

IS THE SPEED OF CONVERGENCE CONSTANT?

Jordan Rappaport

DECEMBER 2000

RWP 00-10

Research Division
Federal Reserve Bank of Kansas City

Jordan Rappaport is an economist at the Federal Reserve Bank of Kansas City. The author is grateful for the advice and feedback of Alberto Alesina, Robert Barro, Francesco Caselli, Steven Durlauf, Edward Glaeser, David Laibson, Jeffrey Sachs, and Jonathan Willis. Michael Lamb provided excellent research assistance. The views expressed herein are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

Rappaport email: jordan.m.rappaport@kc.frb.org.

Abstract

Empirical attempts to measure the speed of convergence -- the rate at which a country's per capita income approaches its steady state relative to its distance from its steady state -- have started from the assumption that it is constant. In contrast, neoclassical models of capital accumulation usually predict that the speed of convergence decreases as income approaches its steady state. Estimating a flexible functional form which allows the speed of convergence to vary suggests that the speed of convergence actually increases as income approaches its steady state. An increasing speed of convergence calls into question structural interpretations of coefficients on conditioning variables in cross-sectional growth regressions. Instead, excluding initial income from cross-sectional growth regressions allows coefficients on exogenous variables to be interpreted as measuring changes in underlying structural relationships.

JEL Classification: O41, O47

Key words: Economic Growth, Income Convergence, Cross-Country Growth Empirics

1 Introduction

Is the speed of convergence — the rate at which income approaches its steady state relative to its distance from its steady state — constant? To date, the cross-country empirical literature has started from the assumption that it is. Conventional wisdom holds that developing countries' per capita income converge towards their steady state level at a constant rate of two percent per year. Alternate empirical methodologies suggest either higher or lower constant rates. But there is no theoretical reason to believe the speed of convergence is constant. On the contrary, linearizing a multiple sector model of capital accumulation around its steady state suggests that the speed of convergence will decrease as economies develop. Indeed, the difficulty calibrating a Ramsey-Cass-Koopmans growth model to avoid such a decreasing speed of convergence has served as the basis for rejecting the adequacy of neoclassical models of capital accumulation to describe economic development (King and Rebelo, 1993).

Generalizing the cross-sectional empirical growth specification to allow for a varying convergence speed, the present paper argues that the speed of convergence is not constant; but nor is it decreasing as predicted by a linearization. Instead, the empirics suggest that the speed of convergence increases as an economy grows towards its steady state. All else equal, poor economies grow quicker than rich ones; but the poor economies' faster growth is less than proportional to their extra distance from their steady state.

The centrality of transitional dynamics within the neoclassical growth framework of capital accumulation underscores the importance of empirically describing how convergence speed behaves. The welfare implications of transitional dynamics may seem to be swamped by the welfare implications of the determinants of productivity and technology. But for poorer countries not on the technological frontier, the distinction between capital accumulation and technological development is arbitrary: technology can be interpreted as just another type of capital to be accumulated. Such an interpretation suggests that models of capital accumulation should be able to more successfully capture development phenomena than previously thought. Kremer and Thomson (1998), for instance, introduce complementarity among different types of capital to slow the speed of convergence. Similarly, Rappaport (2000b) introduces a capital adjustment cost more convex than the typical quadratic speci-

fication to achieve a reasonably accurate calibration of the Ramsey-Cass-Koopmans model.

A second motivation for the present paper is to better understand what we learn from cross-sectional growth regressions. With a constant convergence speed, regressing country growth rates on initial per capita income and a vector of exogenous country attributes allows for a structural interpretation of the coefficients on the latter. For example, the common finding of a statistically significant, positive coefficient on various measures of education can be interpreted as implying a positive correlation between education and steady-state per capita income. The partial correlation of education with growth comes only because controlling for initial income, countries with higher levels of education are further below their steady-state level of income. But if the speed of convergence varies, the structural interpretation breaks down: a positive correlation between education and per capita income growth need not imply a positive correlation between education and steady-state per capita income. By excluding initial income as a conditioning variable in cross-sectional growth regressions, coefficients on exogenous variables can more sensibly be interpreted as measuring *changes* in underlying structural relationships.

The paper proceeds as follows: Section 2 formalizes some language for discussing a varying speed of convergence and then reviews past research efforts to measure a constant convergence speed. Section 3 briefly lays out the theoretical case that we should expect the speed of convergence to vary. Section 4 nests a constant speed of convergence in an empirical specification which allows convergence speed to vary. While the null hypothesis of a constant convergence speed cannot be rejected, nonlinear-least-squares point estimates from each of three different specifications find the speed of convergence to increase from less than 1 percent per year for a country with income far below its steady-state level to above 4 percent per year as a country approaches its steady state. The finding of an increasing speed of convergence is robust in the sense that the likelihood surface is smooth with respect to the coefficients parameterizing the speed of convergence; and large swathes of the parameter space characterized by an increasing convergence have a likelihood above the constant-convergence-speed level. Section 5 concludes.

2 Background: The Speed of Convergence

A natural starting point is to formally define the speed of convergence:

$$\Lambda(\log y(t) - \log y^*(t)) \equiv \frac{-\frac{d}{dt}(\log y(t) - \log y^*(t))}{\log y(t) - \log y^*(t)}$$

Here, the numerator measures the rate at which income, $y(t)$, moves towards its contemporary steady-state value, $y^*(t)$. The denominator measures the distance of income from this steady-state. As defined, the speed of convergence will be positive so long as income is indeed moving towards its steady state.

Consider the special case where the speed of convergence does not depend on income's distance from its steady state:

$$\Lambda(\log y(t) - \log y^*(t)) = \lambda \tag{1}$$

Substituting the definition for the speed of convergence and rearranging gives,

$$\frac{d}{dt}(\log y(t) - \log y^*(t)) = -\lambda(\log y(t) - \log y^*(t)) \tag{2a}$$

$$\frac{\frac{d}{dt}y(t)}{y(t)} = -\lambda(\log y(t) - \log y^*(t)) + \frac{\frac{d}{dt}y^*(t)}{y^*(t)} \tag{2b}$$

The first of these forms, (2a), also emerges from the linearization of a one sector neoclassical growth model around its steady state. The latter form, (2b), structurally underlies most cross-sectional growth regressions. Assuming $\log y^*$ to be measured by $\mathbf{x}'\boldsymbol{\beta}$, a linear combination of exogenous attributes, then by regressing growth rates on initial income and a vector of such attributes, a constant speed of convergence will be measured by the coefficient on initial income (making an adjustment for the fact that the above equations hold in continuous time whereas cross-sectional growth empirics are based on discrete time observations).

Considerable effort has gone into measuring such a constant speed of convergence. A first main result is that failure to control for attributes which vary across countries and hence which may imply different steady states leads to non-significant estimates of λ : looking across countries, we do not see *unconditional* income convergence (Barro, 1991; Barro and Sala-i-Martin, 1992, 1995; Sala-i-Martin, 1994). But after controlling for a small set of country attributes (e.g., education levels, life expectancy, market distortions), λ is usually estimated

in the neighborhood of 0.02. Hence the conventional wisdom that countries converge towards their steady-state income levels at a roughly 2 percent annual rate.

Several methodological problems, however, caution against inferring the magnitude of a constant convergence speed from the coefficient on initial income. Such problems include the failure to control for unobservable country attributes, the endogeneity of conditioning regressors, mismeasurement of initial income, and finite sample bias (e.g., Ales and Glaeser, 1994; Islam, 1995; Caselli, Esquivel, Lefort, 1997; Evans, 1997; Quah, 1997). In an effort to address some of these, Islam (1995) and Caselli, Esquivel, Lefort (1997) suggest a panel data methodology which yields estimates of λ in the range of 0.04 to 0.10. As with the conventional wisdom, however, the panel methodology starts from the assumption of a constant convergence speed. Alternative methodologies proposed by Evans (1997) and Lee, Pesaran, and Smith (1998) allow the speed of convergence to differ across economies but continue to impose that these different speeds remain constant along transition paths. The present author is unaware of any research which explicitly tests whether economies' speed of convergence remains constant.

A more general critique of conditional convergence is that it imposes no behavioral restrictions on cross-country growth data: for any pattern of growth among countries, there exists a set of assumed steady-state income levels which can “explain” the growth as arising from conditional convergence (Durlauf and Quah, 1999).

3 Neoclassical Theory: The Speed of Convergence Varies

Given its importance to cross-sectional growth empirics, an identifying assumption that the speed of convergence is constant might be expected to derive from neoclassical growth theory. On the contrary, neoclassical growth theory has tremendous difficulty admitting a constant convergence speed; instead, most calibrations of neoclassical growth models suggest that the speed of convergence is high for economies far below their steady states and then decreases as economies develop. For many calibrations, the growth rates predicted for developing economies are simply implausible.

Of course, “difficulty” in calibrating a neoclassical growth model to match a constant

speed of convergence is not the same as “impossibility”. On the contrary, Rappaport (2000b) shows that introducing a convex, increasing *average* adjustment cost to gross capital investment allows the neoclassical model to admit a constant and even an increasing speed of convergence.¹ The point is, neoclassical theory in no way implies a constant convergence speed.

One way by which neoclassical theory falsely might seem to imply a constant convergence speed is linearizing a growth system around its steady state in the very special case in which there is only a single state (i.e., “non-jumping”) variable. Such a linearization indeed suggests that income grows at a speed proportional to its distance from its steady-state level and hence that the speed of convergence is constant. But as soon as we allow for a growth system in which there are multiple linearly independent state variables (e.g. physical and human capital), linearizing around the steady state suggests that the speed at which any of the state variables converges towards its steady state will be decreasing as countries develop. (See Appendix A; a similar point is made by Eichner and Turnovsky, 1999.)

This last result should not be interpreted as suggesting that the actual speed of income convergence decreases as countries develop. Rather it cautions against using linearizations to describe behavior of growth systems away from their steady state. Rappaport (2000a), for instance, shows that extending the neoclassical framework to allow for labor mobility causes the speed of convergence to vary widely in a neighborhood very close to the steady state; the asymptotic speed of convergence turns out to be essentially irrelevant for describing the system’s behavior.

More generally, the behavior of the speed of convergence depends quite closely on what is meant by a “steady state”. From the perspective of dynamic theory, the defining characteristic of a steady state is that the endogenous forces contributing to normalized capital accumulation and decumulation just offset each other. Moving from theory to empirics, such a definition admits multiple interpretations given the arbitrary modeling assumptions separating endogenous from exogenous forces. For example, a low steady-state per capita level of a first type of capital good may be implied by a fixed low level of a second type of capital

¹Such an assumption is stronger than the convex *total* adjustment costs more generally used (e.g. quadratic total adjustment costs).

good; but high steady-state per capita levels of both capital goods may follow by allowing the second capital good to vary. In estimating the speed of convergence of the first capital good — e.g. using (2b) — whether or not one should control for the level of the second capital good depends on the interpretation one wants to impart.²

Finally, note that interpreting a steady state as a long run equilibrium which is actually attained, then asymptotically it must be approached at an infinite speed and so the speed of convergence must eventually be increasing. However in practice, a moderate constant speed of convergence is sufficient for a stock variable to very nearly attain its steady state level within a reasonable timespan. For instance, a stock variable which converges at a constant ten percent speed takes approximately 37 years to go from 50 percent to 99 percent its steady state level. At a constant two percent speed of convergence, covering the same distance takes approximately 185 years.

4 Empirics: An Increasing Speed of Convergence

The empirical analysis begins by specifying a flexible functional form which allows the speed of convergence to be either constant or varying. Numerical examples of the growth rates implied by “high” and “low” constant convergence speeds are then compared to the actual growth rates experienced by some selected nations. A heuristic argument is made that even the highest actual growth rates experienced by relatively poor nations are well below those predicted by a high constant convergence speed. Conversely, the high actual growth rates experienced by numerous moderate income nations are well above those predicted by a low constant convergence speed. The section then proceeds to a more formal estimation of the flexible functional form using several different ways to estimate steady state income. While all point estimates indicate an increasing speed of convergence, a constant speed of convergence cannot be rejected. A last subsection, however, argues that starting from neutral priors, the posterior odds strongly support that the speed of convergence is increasing.

²The “second” capital good most obviously corresponds to human capital, which is usually included as an independent variable in cross-sectional growth regressions. More expansively, the intangibles in which a society collectively invests such as “technology” and “rule of law” also can be interpreted as capital goods. Of course, including endogenous capital goods as controls in growth regressions causes econometric problems.

4.1 A Flexible Functional Form

To nest the possibility of a constant speed of convergence within a more general framework, The speed of convergence is assumed to be characterized by,

$$\Lambda (\log y (t) - \log y^* (t)) = (\lambda_r - \lambda_p) \cdot e^{-\kappa|\log y(t) - \log y^*(t)|} + \lambda_p \quad (3)$$

This implies a growth rate,

$$\frac{\frac{d}{dt}y(t)}{y(t)} = - \left((\lambda_r - \lambda_p) \cdot e^{-\kappa|\log y(t) - \log y^*(t)|} + \lambda_p \right) \cdot (\log y(t) - \log y^*(t)) + \frac{\frac{d}{dt}y^*(t)}{y^*(t)} \quad (4)$$

The speed of convergence for a country far below its steady state is measured by λ_p while the speed of convergence for a country close to its steady state is measured by λ_r . Intuitively, these correspond to the speed of convergence when “poor” and “rich”, respectively. (Where “poor” and “rich” describe a country’s output *relative* to its steady state level.) The rate at which the speed of convergence transitions from λ_p to λ_r , is measured by $\kappa > 0$.

Note that *there is no theoretical basis for assuming the data generating process underlying country development can be characterized by (3)*. Rather, the main advantage of (3) is its flexibility. In the special case where $\lambda_p = \lambda_r$, (3) reduces to (1) with λ as above measuring a constant speed of convergence everywhere. More generally, $\lambda_p > \lambda_r$ captures a speed of convergence which decreases as income approaches its steady state and $\lambda_p < \lambda_r$, a speed of convergence which increases as income approaches its steady state. The implied sign convention is summarized in Table 1.

