ, and the coefficient of relative risk aversion, θ, we use the Euler equation, using the current period real interest rate as an instrument in the equation:

(B.5) $E\left[4\Delta\ln c_t - \left(\gamma - \dfrac{(\rho + \theta\gamma)}{\theta\|1 + (\rho - \eta - (1-\theta)\gamma)\|}\right) - \dfrac{1}{\theta\|1 + (\rho - \eta - (1-\theta)\gamma)\|}r_t\right] = 0$

(6) To estimate the output scale coefficient in the Cobb-Douglas production function, A, we use the following moment equation;

(B6.) $E\left[\left(\ln y_t - \dfrac{\gamma}{4}t\right) - \alpha\left(\ln k_t - \dfrac{\gamma}{4}t\right) - \ln A\right] = 0$

(7) To estimate the autoregressive coefficient in the stochastic technology shock, we use the following moment equation:

(B.7) $E\left[\ln z_t - \omega\ln z_t\right] = 0$

where: $\ln z_t = \left(\ln y_t - \dfrac{\gamma}{4}t\right) - \alpha\left(\ln k_t - \dfrac{\gamma}{4}t\right) - \ln A$

(8) to obtain the variance of the innovation to the stochastic technology shock, we use the following moment equation:

(B.8) $E\left[(\ln z_t - \omega\ln z_t)^2\right] = \sigma_\varepsilon^2$