The parameter κ introduces even greater flexibility. As κ decreases to zero, λ_p applies only to countries with essentially zero output. Conversely, as κ increases without bound, λ_r applies only to countries which are asymptotically approaching their steady state. In fact, an unbounded κ makes (3) too flexible for the purposes of estimation. For countries neither completely destitute nor immediately adjacent to their steady state, an approximately constant convergence speed, $\bar{\lambda}$, is equivalently described by λ_p and λ_r both close to $\bar{\lambda}$ (regardless of the value of κ); by λ_p close to $\bar{\lambda}$ and κ large (regardless of the value of λ_r); and by λ_r close to $\bar{\lambda}$ and κ close to zero (regardless of the value of λ_p). To avoid such indeterminacy and so insure that λ_p not equal to λ_r actually meaningfully implies a varying convergence speed, the estimation below constrains κ to lie in the interval [1, 5].

A second advantage of (3) is its relative parsimony. Two parameters, λ_p and λ_r , are the minimum necessary to allow a speed of convergence which can vary. The third parameter, κ , is necessary to avoid imposing a rigid functional form on convergence speed transitional dynamics. In the interest of parsimony, (3) is specified so that the speed of convergence is symmetric with respect to income’s distance from steady state (i.e. the absolute value in the exponential term). But it is easy to imagine that a more accurate representation would allow the behavior of the speed of convergence to depend on whether income is below or above its steady state.

A third advantage of (3) is its relative stability. The implied growth rate, (4), turns out to be highly nonlinear with respect to the parameters determining $\log y^*$. But with sufficient computational power, the parameters λ_p , λ_r , and κ in (4) can be estimated. Alternative candidate descriptions of transitional convergence speed behavior proved to have even greater instability and so could not be estimated.³

4.2 Heuristic Argument

To get a more intuitive sense of magnitude, Figure 1 shows speeds of convergence and implied growth rates under alternative assumptions for the values of λ_p , λ_r , and κ in (3). The bottom locus in each of the panels corresponds to a “low” constant speed of convergence of 2 percent per year (the conventional wisdom value). The top locus in each of the panels corresponds to a “high” constant speed of convergence of 10 percent per year (the value estimated by Caselli, Esquivel, and Lefort, 1997). The loci in between these low constant and high constant time paths correspond to a speed of convergence which varies from the former to the latter. Thus for both of these middle loci, $\lambda_p = 0.02$ and $\lambda_r = 0.10$. The middle loci differ in how quickly the speed of convergence transitions from λ_p to λ_r as calibrated by κ . The upper middle locus shows the gradual transition implied by $\kappa = 1$ while the lower middle locus shows a “delayed” but then more rapid transition implied by $\kappa = 5$.

As shown by the top locus in Panel B, a high constant speed of convergence implies growth rates for nations poor relative to their steady state which seem “too high”. With an 10 percent constant speed of convergence, a nation with income 10 percent its steady state

³For instance, $\Lambda = (\lambda_p \cdot (\log y - \log y^*) - \kappa \lambda_r) / ((\log y - \log y^*) - \kappa)$.

Table 1: Sign Conventions for Convergence Speed
Change in Convergence Speed as Income Approaches Steady State

convention	speed when “poor” vs. when “rich”	derivative with respect to distance from steady state	example
decreasing	$\lambda_p > \lambda_r$	$\frac{\partial \Delta}{\partial \log y(t) - \log y^* } > 0$	traditional neoclassical growth models with linear increasing average adjustment costs
constant	$\lambda_p = \lambda_r$	$\frac{\partial \Delta}{\partial \log y(t) - \log y^* } = 0$	identifying assumption cross-country growth empirics
increasing	$\lambda_p < \lambda_r$	$\frac{\partial \Delta}{\partial \log y(t) - \log y^* } < 0$	generalized neoclassical growth model with convex increasing average adjustment costs (Rappaport, 2000b)

level would be expected to grow 23 percentage points per year faster than a nation already at its steady state. For a nation with income 20 percent of its steady-state level, the additional expected growth would be 16 percentage points per year.

To allow for comparison against such benchmark growth rates, let the level and growth rate of per capita income in the United States proxy for all nations’ steady-state income and rate of technological progress. The economy which comes closest to sustaining a high-convergence-speed growth rate is Singapore over the period 1966 to 1975 during which time its per capita income grew 9.5 percentage points higher than that of the U.S. starting from an initial income 17 percent the U.S. level. Next closest is Japan over the period 1955 to 1964, with an average annual rate of per capita income growth 6.9 percentage points higher than that of the U.S. from an initial relative income 21 percent the U.S. level. In addition, a handful of nations with initial incomes between 10 and 15 percent the U.S. level have been

able to sustain five-year (or longer) average annual growth rates between 6 and 7 percentage points faster than that of the U.S. including South Korea (1968-1974), Taiwan (1969-1974), Brazil (1970-1975), and Thailand (1986-1992).⁴

Even these “success” stories, however, fall far short of the predicted growth rates with a a 10 percent constant speed of convergence. Interpreting actual growth rates as draws from some random distribution, the *expected* growth rates for poor nations should fall somewhat below these highest observations. Observed growth rates would therefore seem to place an upper bound on the speed of convergence for poor nations well below 10 percent.

Alternatively, one can argue that we never observe countries at income levels far below their steady state. If in fact the high-growth countries just listed have steady state per capita income levels significantly below that of the United States, the heuristic argument against a high constant speed of convergence fails. However, as discussed in the previous section, the very notion of a “steady state” is subject to ambiguity given arbitrary assumptions distinguishing exogenous from endogenous forces. In the present case, arguing that Singapore in 1966 or Japan in 1955 had steady state income levels significantly below that of United States *given* their contemporary economic conditions does not make sense. The subsequent growth by these countries to per capita income levels approaching that of United States surely belies the fallacy of taking their contemporary economic conditions as given. Instead, contemporary observables which can be interpreted as suggesting low steady state incomes are more likely to be indicative of low *current* incomes.

Equally evident from Figure 1 Panel B is that for nations which are even moderately close to their steady state, the growth rates implied by a low constant speed of convergence seem “too low”. With a 2 percent constant speed of convergence, a nation with income 50 percent its steady-state level would be expected to grow only 1.4 percentage points per year faster than a nation already at its steady state. For a nation with income 60 percent of steady state, the additional growth would be only 1.0 percentage points per year.

In contrast, numerous countries with initial incomes between 50 and 60 percent the U.S. level have been able to sustain five-year or longer average annual growth rates from 3 to 6 percentage points faster than that of United States including Venezuela (1951-1958), West

⁴All growth rates here and below are based on Summers and Heston (1995).

Germany(1954-1962), Denmark (1955-1962), France (1956-1961), Finland(1967-1974), Iceland (1970-1978), Trinidad and Tobago (1973-1980), the Bahamas (1977-1983), Hong Kong (1978-1984), and Singapore (1985-1992). So many potential counterexamples are unlikely to be explained by chance. Assuming these countries have steady-state income levels below that of United States only worsens the fit of a constant 2 percent speed of convergence.⁵

4.3 Empirical Implementation

To test the null hypothesis of a constant speed of convergence, I use cross-country data to estimate the nonlinear equation,

$$\begin{aligned} \frac{\log y_{i,t+T} - \log y_{i,t}}{T} = & - \left((\tilde{\lambda}_r - \tilde{\lambda}_p) \cdot \exp \left(\tilde{\kappa} \left| \log y_{i,t} - \left(\beta_{0,t} + \mathbf{x}'_{i,t} \boldsymbol{\beta} \right) \right| \right) + \tilde{\lambda}_p \right) \\ & \cdot \left(\log y_{i,t} - \left(\beta_{0,t} + \mathbf{x}'_{i,t} \boldsymbol{\beta} \right) \right) + \frac{d \log y^*(t)}{dt} + \varepsilon_{i,t} \end{aligned} \quad (5)$$

The term, $\beta_{0,t} + \mathbf{x}'_{i,t} \boldsymbol{\beta}$, is assumed to measure steady-state income, $\log y^*_{i,t}$. The time subscript on steady-state income captures, first, that this is growing by the rate of exogenous technological progress, $\frac{d \log y^*(t)}{dt}$, and second, the possibility of exogenous changes in the country attributes, $\mathbf{x}_{i,t}$, which determine steady-state income. Imposing additional structure on (5) so as to minimize the number of coefficients which must be estimated, the level intercept term, $\beta_{0,t}$, is assumed to be linked across time by technological progress,⁶

$$\beta_{0,t+T} = \beta_{0,t} + \frac{d \log y^*(t)}{dt} \cdot T \quad (6)$$

Finally, I assume the intercept term, $\varepsilon_{i,t}$, is independent in cross sections but allow for a country random effect when estimating standard errors using stacked cross sections.

$$\begin{aligned} E(\varepsilon_{i,t}) &= 0 \\ E(\varepsilon_{i,t} \varepsilon_{j,\tau}) &= 0 \text{ for } i \neq j \end{aligned}$$

⁵An alternative explanation which reconciles such examples with a slow constant speed of convergence is that these countries have steady-state income levels significantly above the contemporary U.S. level. Such an explanation would seem plausible only if U.S. income itself were below its steady-state level; if so, U.S. growth over the respective periods would largely be due to convergence rather than technological progress.

⁶Note that this restriction does not necessarily follow from neoclassical theory: to the extent that technological progress implies capital deepening, the growth rate of steady-state income may exceed that of actual income.

Table 2: Specification Explanatory Variables

“standard”	“exogenous”
average secondary education	% population within 100km of coast
log(life expectancy at birth)	% population in moderate climate
% population in 20°N to 20°S latitude	% population in 20°N to 20°S latitude
	log (distance to nearest of Rotterdam, Tokyo, New York)

An additional “unconditional” specification uses initial income as only explanatory variable but limits observations to the 22 OECD nations as of 1973 excluding Greece and Turkey. All three specifications include time dummies. Growth rates and initial incomes are from Summers and Heston (1995), Penn World Tables, Version 5.6. Remaining “standard” explanatory variables are from Barro and Lee (1994). Remaining “exogenous” variables are from Bloom and Sachs (1998).

The specification, (5), corresponds exactly to a discrete time implementation of (4) with $\tilde{\lambda}_p$, $\tilde{\lambda}_r$, and $\tilde{\kappa}$ as the discrete time analogs of λ_p , λ_r , and κ . When λ_p equals λ_r ,

$$\lambda = \left(\frac{1}{T}\right) \log(1 + \tilde{\lambda}T) \quad (7)$$

For λ_p not equal to λ_r , however, no such formula links the discrete and continuous time convergence coefficients. Hence caution is warranted in using the point estimates reported below to calibrate a continuous time model.⁷

The fit of (5) as measured by the sum of squared residuals is highly nonlinear with respect to the coefficients, β , so that estimation is greatly facilitated by a parsimonious set

⁷This “discrete time bias” arises because the single growth rate measured over a discrete interval, t to $t + T$, in fact is an average of a varying growth rate. For a constant speed of convergence, $\lambda_p = \lambda_r = \lambda$, the coefficient on initial income, $\tilde{\lambda}$, effects this average growth rate at the beginning-of-interval income. But growth due to convergence will be faster at t than at $t + T$. Holding constant the combined conditioning coefficients, $\lambda\beta = \tilde{\lambda}\tilde{\beta}$, then $\lambda < \tilde{\lambda}$ effects this faster growth at the beginning of the interval and higher “initial” income later in the interval effects slower growth. So for instance, with growth measured over a ten-year period as in the estimation below, (7) gives that $\tilde{\lambda} = 0.020$ implies $\lambda = 0.018$ and that $\tilde{\lambda} = 0.080$ implies $\lambda = 0.059$. With a varying speed of convergence, $\tilde{\lambda}_p$ is additionally biased towards λ_r given the ongoing transition from λ_p to λ_r over the course of the discrete interval.

of country attributes, $\mathbf{x}_{i,t}$.

As summarized in the first column of Table 2, a first “standard” specification controls for varying country steady-state incomes using average secondary education, life expectancy at birth, and the percentage of a country’s population residing between the 20° North and 20° South latitude bands. Variations on the first two of these tend to be included in the base specification of most cross-sectional growth regressions. The latitude variable is included because of a relatively high marginal contribution to explanatory power.

Returning to the methodological issues discussed in the background section above, a first problem estimating (5) using the standard specification is the potential endogeneity of the conditioning variables. For the most part, the concern has been of a reverse causal link from income growth to any of the “flow” conditioning variables (i.e. right-hand-side variables which can “jump”, such as life expectancy). The most common approach to control for such endogeneity is to instrument using lagged values. Of equal concern, however, is the potential for a reverse causal link from the current income level to any of the “stock” conditioning variables (i.e., right-hand-side variables constrained to a finite time derivative). To the extent that an included right-hand-side stock variable is a normal good, its level will increase with income; education and public capital seem obvious examples. The persistence of stock variables along with optimization by forward-looking agents rule out using lagged values as instruments.

Estimating steady-state income using variables for which the main direction of causality is from the level of current income to the level of the variable may cause countries to appear falsely close to their steady states. To see this, suppose an underlying data generating process characterized by an increasing speed of convergence ($\lambda_p < \lambda_r$). For a cross section of countries with different current incomes but all sharing the same steady-state income, including in $\mathbf{x}_{i,t}$ a variable which depends positively on the level of current income provides exactly the mechanism by which (5) can incorrectly find that $\lambda_p = \lambda_r$. Let k_2 be a function $f(\cdot)$ of current income y . For a country with current income y_r close to its steady-state level y^* , $f^{-1}(k_2(y_r))$ will reasonably accurately measure steady state income. But for a different country with current income y_p far below an identical steady-state-level, $f^{-1}(k_2(y_p))$ will falsely measure a low steady state income. For a country growing at rate γ , the closer it is

estimated to be to its steady state, the higher will be its estimated speed of convergence, Λ . So (5) will tend to estimate that the “poor” country is converging at rate $\hat{\lambda}_p > \lambda_p$ towards $f^{-1}(k_2(y_p))$ and that the “rich” country is converging at rate $\hat{\lambda}_r \approx \lambda_r$ towards $f^{-1}(k_2(y_r))$.⁸

Endogeneity in levels suggests estimating (5) using time invariant country attributes as controls. As summarized in Table 2 Column 2, a second “exogenous” specification controls for varying country steady-state incomes using the percentage of a country’s population living within 100 kilometers of an ocean coast, living in moderate climate zones, living between the 20° North and 20° South latitude bands as well as the logarithm of the distance from the country’s capital city to the nearest of Rotterdam, Tokyo, or New York.⁹

A great advantage of the exogenous specification is that it allows for regressing the current level of income on the exogenous attributes without the possibility of reverse causality. Assuming that deviations of countries’ current income from its steady-state level is randomly distributed, such a “level” regression directly estimates the structural determinants of steady-state income. Such level estimates can then be used to impose behavioral restrictions when estimating (5).

An alternative strategy of addressing the endogeneity in levels is to estimate (5) without controlling for any country attributes (other than initial income) but limiting the observed growth rates to those of countries which might reasonably be believed to share the same steady-state income level. This third “unconditional” approach is implemented using the OECD nations as of 1973 less Greece and Turkey.¹⁰

Regardless of specification, a remaining problem estimating (5) is the failure to control for unobserved country effects. Caselli, Esquivel, Lefort (1996) argue that (5) be augmented

⁸On the other hand, control variables may depend more on steady-state income than on current income, $k_i = k(y_i^*)$. Including endogenous variables of this type as controls for the purposes of estimating λ_p and λ_r may be desirable.

⁹These variables are drawn from Bloom and Sachs (1998). The temperate climate zone designation is based on the Koeppen Geiger classification system. In particular, it combines the temperate dry summer, temperate without dry season, and cold without dry season zones (designated “cf”, “cs”, and “df” respectively). For a further description, see www.fao.org/waicent/faoinfo/sustdev/EIdirect/climate/EIsp0002.htm.

¹⁰Including Greece and Turkey, relatively poor OECD nations which experienced slow growth, strengthens the increasing speed of convergence results reported below.

to include in addition to $\frac{d \log y^*(t)}{dt}$ and $\varepsilon_{i,t}$, a country-specific intercept term η_i .¹¹ Then with a constant speed of convergence ($\lambda_p = \lambda_r$), the $\hat{\lambda}$ estimate of $\tilde{\lambda}$ will be biased towards zero. Without disagreeing with the validity of this problem, herein it is simply assumed that conditioning regressors fully account for all cross-country variation in steady states so that the residual, $\varepsilon_{i,t}$, indeed satisfies the necessary orthogonality condition.

4.4 Results

For each of the three specifications, I run an OLS “level” regression of per capita income on the various attributes (including appropriate time dummies), an OLS growth regression of per capita income growth on the various attributes conditional on initial income, an OLS growth regression of per capita income growth on the various attributes excluding initial income, and an NLS estimate of (5). Results are reported in Tables 4 (standard specification), 5 (exogenous specification), and 6 (unconditional specification).

4.4.1 Standard Specification

Table 4 Column 1a reports the partial correlations between the logarithm of per capita income and the standard set of controls. Unsurprisingly, per capita income is positively correlated with each of life expectancy at birth and average secondary education and negatively correlated with the percentage of the population living in the 20° North to 20° South latitude bands. Except possibly for the lattermost relationship, certainly no inference of causality is warranted. Together with separate intercept terms for each of 1960, 1970, and 1980, the standard right-hand-side variables account for more than 81 percent of the variation in the level of per capita income.

Table 4 Column 2a reports corresponding partial correlations of the growth rate of per capita income with initial income and each of the standard set of controls. The estimated coefficient on initial income, -0.019, corresponds to a constant speed of convergence of 1.9 percent per year (or 1.7 percent adjusting for the discrete time bias) and so is consistent with the conventional wisdom of a 2 percent speed of convergence. The coefficients on the

¹¹Their discussion, however, suggests that what they have in mind is more a country-specific steady-state level intercept term, $\beta_{0,i}$.

standard control variables are all identical in sign to their counterparts in the level regression. Column 2b reports the results from dividing the coefficients in Column 2a through by 0.019 to recover structural parameters. The implied partial correlations between steady-state income and each of life expectancy at birth and average secondary education are approximately the same magnitude as the level partial correlations in Column 1a. The implied negative partial correlation between steady-state income and the percentage of the population living in the tropical latitudes is somewhat larger than the negative partial correlation in the level regression.¹²

Table 4 Column 3 reports the results from regressing per capita income growth on the standard controls while excluding initial income. Such a specification removes any justification for interpreting coefficients as measuring steady-state structural relationships. But to the extent that a right-hand-side variable is indeed exogenous, the coefficient on it can be interpreted as indicating the sign (and possibly magnitude) of the change in an underlying structural relationship. Given the very high statistical significance of the coefficient on initial income reported in Column 2a, unsurprisingly excluding initial income causes a large drop in R^2 from 0.382 to 0.287. But the coefficients which are statistically significant in the regression including initial income remain so when excluding initial income. The coefficient on percentage population in tropical latitudes additionally remains similar in magnitude. So the regression reported in Column 2a can be interpreted as implying a negative partial correlation between steady-state income and tropical latitude; the regression reported in Column 3 finds a nearly identical negative partial correlation between growth and tropical latitude. There is no reason to expect these two partial correlations to be similar. Which of course suggests that the coefficients from the two specifications are measuring the same underlying relationship.

Table 4 Column 4 reports the nonlinear least square estimates of the structural coefficients in (5) using the standard specification set of controls. Because the sum of squared residuals from (5) is highly nonlinear with respect to the coefficients, a computationally intensive modified grid search is used. Doing so requires bounding above and below the range

¹²The OLS regression does not allow for separate identification of the rates of technological progress, $\frac{dy^*(t)}{dt}$, from the initial level intercept of steady-state income, β_0 , even after imposing (6).

for the various coefficients. For the structural coefficients linking the standard attributes to steady-state income, these bounds are chosen to create the smallest possible convex interval including both the three-standard-deviation interval around the OLS level coefficients reported in Column 1a and the two-standard-deviation interval around the structural parameters recovered from the OLS growth regression reported in Column 2b. Decade-specific growth due to technological progress is constrained to lie between 0 and 4 percent per year. And the coefficient capturing the transition from the speed of convergence when poor to the speed of convergence when rich, κ , is constrained to lie between 1 and 5 (the same bounds shown in Figure 1). Where a lower or upper bound binds, reported coefficients include an “l” or “u” subscript.

The standard-specification NLS point estimate suggests that the speed of convergence increases from $\hat{\lambda}_p = 0.009$ to $\hat{\lambda}_r = 0.040$ with the transition constrained to its lower boundary, $\kappa = 1$ (i.e., the fit would be improved by allowing an earlier and more rapid transition from λ_p to λ_r).¹³ The coefficients on life expectancy, secondary education, and the tropical latitude population are similar in magnitude to those from the structural parameters implied by the OLS growth estimate. Technological progress, $\frac{dy^*(t)}{dt}$, is estimated at its upper bound value of 4 percent per year for the 1960s, at 2.6 percent per year for the 1970s, and at its lower bound value of 0 for the 1980s.

Table 4 also reports asymptotic standard errors for the NLS point estimates; these should be viewed with considerable skepticism. Especially problematic is that both the transition parameter, κ , and the initial-period level intercept, $\beta_{0,t0}$ are not identified under the null hypothesis that $\lambda_p = \lambda_r$. As a result, the NLS standard errors are consistent only under the alternative hypothesis, $\lambda_p \neq \lambda_r$, and so are not appropriate for testing the null hypothesis of a constant convergence speed (Davies, 1977, 1987; Andrews, 1993; Hansen, 1996). Even assuming a varying speed of convergence, the relatively small sample size suggests additional skepticism of the asymptotic-based standard error estimates.¹⁴ So while the point estimate

¹³Note that the “~” designation for the discrete time estimation of continuous time variables has been dropped.

¹⁴For $y = X(\beta)$, the estimated standard errors are calculated as the square root of the diagonal elements of $s^2 \cdot (\mathbf{J}'\mathbf{J})^{-1}$ where \mathbf{J} is the N-by-K Jacobian partial derivative matrix $\frac{\partial \mathbf{y}}{\partial \beta}$ evaluated at the NLS-estimated parameter values and s^2 is the estimated variance of an underlying i.i.d. observation error term. Asymp-

of the change in convergence speed reported at the bottom of the table is large relative to its asymptotic standard error, the inference that $\lambda_r > \lambda_p$ does not follow. Rather, Monte Carlo simulation imposing a data generating process equivalent to the OLS growth relationship estimated in Column 2a suggests a p-value of 0.395 (i.e., with $\lambda_p = \lambda_r$, 39.5 percent of the time the NLS algorithm will estimate $d\lambda = (\lambda_r - \lambda_p) \geq 0.031$).

The Monte Carlo simulation implies that the NLS point estimate of an increasing convergence speed may derive either from an underlying data generating process which is indeed so or from a linear data generating process along with randomness. This low power against the null of a constant convergence speed is not surprising with the inclusion of life expectancy at birth and average secondary school education as right-hand-side variables, both of which may very plausibly depend on the current level of income. As discussed in the previous subsection, estimating steady-state income using an attribute which endogenously depends on the current level of income provides exactly the mechanism to make an increasing convergence speed DGP falsely appear to be a constant convergence speed DGP.

A different source of doubt about the increasing convergence speed NLS point estimate which can be rejected is that it derives from some idiosyncratic property of the estimation procedure itself. Because (5) is highly nonlinear with respect to its parameters, the point estimate might represent a single, narrow “peak” of a likelihood function which otherwise suggests a constant convergence speed. On the contrary, Figure 2 shows the likelihood surface estimated with the standard specification to be smooth over a broad range of λ_p - λ_r space. In particular, the R^2 contours shown in Figure 2 Panel B define a large swath of parameter space with likelihood above its OLS constant convergence speed level. (For a given set of observed growth rates, R^2 maps one-to-one with the likelihood function.) Moreover, this swath lies entirely in the region defining an increasing speed of convergence, $\lambda_p < \lambda_r$. Figure 2 Panel C, however, shows that as λ_r increasingly exceeds λ_p , the the rise in explanatory power is small.

totically, these will converge to the true standard errors of the estimated parameter values. (Davidson and MacKinnon, 1993, Chapter 5)

4.4.2 Exogenous Specification

The exogenous specification estimates make a stronger case for an increasing speed of convergence. Table 5 Column 1a reports partial correlations between the level of per capita income and the exogenous specification set of controls. Positive coefficients on coastal proximity and temperate weather and negative coefficients on tropical latitude and distance to major trading centers are all highly significant and unsurprising in sign. Together with separate intercept terms for each of the decennial years, 1950 to 1990, the exogenous right-hand-side variables account for 59 percent of the variation in per capita income. The coefficients on the time dummies imply annual rates of technological progress (or more precisely, the rate of change for steady-state income) ranging from a low 0.4 percent per year for the 1980s to a high of 3.0 percent per year for the 1960s.

Table 5 Column 2a reports the partial correlations from the OLS regression of per capita income growth on initial income and the same set of exogenous right-hand-side variables. The coefficient on initial income implies a constant speed of the convergence of 0.4 percent per year, which is much lower than the conventional wisdom value of 2 percent per year. One possible explanation is that the specification does not sufficiently control for variations in steady state income (i.e. the distinction between absolute versus conditional convergence). But such an explanation seems doubtful given the high explanatory power of the same exogenous controls for the level variation in per capita income.¹⁵ Dividing the coefficients on the exogenous attributes through by 0.004 implies structural parameters somewhat larger in absolute value than the estimates from the level regression. But the level estimates remain within 2.1 standard errors of the implicit parameters.

A negative, significant coefficient on temperate weather stands out as the big surprise of the exogenous OLS conditional convergence growth regression. This negative partial correlation between per capita income growth and temperate weather proves extremely robust so long as the tropical latitude variable is also included as a right-hand-side control (The correlation between percent of the population living between 20 degrees south and 20 degrees north latitude and percent of the population living in temperate weather is -0.57.) Intro-

¹⁵Running the OLS level regression with the 379 observations from the growth regression yields R^2 values of 0.574 and 0.606 using start of decade and end of decade per capita income, respectively.

ducing additional right-hand-side variables such as secondary education and life expectancy strengthens the negative partial correlation and causes the implied negative structural effect on steady state income by temperate weather to differ statistically from zero as well. Of course, such a structural interpretation is nonsense. A priori, if any attributes underlie the wealth of nations, surely temperate weather is among them. Higher agricultural productivity and an environment less hospitable to vector borne disease transmission are just two of the many possible reasons why this is so (Bloom and Sachs, 1998). The positive partial correlation between current income and temperate weather in the level regression reaffirms such a view.

A more plausible interpretation of the negative partial correlation is that while temperate weather is positively correlated with high steady-state income, this positive correlation is smaller today than it was at some time in the past. Consistent with such an interpretation, Table 5 Column 3 shows that the coefficients on the exogenous controls from an OLS growth regression not conditioned on initial income are nearly identical to the coefficients from the conditional convergence regression. Nor is explanatory power of reduced much by excluding initial income; R^2 drops from 0.319 to 0.309. Using the terminology of Easterly, Kremer, Pritchett, and Summers (1993), residents of areas lacking temperate weather have the good luck to be living there today rather than in the past.

Table 5 Column 4 reports the nonlinear least square estimates of the structural coefficients in (5) using the exogenous specification set of controls. (The “wide constrained” column heading is meant to distinguish this regression for the one reported in Column 5; as discussed above, all of the NLS estimates of (5) impose constraints on feasible coefficients.) The estimated speed of convergence increases from $\hat{\lambda}_p = 0.003$ to $\hat{\lambda}_r = 0.269$ with the transition constrained to its upper boundary, $\kappa = 5$ (i.e., the speed remains low until relatively high incomes where it increases rapidly as income approaches its steady state). The estimated change in convergence speed, $(\lambda_r - \lambda_p)$, is 26.6 percent which is nearly three times its asymptotic standard error. However, the presence of nuisance parameters not identified under the null hypothesis, $\lambda_p = \lambda_r$, implies that we can not reject a constant speed of convergence. Instead, Monte Carlo simulations suggest that with a constant convergence speed data generating process equivalent to that estimated in Column 2a, the NLS estimation of

(5) using the exogenous controls will estimate $(\lambda_r - \lambda_p) \geq 0.266$ 14.2 percent of the time. The estimated annual rate of technological progress ranges from 2.1 percent to 2.6 percent for the 1950s through the 1970s falling to 0.2 percent for the 1980s.

Generalizing the cross-country growth specification to allow for a varying speed of convergence does not seem to improve the ability to recover underlying steady-state structure. The coefficients on the exogenous controls reported in Column 4 are similar to the implied structural parameters reported in Column 2b from the conditional convergence growth regression. In particular, the negative coefficient on temperate weather remains. Examining fitted values for steady-state annual per capita income in 1990 emphasizes the failure of (5) to recover structure. Japan and Korea’s steady-state annual per capita incomes are estimated at \$9.7 million and \$5.5 million, respectively. Seven additional nations are estimated to have steady-state annual per capita income levels above \$1 million (the Netherlands, Taiwan, Denmark, Norway, the United Kingdom, Belgium, and Ireland). Such high values seem fantastical even under the broadest interpretations of capital. At the other extreme, eleven African nations are estimated to have annual steady-state incomes in 1990 from \$16 to \$33 per capita (Malawi, Zambia, Kenya, Burundi, Rwanda, Uganda, the Central African Republic, Burkino Faso, Chad, Mali, and Sudan). Such low values imply that these country’s already poverty-level 1990 incomes are from 150 to more than 500 times their steady-state levels. But that sub-Saharan Africa is not capable of supporting human life is clearly disproved by ten thousand years of historical and anthropological evidence.

Figure 3 Panel A shows that, as with the standard specification, the exogenous-specification likelihood surface is smooth with respect to the convergence parameters, λ_p and λ_r . For the parameter range shown, $\{0 \leq \lambda_p \leq 0.016\} \times \{0 \leq \lambda_r \leq 0.320\}$, by far the lowest likelihood is associated with zero convergence (or equivalently, attributing all growth to technological progress). The associated R^2 value of 0.136 (front corner of Figure 3 Panel A) compares with a maximum constant convergence speed R^2 value of 0.319 (“ * ” in Figure 3 Panel B) and the global maximum increasing convergence speed R^2 value of 0.354 (“ • ” in Figure 3 Panel B).

The R^2 contours in Figure 3 Panel B delimit large swaths of λ_p - λ_r parameter space with likelihood above the OLS constant-convergence-speed level. By far the largest portion of this

parameter space with likelihood above the OLS level is characterized by an increasing speed of convergence (it lies to the southeast of the ray $\lambda_p = \lambda_r$). However there also exists a much smaller region with likelihood above the OLS level but characterized by a decreasing speed of convergence (to the northwest of the ray $\lambda_p = \lambda_r$). In fact, while the constant-convergence-speed likelihood represents a local maximum along the ray $\lambda_p = \lambda_r$, it represents a local minimum along the orthogonal vector, $\lambda_p = -\lambda_r$. This local minimum property is intuitive since relaxing the constraint of a constant convergence speed introduces to (5) three extra parameters with which to fit the observed growth rates.

Figure 3 Panel C shows the maximum R^2 attainable for any given difference in convergence speed, $d\lambda = \lambda_r - \lambda_p$. Explanatory power rises monotonically as this difference increases from zero attaining its highest value at $d\lambda = 0.266$ before declining slightly. Explanatory power also briefly rises as $d\lambda$ decreases from zero before dropping sharply. The constant-convergence-speed local minimum again derives from the introduction of the three extra parameters and so in itself is not interesting. But the figure highlights that the rise in explanatory power with a decreasing speed of convergence is small both in magnitude and with respect to the parameter space it covers relative to the the rise in explanatory power that occurs with an increasing speed of convergence.

The negative coefficient on temperate weather in both the OLS and NLS growth specifications exemplifies the criticism that in the absence of a reasonably accurate proxy for unobserved steady-state income, testing for conditional convergence (whether constant or varying) imposes minimal behavior restrictions on observed growth rates. To address such a concern, (5) is reestimated imposing the constraint that the structural coefficients lie within two standard deviations of their “level” estimates in Table 5 Column 1. As reported in Table 5 Column 6, this “narrow constrained” exogenous specification again suggests an increasing speed of convergence with $\hat{\lambda}_p = 0.003$ and $\hat{\lambda}_r = 0.045$. Interestingly, at the NLS maximum likelihood estimate, the only constraints which bind are the lower bounds on the negative coefficients on tropical latitude and distance to major markets; in particular, the constraint on the estimated positive coefficient on temperate weather remains slack.

The narrow constrained exogenous likelihood surface shown in Figure 4 looks quite similar to its wide constrained counterpart in Figure 3. By construction, the narrow constrained

surface lies at a somewhat lower R^2 level; the main other difference is that explanatory power attains its maximum at a much lower estimated value of λ_r . (Note that the horizontal scale used in the Figure 4 differs substantially from that used in Figure 3.) But the exogenous narrow constrained surface remains smooth with a large swath of parameter space characterized by an increasing speed of convergence lying above the constant convergence speed level and with the constant-convergence-speed likelihood representing a local minimum.

4.4.3 Unconditional Specification

The unconditional specification estimates reinforce the conclusion that the speed of convergence is increasing. The level regression reported in Table 6 Column 1a establishes that time dummies alone account for 60.6 percent of the variation in the level of per capita income across these OECD nations (almost identical to level explanatory power in the exogenous specification). The growth regression reported in Table 6 Column 2a establishes that time dummies together with initial income account for 62.9 percent of the variation in the growth rate of per capita income. Growth for these OECD nations is strongly characterized by unconditional convergence; the -0.025 coefficient on initial income is significant at the 0.001 level. Not including initial income and the growth regression results in a dramatic decrease in explanatory power; R^2 drops from 0.629 to 0.219 (Column 3). The NLS estimate of (5) gives $\hat{\lambda}_p = 0.006$ and $\hat{\lambda}_r = 0.043$. The relatively small sample size leads to a large asymptotic standard error relative to the estimated difference $\widehat{\lambda}_r - \widehat{\lambda}_p$. The Monte Carlo simulation suggests a 0.275 p-value for rejecting a constant speed of convergence.

The unconditional likelihood surface shown in Figure 5 again remains smooth with a large swath of parameter space characterized by an increasing speed of convergence lying above the constant convergence speed level. As with the exogenous specifications, the constant-convergence-speed likelihood represents a local minimum with respect to the difference $d\lambda = \lambda_r - \lambda_p$.¹⁶

Compared with the exogenous surfaces, the unconditional surface does differ in that

¹⁶The decreasing convergence speed local maximum occurs at $d\lambda = -0.024$ (not shown) and represents a 0.004 rise in R^2 from its constant convergence speed level (versus a 0.021 rise in R^2 at the increasing convergence speed global maximum).

the area of the parameter space where the unconditional likelihood lies above its constant convergence speed level extends to somewhat higher values of λ_p . For instance, the area with R^2 at least 0.01 above the constant convergence speed level extends up to $\lambda_p = 0.026$ with the unconditional specification versus up to $\lambda_p = 0.006$ and $\lambda_p = 0.011$ with the wide and narrow constrained exogenous specifications, respectively.

Using the OECD nations to estimate the speed of convergence for countries which are poor relative to their steady state is probably misleading. Few of these OECD observations are estimated to have initial incomes very low relative to their fitted steady state (which is the same for all 22 countries for a given year). For instance, only 4 of the 88 observations have initial incomes less than 40 percent of the estimated decade-specific common steady state. And only 6 more have initial incomes less than 60 percent of their estimated steady state. Moreover all 10 of these observations experienced rapid growth over the subsequent decade.¹⁷ In the heuristic argument above, it was exactly the relatively fast growth of countries with initial incomes between 50 percent and 60 percent of a proxy steady state that was meant to motivate the need for a λ_r above a λ_p which would apply to somewhat poorer countries. In the present case, the countries which are estimated to be poor relative to their steady state are simply absent.

A second problem with the unconditional specification is that the estimated decade-specific common steady states seem too low. Steady-state real per capita annual income (in 1990 dollars) for the 22 OECD nations is estimated to be \$4,024 in 1950; \$5,653 in 1960; \$8,433 in 1970; \$10,925 in 1980; and \$13,537 in 1990. Comparing actual incomes to such steady-state values illustrates the implausibility of the estimates. Focusing first on 1950, it seems reasonable that nations which had not had their capital stocks destroyed during World War II might have current income above the common steady-state level. What does not seem reasonable is the extent to which they exceed the estimated steady state. Most extreme is the United States with current income 218% the estimated steady state. And

¹⁷The country, initial year, initial income as a percentage of steady state, and average annual growth rate over the subsequent decade for the ten observations: Portugal, 1950, 30%, 4.4%; Portugal, 1960, 33%, 5.7%; Japan, 1950, 36%, 7.3%; Portugal, 1970, 39%, 4.1%; Portugal, 1980, 46%, 4.1%; Spain, 1950, 48%, 4.9%; Japan, 1960, 52%, 9.1%; Spain, 1960, 55%, 6.3%; Ireland, 1960, 59%, 4.2%; Ireland, 1970, 59%, 3.1%.

the incomes of Australia, New Zealand, and Switzerland range from 166% to 169% the estimated steady state. For nations which did have their capital stocks destroyed during the war, the estimated 1950 steady-state income level implies remarkably quick recoveries. France, Norway, Belgium, the Netherlands, Denmark, and the United Kingdom all had 1950 per capita income above the estimated steady state. Even West Germany's per capita income had already reattained 85% the estimated steady-state level. As (6) links steady state income across decades by the rate of technological progress, a low estimate of steady-state income in 1950 implies low subsequent estimates of steady-state income. In each of the decades, either 12 or 13 of the 22 nations have initial incomes above the estimated common steady-state level. And jumping forward to 1990, the United States' per capita income remains 133% above the estimated steady state.

4.4.4 Posterior Odds: Increasing Versus Constant Convergence Speed

Each of three specifications effects point estimates that the speed at which per capita output converges toward its steady-state level increases as economies develop. While the null hypothesis of a constant convergence speed cannot be rejected, no evidence is found to support such a null. In particular, the likelihood surfaces for each of the specifications is smooth with respect to the coefficients parameterizing the speed of convergence. Large swathes of the parameter space characterized by an increasing convergence have a likelihood above the constant convergence speed level. And the finding of an increasing speed of convergence is robust to imposing behavioral constraints bounding the structural coefficients.

Accepting that there is nothing natural or focal about a constant speed of convergence, posterior odds nicely summarize the empirical evidence favoring an increasing speed of convergence. Such posterior odds are simply the ratio of the probability given the data that the change in convergence speed, $\lambda_r - \lambda_p$, is greater than some arbitrary cutoff relative to the probability given the data that the change is less than or equal to the same cutoff. Assuming the error term in (5) is normally distributed, the posterior odds can be calculated by *integrating* (rather than maximizing) the likelihood function over $\lambda_p \geq 0$ times $\lambda_r \geq 0$

(Kass and Raftery, 1995).

$$\text{Posterior Odds}(\lambda_r - \lambda_p > x) = \frac{\iint_{\lambda_r - \lambda_p > x} \text{Likelihood}(\text{observed data} | \lambda_p, \lambda_r) d\lambda_p d\lambda_r}{\iint_{\lambda_r - \lambda_p \leq x} \text{Likelihood}(\text{observed data} | \lambda_p, \lambda_r) d\lambda_p d\lambda_r} \quad (8)$$

Intuitively, the posterior odds just is just the ratio of the area under the surfaces in Panel A of Figures 2 through 5 on either side of the plane defined by $\lambda_r - \lambda_p = x$. In practice, the integrals are approximated by Riemann sums on a discrete partition of $\lambda_p \in [0, 0.03]$ times $\lambda_r \in [0, 0.32]$. The likelihood function in the integral is conditioned on λ_p and λ_r but is maximized with respect to the remaining parameters (i.e. κ , et cetera).¹⁸

Figure 6 Panel A shows the posterior odds for the standard and unconditional specifications. Figure 6 Panel B does the same for the two exogenous specifications. The various loci give the estimated odds that the change in convergence speed is greater than the corresponding cutoff along the horizontal axis.

Focusing first on Panel A, the vertical intercepts for both the standard and unconditional specifications place the odds that the speed of convergence is at least slightly increasing at greater than 40 to 1. Moving to the right along the horizontal axis, the odds steadily fall that the change in convergence speed exceeds the given cutoff. The standard and unconditional specification odds that convergence speed increases by more than two percent are 6.4 and 8.4, respectively. The corresponding odds that convergence speed increases by more than four percent are 0.78 and 1.06. A unitary odds implies equal probability that the change in convergence speed is above or below the associated cutoff. So the standard specification estimates that the probability that the change in convergence speed is less than four percent exceeds the probability that it is above four percent. The cutoffs associated with unitary odds are 3.7% ($\lambda_r - \lambda_p = 0.037$) for the standard specification and 4.0% ($\lambda_r - \lambda_p = 0.040$) for the unconditional specification.

The equal probability above and below a cutoff captured by a unitary posterior odds represents a very weak criterion by which to judge the behavior of the speed of convergence. Kass and Raftery (1995) suggest instead that 20 be used as the minimum odds which suggest

¹⁸In addition, for computational reasons the likelihood function is constrained to a minimum value equal to the likelihood given zero convergence (i.e. the likelihood that the data is generated by decade-specific technological progress and a residual term only).

strong evidence in favor of one hypothesis over another.¹⁹ The standard specification yields odds of at least 20 for an increase in convergence speed up to 0.9%. The unconditional specification yields odds of at least 20 for an increase in convergence speed up to 1.1%. Conversely, an odds less than 1/20 can be considered strong evidence that the increase in convergence speed is less than the associated cutoff. The standard specification yields odds below 1/20 for an increase in convergence speed of at least 10.9%. The unconditional specification yields odds below 1/20 for an increase in convergence speed of at least 7.7%. So together the standard and unconditional specifications strongly suggest that the speed of convergence is increasing with the change in convergence speed, $\lambda_r - \lambda_p$, probably falling somewhere between 0.9% and 10.9%.²⁰

Focusing next on the wide and narrow constrained exogenous specifications summarized in Figure 6 Panel B, the vertical intercepts place the respective odds of at least a slightly increasing speed of convergence at 63,000 to 1 and 20,000 to 1. The narrow constrained specification yields odds of at least 20 for an increase in convergence speed up to 2.5%; it yields even odds for an increase in convergence speed of 4.7%; and it yields odds below 1/20 for an increase in convergence speed above 8.5%. So the narrow constrained exogenous specification suggests that the speed of convergence is increasing with the change in speed probably falling somewhere between 2.5% and 8.5%.

The wide constrained specification suggests that the speed of convergence is increasing by an even greater amount. It yields odds of at least 20 for an increase in convergence speed

¹⁹The Kass and Raftery criterion actually applies to Bayes Factors which are the ratio of posterior odds to prior odds. The justification for focusing on posterior odds in the present case is the lack of any compelling priors on the behavior of convergence speed. Applying the Kass and Raftery criterion to posterior odds implicitly assumes unitary prior odds that convergence speed is above or below *any* candidate cutoff. Alternatively, one can assume a uniform prior distribution of $\lambda_p \in [0, 0.03]$ times $\lambda_r \in [0, 0.032]$ which is the space over which the likelihood function for the data generating process is calculated. But such a prior already strongly favors an increasing speed of convergence; normalizing by it obscures the evidence contained in the data.

²⁰The standard specification yields 9.5 to 1 odds that the change in convergence speed falls in [0.009, 0.109]. The unconditional specification yields 9.5 to 1 odds that the change in convergence speed falls in [0.011, 0.077]. These essentially correspond to 90 percent confidence intervals. Note that the standard specification interval spans the unconditional specification interval.

up to 11.4% and yields even odds at 23%.²¹ However, these high odds for a large increase in convergence speed are calculated based on convergence to estimated steady states which, as discussed above, are clearly implausible. Hence the results for the wide constrained exogenous specification are suspect.

5 Conclusions

The present paper has argued that the speed of convergence varies. Empirical evidence suggests that convergence speed increases as countries develop: from below 1 percent per year for economies with low relative outputs to above 4 percent per year (and possibly much higher) for economies with output approaching its steady state. The null hypothesis of a constant convergence speed cannot be rejected; but there is no theoretical basis for making a constant convergence speed focal in the first place.

With a varying speed of convergence, a level structural interpretation of coefficients on conditioning attributes in cross-sectional growth regressions breaks down; the partial correlations between growth and the conditioning attributes may differ in sign from the partial correlations between steady-state income and the conditioning attributes. Within a neoclassical framework, understanding the determinants of steady-state income clearly stands out as the more important relationship. Even though growth regressions cannot identify these underlying determinants, nevertheless they can contribute to the effort to do so. In particular, regressing per capita income growth on a set of exogenous attributes while excluding initial income can show how structural relationships have changed across time.

Equally important, an increasing speed of convergence presents a new benchmark against which to calibrate neoclassical growth models. Given the tendency of the Ramsey-Cass-Koopmans model to effect a sharply decreasing speed of convergence, the results herein would seem to make futile what is already a difficult challenge. But on the contrary, it

²¹Calculating the increase in convergence speed at which the wide constrained exogenous specification yields odds below 1/20 would require expanding the λ_r search space above 0.32. As illustrated in Figure 3 Panel B, there should be a large swathe of area with $\lambda_r > 0.32$ and likelihood above the OLS constant convergence speed level. Expanding the search space would also increase the cutoffs associated with the 20 to 1 and even odds.

turns out that adopting an adjustment cost structure which effects an increasing speed of convergence allows the Ramsey-Cass-Koopmans model to match the development paths of other real world observables including real interest rates, savings rates, and shadow values of capital (Rappaport, 2000b). That allowing for an increasing speed of convergence proves the key to successfully calibrating this core macroeconomic model might be considered further evidence that the speed of convergence is indeed so.

Appendices

A Convergence with More than One State Variable: Some Algebra

Two key characteristics of the dynamic model are, first, that its steady-state is not uniquely determined but rather depends on history; and second, that the speed of convergence – the rate at which income and population approach their steady-state levels normalized by their distances from their respective steady-state levels – varies greatly, even in a neighborhood quite close to the steady-state. It turns out that both properties, history dependence and a varying speed of convergence, are generic with multiple state (i.e. “non-jumping”) variables. The “proof” lies mainly in pointing out the necessity of an N-dimensional surface to span an N-dimensional space. That is, to assure that *some* steady state can be reached from any feasible starting-value combination of state variables, the dimensionality of possibly multiple steady states plus the dimensionality of the transition surface to each of these must sum to the number of state variables.

A steady state with dimensionality one or more is equivalent to history dependence. Such history dependence is more common than is often believed. Barro (1979) shows there is no one optimal level of government debt; rather, a country’s optimal debt depends on its specific history of shocks (i.e. wars, famines, baby booms, natural resource finds, etc.). In two sector endogenous growth models (Mulligan and Sala-i-Martin, 1993; Caballe and Santos, 1993), the ratio but not the level of human to physical capital is determinate (the level however is less interesting within an endogenous growth framework).

The *linearization* of a transition path with two or more dimensions will always show an increasing speed of convergence: near the steady-state the negative eigenvalue lowest in absolute value will dominate; as one moves away from the steady-state, the negative eigenvalue highest in absolute

value will dominate. The algebra showing this follows immediately below. That the asymptotic speed of convergence is given by the negative eigenvalue lowest in absolute value is also pointed out by Eichner and Turnovsky (1999). The actual transition path (rather than its linearized approximation) may show an increasing, constant, or decreasing speed of convergence. A constant speed of convergence, however, would be a razor thin result.

Some definitions:

$$\begin{aligned}
 \mathbf{b} &\equiv \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{bmatrix} & \mathbf{z} &\equiv \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ \vdots \\ z_n \end{bmatrix} & \mathbf{l} &\equiv \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} \left. \vphantom{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix}} \right\} \text{n times} \\
 \\
 \mathbf{b} \odot \mathbf{z} &\equiv \begin{bmatrix} b_1 z_1 \\ b_2 z_2 \\ \vdots \\ \vdots \\ b_n z_n \end{bmatrix} & \mathbf{b} \oslash \mathbf{z} &\equiv \begin{bmatrix} b_1/z_1 \\ b_2/z_2 \\ \vdots \\ \vdots \\ b_n/z_n \end{bmatrix} & \log \mathbf{z} &\equiv \begin{bmatrix} \log z_1 \\ \log z_2 \\ \vdots \\ \vdots \\ \log z_n \end{bmatrix} \\
 \\
 \mathbf{A}(\mathbf{z}) &\equiv \begin{bmatrix} A_{11}(\mathbf{z}) & A_{12}(\mathbf{z}) & \cdots & \cdots & A_{1n}(\mathbf{z}) \\ A_{21}(\mathbf{z}) & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ A_{n1}(\mathbf{z}) & \cdots & \cdots & \cdots & A_{nn}(\mathbf{z}) \end{bmatrix} \\
 \\
 \frac{\partial \mathbf{A}(\mathbf{z})}{\partial \mathbf{z}} &\equiv \begin{bmatrix} \frac{\partial A_{11}(\mathbf{z})}{\partial z_1} & \frac{\partial A_{12}(\mathbf{z})}{\partial z_2} & \cdots & \cdots & \frac{\partial A_{1n}(\mathbf{z})}{\partial z_n} \\ \frac{\partial A_{21}(\mathbf{z})}{\partial z_1} & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \frac{\partial A_{n1}(\mathbf{z})}{\partial z_1} & \cdots & \cdots & \cdots & \frac{\partial A_{nn}(\mathbf{z})}{\partial z_n} \end{bmatrix}
 \end{aligned}$$

Suppose a system of differential equations defined in terms of the logarithms of the vector of variables, \mathbf{z} :

$$\frac{d}{dt} \log \mathbf{z} = \mathbf{A}(\log \mathbf{z}) \tag{A.1}$$

Take a Taylor expansion around the system's steady-state:

$$\frac{d}{dt} \log \mathbf{z} \approx \mathbf{A}(\log \mathbf{z}^*) + \left. \frac{\partial \mathbf{A}(\log \mathbf{z})}{\partial \log \mathbf{z}} \right|_{\log \mathbf{z} = \log \mathbf{z}^*} \cdot (\log \mathbf{z} - \log \mathbf{z}^*) \quad (\text{A.2a})$$

$$= \mathbf{J} \log(\mathbf{z} \odot \mathbf{z}^*) \quad (\text{A.2b})$$

$$\mathbf{J} \equiv \left. \frac{\partial \mathbf{A}(\log \mathbf{z})}{\partial \log \mathbf{z}} \right|_{\log \mathbf{z} = \log \mathbf{z}^*}$$

Let $|\lambda_f| > |\lambda_s|$ represent two *negative* eigenvalues of \mathbf{J} with corresponding eigenvectors \mathbf{v}_f and \mathbf{v}_s , and arbitrary weighting scalars, a_f and a_s (“*f*” is meant to connote “fast” and “*s*”, “slow”).

$$\mathbf{J} (a_f \mathbf{v}_f + a_s \mathbf{v}_s) = -(\lambda_f a_f \mathbf{v}_f + \lambda_s a_s \mathbf{v}_s) \quad (\text{A.3})$$

Then the solution to (A.2b) can be written as,

$$\log(\mathbf{z} \odot \mathbf{z}^*) \approx a_f \mathbf{v}_f e^{-\lambda_f t} + a_s \mathbf{v}_s e^{-\lambda_s t} \quad (\text{A.4})$$

Take the derivative of (A.4) with respect to t ,

$$\frac{d}{dt} \log(\mathbf{z} \odot \mathbf{z}^*) \approx -(\lambda_f a_f \mathbf{v}_f e^{-\lambda_f t} + \lambda_s a_s \mathbf{v}_s e^{-\lambda_s t}) \quad (\text{A.5})$$

The canonical underlying structural form for growth regressions is,

$$\frac{d}{dt} \log z = \frac{d}{dt} (\log z - \log z^*) = -\lambda \cdot (\log z - \log z^*)$$

For the vector analog to this we want,

$$\boldsymbol{\lambda} \text{ s.t. } \frac{d}{dt} \log(\mathbf{z} \odot \mathbf{z}^*) = -\boldsymbol{\lambda} \odot \log(\mathbf{z} \odot \mathbf{z}^*)$$

Substituting using (A.4) and (A.5) gives,

$$\boldsymbol{\lambda} \text{ s.t. } (\lambda_f a_f \mathbf{v}_f e^{-\lambda_f t} + \lambda_s a_s \mathbf{v}_s e^{-\lambda_s t}) \approx \boldsymbol{\lambda} \odot (a_f \mathbf{v}_f e^{-\lambda_f t} + a_s \mathbf{v}_s e^{-\lambda_s t}) \quad (\text{A.6})$$

It is immediately evident that unless either a_f or a_s equal zero, $\boldsymbol{\lambda}$ will differ in its elements *and* vary with time. I now formally define $\boldsymbol{\lambda}(t)$ as,

$$\boldsymbol{\lambda}(t) \equiv -\frac{d}{dt} \log(\mathbf{z} \odot \mathbf{z}^*) \odot \log(\mathbf{z} \odot \mathbf{z}^*)$$

The log linearization therefore approximates the speed of convergence as,

$$\boldsymbol{\lambda}(t) \approx (\lambda_f a_f \mathbf{v}_f e^{-\lambda_f t} + \lambda_s a_s \mathbf{v}_s e^{-\lambda_s t}) \odot (a_f \mathbf{v}_f e^{-\lambda_f t} + a_s \mathbf{v}_s e^{-\lambda_s t}) \quad (\text{A.7})$$

Normalize the eigenvectors so that the first element of each equals one. Then the speed of convergence corresponding to this first element is given by,

$$\lambda_1(t) \approx \frac{\lambda_f a_f e^{-\lambda_f t} + \lambda_s a_s e^{-\lambda_s t}}{a_f e^{-\lambda_f t} + a_s e^{-\lambda_s t}} \quad (\text{A.8})$$

So except in the special case when a_f or a_s equal zero, the linearization implies the speed of convergence for this representative first element will go from a_f to a_s as time goes from negative to positive infinity. If a_f and a_s are oppositely signed, it will also asymptote to positive and negative infinity at some intermediate time.

Bibliography

Ades, Alberto F. and Edward L. Glaeser (1994). "Evidence on Growth, Increasing Returns, and the Extent of the Market." NBER Working Paper No. 4714 (April).

Andrews, Donald W. K. (1993). "Test for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 61, pp. 821–856.

Barro, Robert J. (1991). "Economic Growth in a Cross-Section of Countries." *Quarterly Journal of Economics* 106, 2 (May), pp. 407–43.

Barro, Robert J. and Jong-Wha Lee (1994). "Data Panel for a Panel of 138 Countries." Available from <http://www.nber.org>.

Barro, Robert J. and Xavier Sala-i-Martin, (1991). "Convergence across States and Regions." *Brookings Papers on Economic Activity*, 1:1991, pp. 107–182.

Barro, Robert J. and Xavier Sala-i-Martin (1992). "Convergence." *Journal of Political Economy* 100, 2 (April), pp. 223–251.

Barro, Robert J. and Xavier Sala-i-Martin (1995). *Economic Growth*. New York: McGraw Hill.

Bloom, David E. and Jeffrey D. Sachs (1998). "Geography, Demography, and Economic Growth in Africa." *Brookings Panel on Economic Activity* 2:1998.

Caselli, Francesco, Gerardo Esquivel, and Fernando Lefort (1996). "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics." *Journal of Economic Growth* 1, 3 (September), pp. 363–389.

- Cass, David (1965). "Optimum Growth in an Aggregative Model of Capital Accumulation." *Review of Economic Studies* 32 (July), pp. 233–240.
- Davies, Robert B. (1977). "Hypothesis Testing when a Nuisance Parameter is Present Only under the Alternative." *Biometrika* 64, 2 (August), pp. 247–254.
- Davies, Robert B. (1987). "Hypothesis Testing when a Nuisance Parameter is Present Only under the Alternatives." *Biometrika* 74, 1 (March), pp. 33–43.
- Durlauf, Steven N. and Danny T. Quah (1999). "The New Empirics of Economic Growth." In John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics* Vol 1A, Amsterdam: Elsevier.
- Easterly, William, Michael Kremer, Lant Pritchett, and Lawrence Summers (1993). "Good Policy or Good Luck." *Journal of Monetary Economics* 32, 3 (Dec.), pp 459–483.
- Eicher, Theo S. and Stephen J. Turnovsky (1999). "Convergence in a Two-Sector Nonscale Growth Model." *Journal of Economic Growth* 4, 4 (Dec.), pp. 413–428.
- Evans, Paul (1997). "How Fast Do Economies Converge?" *Review of Economics and Statistics* 79, 2 (May), pp. 219–225.
- Hansen, Bruce E. (1996). "Inference When a Nuisance Parameter Is Not Identified under the Null Hypothesis." *Econometrica* 64, 2 (March), pp. 413–430.
- Islam, Nazrul (1995). "Growth Empirics: a Panel Data Approach." *Quarterly Journal of Economics* 110, 4 (November), pp. 1127–1170.
- Kass, Robert E. and Adrian E. Raftery (1995), "Bayes Factors." *Journal of the American Statistical Association* 90, 430 (June), pp. 773–794
- King, Robert G. and Sergio T. Rebelo (1993). "Transitional Dynamics and Economic Growth in the Neoclassical Model." *American Economic Review* 83, 4 (September), pp. 908–931.
- Koopmans, Tjalling C. (1965). "On the Concept of Optimal Economic Growth." *The Econometric Approach to Development Planning*. Amsterdam: North Holland.
- Kremer, Michael and James Thomson (1998). "Why Isn't Convergence Instantaneous? Young Workers, Old Workers, and Gradual Adjustment." *Journal of Economic Growth* 3, 1 (March), pp. 5–28.
- Lee, Kevin, M. Hashem Pesaran, and Ron Smith (1998). "Growth Empirics: A Panel Data Approach - A Comment." *Quarterly Journal of Economics* 113, 1 (February), pp. 319–324.

Ramsey, Frank (1928). “A Mathematical Theory of Saving.” *Economic Journal* 38 (Dec.), pp. 543–559.

Rappaport, Jordan (2000a). “How Does Labor Mobility Affect Income Convergence?” Federal Reserve Bank of Kansas City, Research Working Paper 99–12 (May).

Rappaport, Jordan (2000b). “How Does Openness to Capital Flows Affect Growth?” Federal Reserve Bank of Kansas City, Research Working Paper 00–11 (December).

Sala-i-Martin, Xavier (1994). “Cross-sectional Regressions and the Empirics of Economic Growth.” *European Economic Review* 38, pp. 739–747.

Solow, Robert M. (1956). “A Contribution to the Theory of Economic Growth.” *Quarterly Journal of Economics* 70, 1, pp. 65–94.

Summers, Robert and Alan Heston (1995). “The Penn World Table (Mark 5.6): An Expanded Set of International Comparisons, 1950–1992.” Available from <http://www.nber.org>.

Table 3: Summary Statistics

	Obs	Mean	Std Dev	Min	Max
Standard Specification					
Life Expectancy at Birth	271	58.0	12.0	32.3	76.0
Average Secondary Education	271	0.806	0.866	0.008	5.106
% Pop. between 20°S to 20°N latitude	95	0.491	0.487	0	1
Real Per Capita Income, 1960	90	2,505	2,342	313	9,895
Real Per Capita Income, 1970	95	3,434	3,240	419	12,963
Real Per Capita Income, 1980	86	4,778	4,133	471	15,295
Real Per Capita Income Growth Rate, 1960-1970	90	0.030	0.020	-0.024	0.091
Real Per Capita Income Growth Rate, 1970-1980	95	0.024	0.028	-0.048	0.087
Real Per Capita Income Growth Rate, 1980-1990	86	0.006	0.024	-0.037	0.077
Exogenous Specification					
% Pop. within 100km of ocean or sea coast	121	0.456	0.366	0	1
% Pop. within in temperate weather	121	0.258	0.358	0	1
% Pop. between 20°S to 20°N latitude	121	0.502	0.478	0	1
air dist. to nearest of NYC, Rotterdam, Tokyo (km) [†]	121	4,258	2,476	140	9,320
Real Per Capita Income, 1950	55	2,505	2,081	221	8,772
Real Per Capita Income, 1960	109	2,230	2,229	257	9,895
Real Per Capita Income, 1970	116	3,144	3,121	296	12,963
Real Per Capita Income, 1980	119	4,456	4,917	322	31,969
Real Per Capita Income, 1990	101	4,796	4,894	399	18,054
Real Per Capita Income Growth Rate, 1950-1960	55	0.025	0.017	-0.010	0.073
Real Per Capita Income Growth Rate, 1960-1970	109	0.028	0.023	-0.063	0.091
Real Per Capita Income Growth Rate, 1970-1980	116	0.023	0.028	-0.055	0.087
Real Per Capita Income Growth Rate, 1980-1990	99	0.004	0.023	-0.039	0.077
Unconditional Specification[‡]					
Real Per Capita Income, 1950	22	4,517	1,980	1,208	8,772
Real Per Capita Income, 1960	22	6,009	2,069	1,869	9,895
Real Per Capita Income, 1970	22	8,662	2,292	3,306	12,963
Real Per Capita Income, 1980	22	11,054	2,360	4,982	15,295
Real Per Capita Income, 1990	22	13,570	2,515	7,478	18,054
Real Per Capita Income Growth Rate, 1950-1960	22	0.033	0.017	0.012	0.073
Real Per Capita Income Growth Rate, 1960-1970	22	0.040	0.016	0.017	0.091
Real Per Capita Income Growth Rate, 1970-1980	22	0.026	0.011	0.010	0.054
Real Per Capita Income Growth Rate, 1980-1990	22	0.021	0.008	0.011	0.041

[†] Measured from capital city. [‡] 22 OECD Nations: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany (West), Iceland, Ireland, Italy, Japan, Luxembourg, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, United States.

Table 4: "Standard" Specification Parameter Estimates

RHS Variables and Parameters	(1a)	(1b)	(2a)	(2b)	(3)	(4)
	OLS Level Regression		OLS Growth Regressions			NLS
	raw coefficients	implied structural parameters	raw coefficients	implied structural parameters	excluding initial income	
cnvrg speed when "poor" (λ_p) OR log(initial gdp)			-0.019 (0.003)	0.019 (0.003)		0.009 (0.013)
cnvrg speed when "rich" (λ_r)						0.040 (0.016)
transition from λ_p to λ_r (κ)						1.00 ¹ (1.26)
life expectancy at birth	0.05 (0.01)		0.002 (0.000)	0.086 (0.010)	0.001 (0.000)	0.09 (0.01)
average secondary education	0.25 (0.05)		0.003 (0.002)	0.171 (0.089)	-0.002 (0.002)	0.06 (0.06)
pct pop in 20°S to 20°N latitude	-0.18 (0.13)		-0.013 (0.004)	-0.689 (0.228)	-0.010 (0.004)	-0.53 (0.18)
1960 dummy OR 1960-1970 level intercept	4.44 (0.33)					2.13 (0.54)
1970 dummy OR 1960-1970 growth intercept OR dy*/dt 1960-1970	4.49 (0.35)	0.005 (0.003)	0.088 (0.016)		0.003 (0.011)	0.040 ^u (0.004)
1980 dummy OR 1970-1980 growth intercept OR dy*/dt 1970-1980	4.44 (0.36)	-0.005 (0.003)	0.081 (0.016)		-0.005 (0.012)	0.026 (0.004)
1990 dummy OR 1980-1990 growth intercept OR dy*/dt 1980-1990			0.060 (0.017)		-0.025 (0.012)	0.000 ¹ (0.003)
N	271		271		271	271
R ²	0.813		0.382		0.287	0.394
Sum of Squared Residuals	47.2		0.113		0.130	0.110
Number of Parameters	6		7		6	10
change in cnvrg speed ($\lambda_r - \lambda_p$)						0.031 (0.012)
p-value ($\lambda_r - \lambda_p$) > 0 from monte carlo simulation						0.395
Likelihood Ratio NLS point estimate vs $\lambda_r = \lambda_p$						13.6

¹ parameter constrained to lower bound; ^u parameter constrained to upper bound. Columns 1a and 1b report estimates of the equation $y = x\beta + \varepsilon$ where y is the log of per capita income. Columns 2a, 2b, and 4 report OLS and NLS estimates of the equation $dy/dt = -[(\lambda_r - \lambda_p)\exp(-\kappa|y - y^*|) + \lambda_p] \cdot (y - y^*) + dy^*/dt + \varepsilon$ where y^* , which is estimated by $x\beta$, is the log of steady-state income. For OLS estimate, $\lambda_r = \lambda_p$ implicitly and κ is not identified. NLS estimates are by modified grid search. Column 3 reports estimate of the equation $dy/dt = x\delta + dy^*/dt + \varepsilon$. Standard errors in parentheses allow for country random effects; standard errors of implied structural parameters are constructed by the "delta" method; standard errors for NLS estimates hold asymptotically for $\lambda_p \neq \lambda_r$.

Table 5: "Exogenous" Specification Parameter Estimates

RHS Variables and Parameters	(1a) (1b) OLS Level Regression		(2a) (2b) (3) OLS Growth Regressions			(4) (5) NLS Growth Regressions	
	raw coefficients	implied structural parameters	raw coefficients	implied structural parameters	excluding initial income	wide constrained	narrow constrained
cnvrg speed when "poor" (λ_p) OR log(initial gdp)			-0.004 (0.002)	0.004 (0.002)		0.003 (0.000)	0.003 (0.005)
cnvrg speed when "rich" (λ_r)						0.269 (0.093)	0.045 (0.019)
transition from λ_p to λ_r (κ)						5.00 ^u (1.32)	1.46 (0.84)
pct pop within 100km of ocean or sea coast	0.76 (0.16)		0.016 (0.004)	3.79 (1.69)	0.013 (0.003)	5.82 (0.33)	0.96 (0.28)
pct pop in temperate weather	0.71 (0.26)		-0.012 (0.006)	-2.79 (2.15)	-0.015 (0.005)	-2.03 (0.21)	1.22 (0.37)
pct pop in 20°S to 20°N latitude	-0.58 (0.18)		-0.023 (0.004)	-5.48 (2.38)	-0.021 (0.004)	-5.50 (0.21)	-0.93 ¹ (0.28)
log(air distance to NYC, Rotterdam, Tokyo)	-0.21 (0.08)		-0.006 (0.001)	-1.37 (0.61)	-0.005 (0.002)	-1.02 (0.07)	-0.36 ¹ (0.10)
1950 dummy OR 1950-1960 level intercept	8.60 (0.67)					16.64 (0.62)	9.14 (0.86)
1960 dummy OR 1950-1960 growth intercept OR dy*/dt 1950-1960	8.73 (0.67)	0.014 (0.006)	0.105 (0.021)		0.071 (0.014)	0.021 (0.002)	0.030 (0.003)
1970 dummy OR 1960-1970 growth intercept OR dy*/dt 1960-1970	9.03 (0.67)	0.030 (0.002)	0.112 (0.020)		0.077 (0.014)	0.026 (0.002)	0.034 (0.002)
1980 dummy OR 1970-1980 growth intercept OR dy*/dt 1970-1980	9.29 (0.67)	0.026 (0.003)	0.107 (0.021)		0.071 (0.014)	0.021 (0.002)	0.028 (0.003)
1990 dummy OR 1980-1990 growth intercept OR dy*/dt 1980-1990	9.33 (0.66)	0.004 (0.004)	0.089 (0.021)		0.052 (0.014)	0.002 (0.002)	0.008 (0.003)
N	500		379		379	379	379
R ²	0.591		0.319		0.309	0.354	0.273
Sum of Squared Residuals	213.3		0.171		0.174	0.163	0.183
Number of Parameters	9		9		8	12	12
cnvrg speed change ($\lambda_r - \lambda_p$)						0.266 (0.093)	0.043 (0.017)
p-value ($\lambda_r - \lambda_p$) > 0 from monte carlo simulation						0.142	
Likelihood Ratio NLS point estimate vs $\lambda_r = \lambda_p$						1.96E+04	4.55E+03

¹ parameter constrained to lower bound; ^u parameter constrained to upper bound. Columns 1a and 1b report estimates of the equation $y = x\beta + \varepsilon$ where y is the log of per capita income. Columns 2a, 2b, 4, and 5 report OLS and NLS estimates of the equation $dy/dt = -[(\lambda_r - \lambda_p)\exp(-\kappa|y - y^*|) + \lambda_p](y - y^*) + dy^*/dt + \varepsilon$ where y^* , which is estimated by $x\beta$, is the log of steady-state income. For OLS estimate, $\lambda_p = \lambda_r$ implicitly and κ is not identified. NLS estimates are by modified grid search; in Column 5, coefficients on non-dummy variables are constrained to lie within two standard deviations of the level estimates in Column 1a. Column 3 reports estimate of the equation $dy/dt = x\delta + dy^*/dt + \varepsilon$. Standard errors in parentheses allow for country random effects; standard errors of implied structural parameters are constructed by the "delta" method; standard errors for NLS estimates hold asymptotically for $\lambda_p \neq \lambda_r$.

Table 6: "Unconditional" Specification Parameter Estimates

RHS Variables and Parameters	(1a) OLS Level Regression		(2a) OLS Growth Regressions			(4) NLS
	raw coefficients	implied structural parameters	raw coefficients	implied structural parameters	excluding initial income	
cnvrg speed when "poor" (λ_p) OR log(initial gdp)			-0.025 (0.005)	0.025 (0.005)		0.006 (0.085)
cnvrg speed when "rich" (λ_r)						0.043 (0.014)
transition from λ_p to λ_r (κ)						1.00 ¹ (4.04)
1950 dummy OR 1950-1960 level intercept	8.30 (0.11)					8.30 (0.25)
1960 dummy OR 1950-1960 growth intercept OR dy*/dt 1950-1960	8.63 (0.09)	0.033 (0.004)	0.242 (0.045)		0.033 (0.004)	0.034 (0.008)
1970 dummy OR 1960-1970 growth intercept OR dy*/dt 1960-1970	9.03 (0.07)	0.040 (0.003)	0.257 (0.046)		0.040 (0.003)	0.040 ^u (0.006)
1980 dummy OR 1970-1980 growth intercept OR dy*/dt 1970-1980	9.28 (0.05)	0.026 (0.002)	0.253 (0.047)		0.026 (0.002)	0.026 (0.005)
1990 dummy OR 1980-1990 growth intercept OR dy*/dt 1980-1990	9.50 (0.04)	0.021 (0.002)	0.256 (0.049)		0.021 (0.002)	0.021 (0.003)
N	110		88		88	88
R ²	0.606		0.629		0.219	0.650
Sum of Squared Residuals	13.39		0.007		0.015	0.007
Number of Parameters	5		5		4	8
change in cnvrg speed ($\lambda_r - \lambda_p$)						0.037 (0.077)
p-value ($\lambda_r - \lambda_p$) > 0 from monte carlo simulation						0.275
Likelihood Ratio NLS point estimate vs $\lambda_r = \lambda_p$						10.9

¹ parameter constrained to lower bound; ^u parameter constrained to upper bound. Observations are the OECD nations as of 1973 excluding Greece and Turkey. Columns 1a and 1b report estimates of the equation $y = x\beta + \varepsilon$ where y is the log of per capita income. Columns 2a, 2b, and 4 report OLS and NLS estimates of the equation $dy/dt = -[(\lambda_r - \lambda_p)\exp(-\kappa|y-y^*|) + \lambda_p](y-y^*) + dy^*/dt + \varepsilon$ where y^* , which is estimated by $x\beta$, is the log of steady-state income. For OLS estimate, $\lambda_r = \lambda_p$ implicitly and κ is not identified. NLS estimates are by modified grid search. Column 3 reports estimate of the equation $dy/dt = x\delta + dy^*/dt + \varepsilon$. Standard errors in parentheses allow for country random effects; standard errors of implied structural parameters are constructed by the "delta" method; standard errors for NLS estimates hold asymptotically for $\lambda_p \neq \lambda_r$.

Figure 1: Constant and Varying Convergence Speed

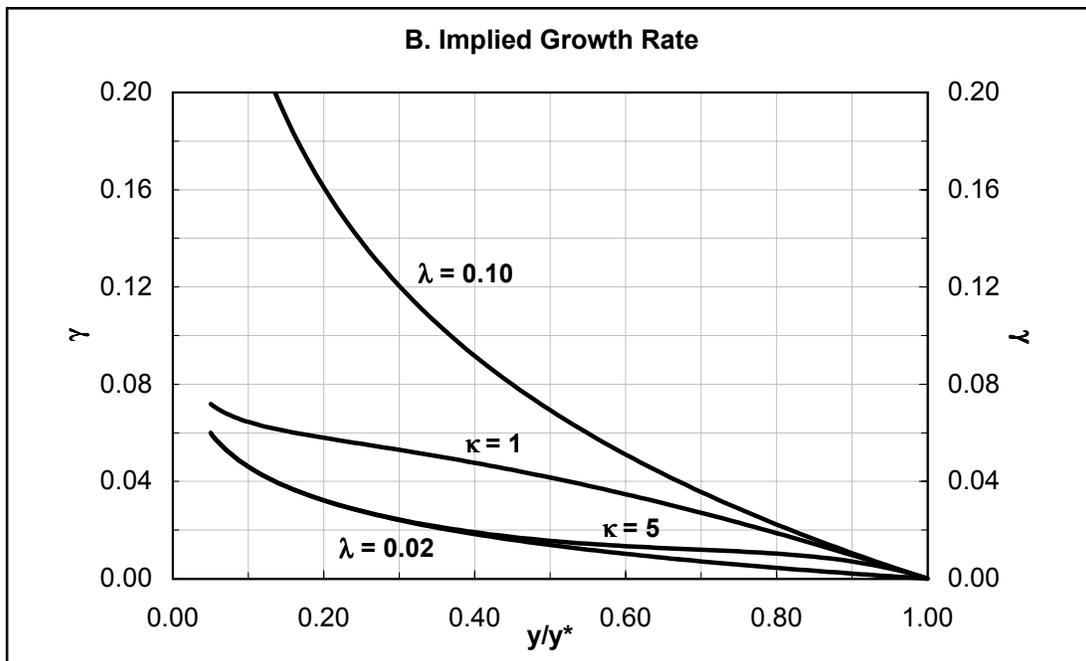
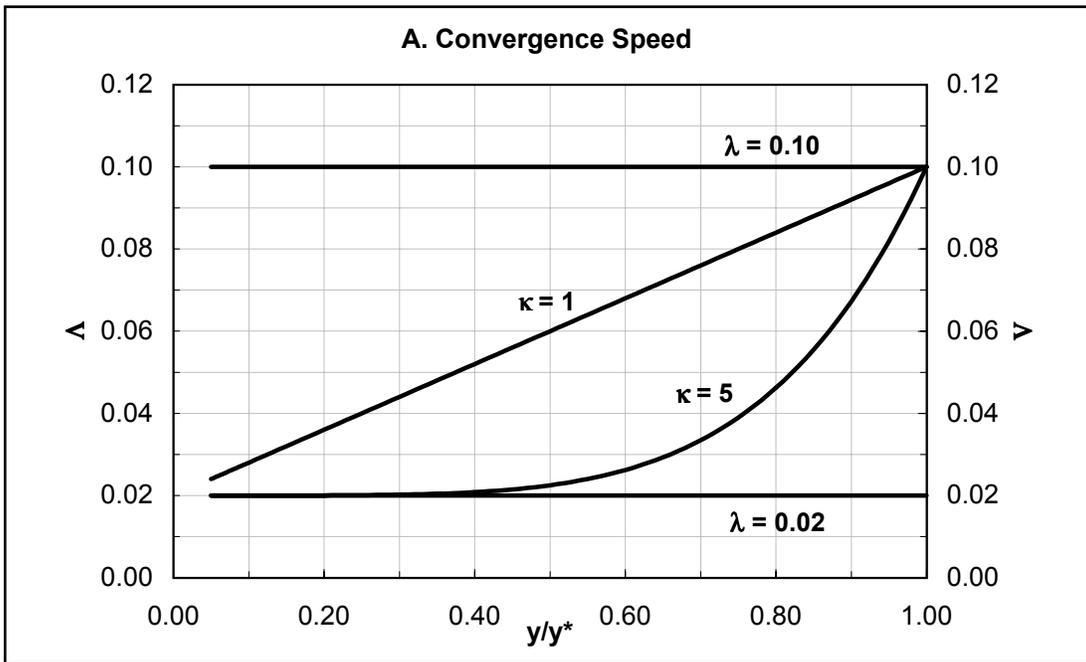


Figure assumes a data generating process:

$$\Delta = (\lambda_r - \lambda_p) \cdot \exp(-\kappa|\log(y) - \log(y^*)|) + \lambda_p$$

$$d/\text{dt} \log(y) = -\Delta \cdot (\log(y) - \log(y^*))$$

Figure 2: "Standard" Nonlinear Least Squares Surface

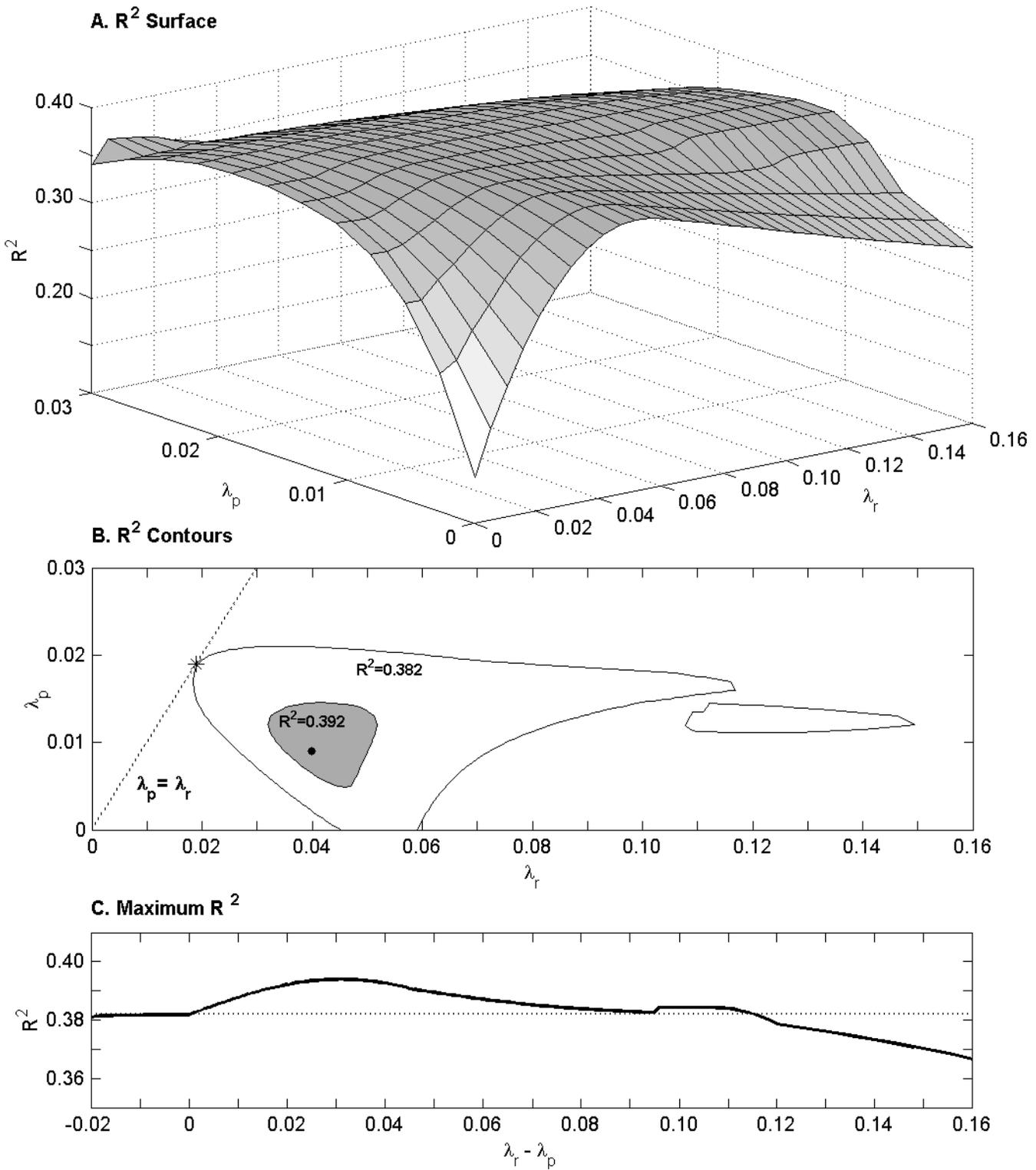


Figure shows nonlinear least square estimate of the equation $d/dt y = -[(\lambda_r - \lambda_p) \exp(-\kappa |y - y^*|) + \lambda_p] \cdot (y - y^*) + dy^*/dt$ where y is the log of initial income and y^* is the log of steady-state income estimated using the standard specification described in Table 2.

Figure 3: "Exogenous Wide Constrained" NLS Surface

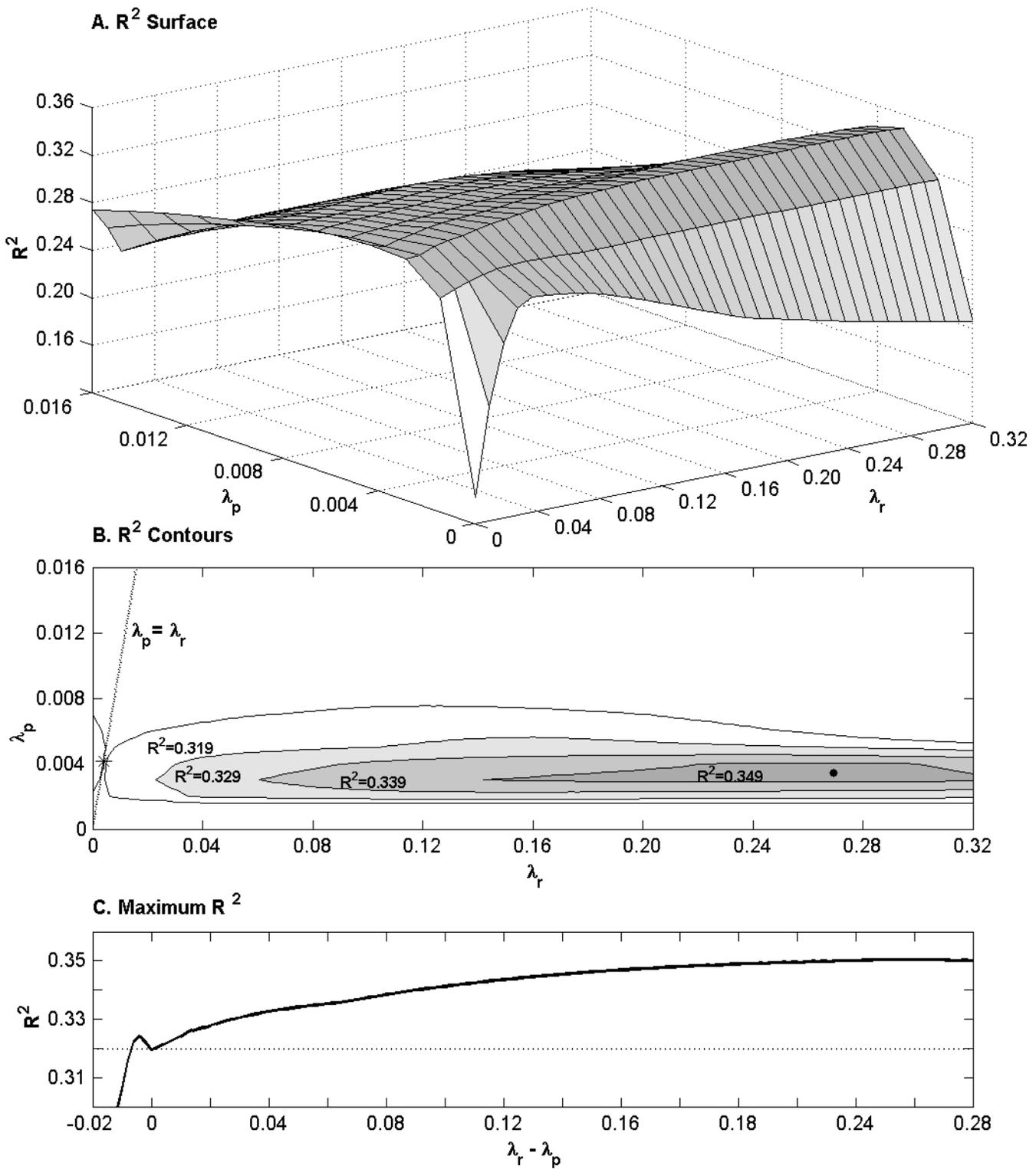


Figure shows nonlinear least square estimate of the equation $d/dt y = -[(\lambda_r - \lambda_p) \exp(-\kappa |y - y^*|) + \lambda_p] \cdot (y - y^*) + dy^*/dt$ where y is the log of initial income and y^* is the log of steady-state income estimated using the exogenous specification described in Table 2.

Figure 4: "Exogenous Narrow Constrained" NLS Surface

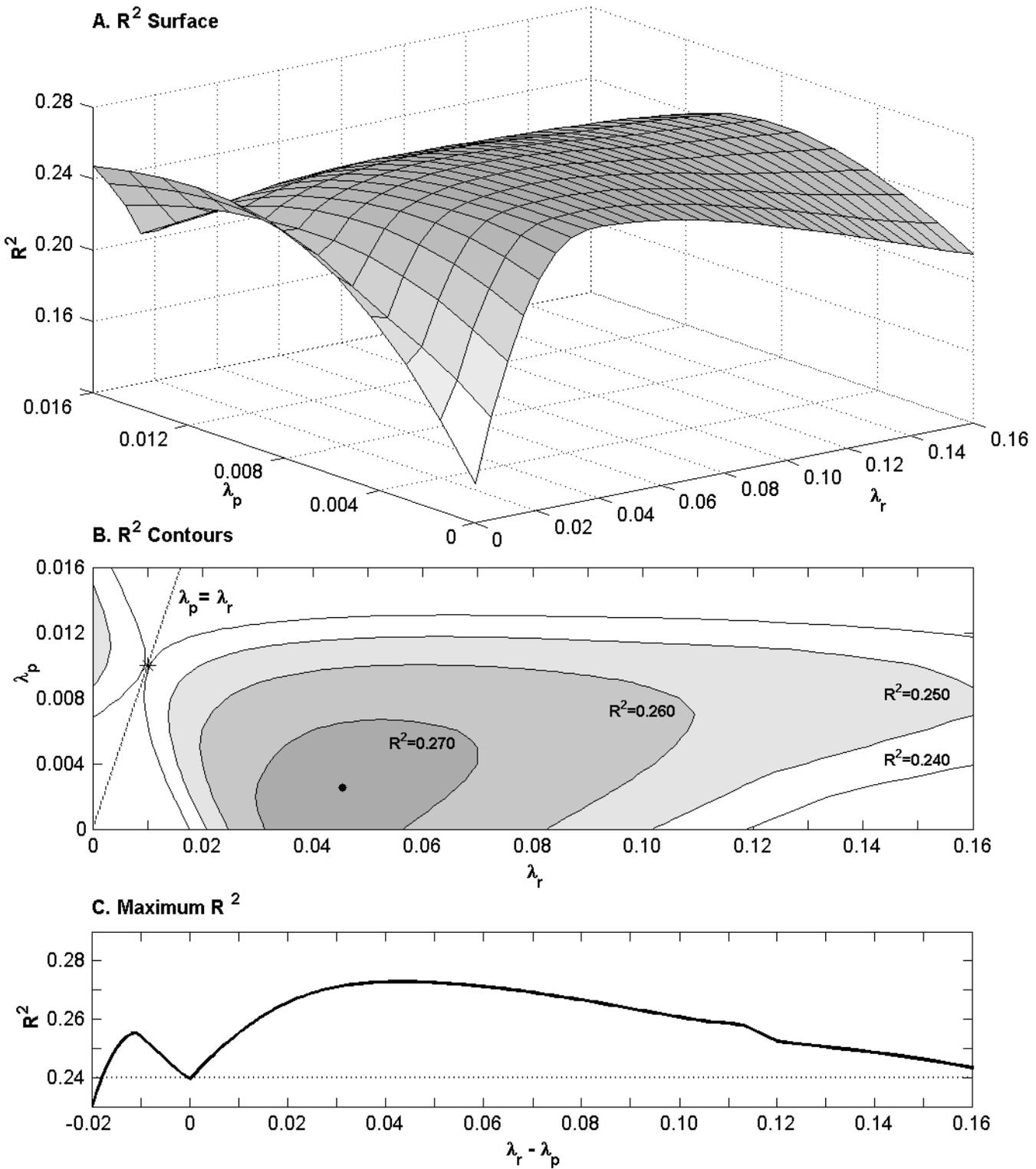


Figure shows nonlinear least square estimate of the equation $d/dt y = -[(\lambda_r - \lambda_p) \exp(-\kappa |y - y^*|) + \lambda_p] \cdot (y - y^*) + dy^*/dt$ where y is the log of initial income and y^* is the log of steady-state income estimated using the exogenous specification described in Table 2 with coefficients constrained to lie within two standard deviations of level estimates in Table 5.

Figure 5: "Unconditional" Nonlinear Least Squares Surface

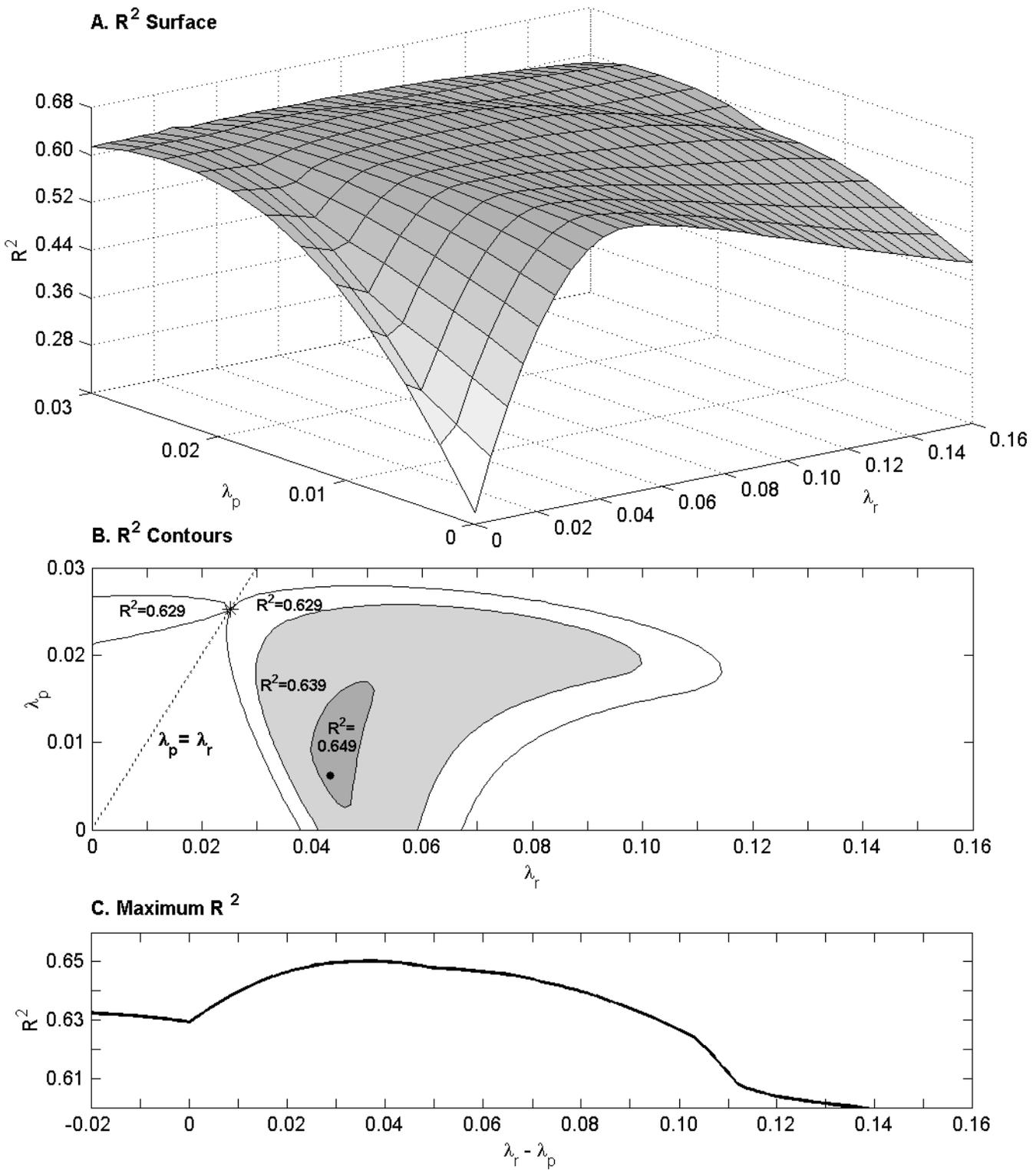


Figure shows nonlinear least square estimate of the equation $d/dt y = -[(\lambda_r - \lambda_p) \exp(-\kappa |y - y^*|) + \lambda_p] \cdot (y - y^*) + dy^*/dt$ where y is the log of initial income and y^* is the log of steady-state income estimated using the unconditional specification described in Table 2

Figure 6: Increasing Convergence Speed Posterior Odds

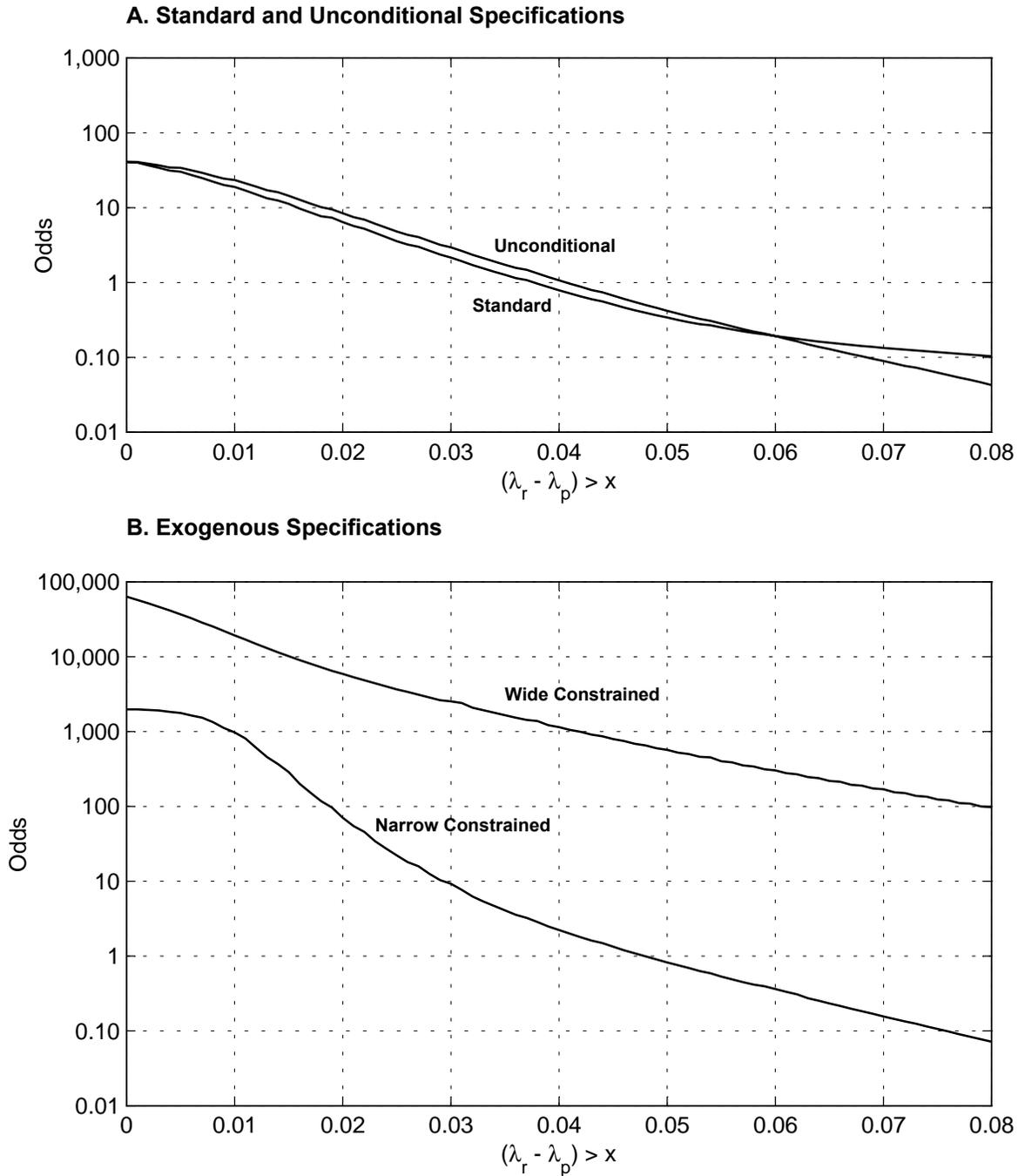


Figure shows posterior ratio of the integral of the likelihood function for $(\lambda_r - \lambda_p) > x$ relative to $(\lambda_r - \lambda_p) \leq x$ for $\lambda_p \in [0, 0.03]$ and $\lambda_r \in [0, 0.32]$.