Regional Convergence and Aggregate Growth

(Preliminary Version)

Stéphane Adjemian*, Jérôme Glachant† and Charles Vellutini‡

January, 2000

Abstract

A striking feature of US states convergence is the link between the spatial speed of convergence and the aggregate growth rate: fast aggregate growth induces a reduction in regional inequalities. This paper uses a neoclassical growth framework with integrated economies in order to capture this phenomena. As it has been stressed by Ventura [1997], the interdependence between regional economies through the access to common markets generates a link between aggregate evolution and spatial convergence dynamics. The paper has two mains results. First, we show how deep parameters of the economy determines quantitatively the magnitude of this link. Second, we propose two directions for testing the model and we provide some empirical evidence using US states data on personal income. These results are mixed, only a part of the convergence pattern is well captured by the model.

1 Introduction

Regional convergence appears to be time varying, and, in most cases, gradually slowing down. In the USA, Sherwood-Call [1996] has documented the divergence in state incomes that took place in the 80's, after decades of convergence. Martin [1997] reports that the speed of convergence among European regions fell from 2% to 1.3%, also in the 80's. Barro and Sala-I-Martin [1995, chap. 11] have evidenced a significant fall in the speed of convergence among Japanese prefectures after 1955. De La Fuente [1998] reports that convergence in Spanish regional incomes also shows a marked tendency to decelerate.

*EPEE, Université d'Evry-Val-d'Essonne, bld. François Mitterrand, 91025 Evry Cedex, France. <*adjemian@eco.univ-evry.fr*>

†EPEE, Université d'Evry-Val-d'Essonne. <*glachant@eco.univ-evry.fr*>

‡EUREQUA, Université de Paris I, 106-112, bld de l'Hôpital - 75647 Paris Cedex 13, France. <*vellutin@club-internet.fr*>

In figure 1, the solid line stands for the speed of convergence among US states' relative incomes[1]. Without performing formal tests, the instability of the spatial speed of convergence seems to be a crucial feature of US regional growth dynamics. In the same figure a dashed line represents the annual average, over ten years intervals, growth rate of US national income. It appears that the series are strongly correlated (the correlation is 0.8055), highlighting the link between aggregate growth and regional convergence.

< Insert figure 1. around here >

This paper examines how a neoclassical growth model of integrated economies may explain such stylized fact. Using US states income data, we estimate an extended system of growth regression, which make explicit the link between spatial convergence and aggregate growth. Our results are mixed, only a part of the convergence pattern is well captured by the model.

We use the neoclassical model that has been constructed by Caselli and Ventura [2000] to study distributive dynamics among infinitely lived individuals in a particular economy. We follow Bliss [1995] and Ventura [1997], who have applied the model to study inequality among regions/countries. Our basic framework is a general equilibrium growth model, as this is a simple and natural way to deal with interactions among regional economies. Factor price equalization is achieved without any restrictions and the optimal behavior of Ramsey savers determines the dynamics of wealth and income.

Bliss [1995] has shown that in this setting, globalization promotes long-run income inequality. More precisely, factor price equalization deters convergence so that initial differences in income persist forever. However, Caselli and Ventura [2000] and Ventura [1997] have shown that convergence *may* exist during the transition of the aggregate economy toward its steady-state. In contrast with the closed model, conditional convergence is due to the properties of intertemporal demand – not diminishing returns. These articles do not indicate exactly how these theoretical results can be confronted with empirical evidence on the distribution of income across regions. There remains an important gap to bridge between these theoretical criticisms and their empirical counterparts. The paper provides an empirical strategy for testing the effect of integration on convergence.

The findings of this paper are as follows. First, the model provides one explanation for the deceleration of convergence reported above, as we simultaneously have a model that can display convergence during the transition and long-run persistence. The relationship between the dynamics of cross-section inequality and aggregate growth is made explicit by a linearization around the steady-state. This local characterization makes it possible to relate the magnitude of the convergence or divergence effect to the parameters describing the fundamentals of economies. Second, we develop testing

[1]These estimates are obtained by estimating a cross-region growth equation (See Barro and Sala-I-Martin [1995]) with regional dummies over moving ten years intervals, on the period 1939-1997

strategies that are based on the linearized version of the model. A panel data study, which consists in estimating an "extended" system of growth rate regressions with fixed effect, can potentially discriminate between the model of the integrated economy presented below and the traditional model of autarchic economies. Third, we provide empirical evidence using US states data on personal income.

The paper is organized as follows. In the second section, we present the basic setup borrowed from Caselli and Ventura [2000]. The third section contains the core theoretical results of the article. We detail how the distribution dynamics displays both persistence and conditional convergence or divergence. Section 4 provides a testing strategy and section 5 the estimation results. We conclude in the last section

2 Basic framework

2.1 Structure of the economy and technology

2.1.1 Static structure

The aggregate economy consists of a collection of regions indexed by their relative labor productivity $\theta \in]0, \theta_{\max}]$. In *per capita* terms, the regional technology is given by:

$$Y(\theta) = f\left[K(\theta), A\theta\right] + A\phi(\theta), \tag{1}$$

where $Y(\theta)$ and $K(\theta)$ are respectively the domestic product and the domestic capital stock of region θ, both in *per capita* terms. A is the aggregate level of technological efficiency growing at a constant exogenous rate x and $f(\cdot)$ is a neoclassical production function. $\phi(\theta)$ is a constant parameter, either positive or negative. $\phi(\theta) < 0$ means that a subsistence consumption level absorbs part of the output. $\phi(\theta) > 0$ can be viewed as a fixed rent increasing output, for example some production using only specific factors that region θ alone possesses.

$(\theta, \phi(\theta))$ is therefore time-invariant and characterizes the technology used by region θ. To do away with certain technical difficulties, these quantities are specified in intensive terms[2]. We will use lower-case letters to denote intensive variables: for all *per capita* variable Z, $z = Z/A$.

We therefore have three sources of region heterogeneity: capital stock $K(\theta)$, labor productivity θ and the parameter $\phi(\theta)$. We note $q(B)$ the number of regions with productivity index $\theta \in B$. We normalize θ and the size of the aggregate economy so that $\int_0^{\theta_{\max}} \theta q(d\theta) = \int_0^{\theta_{\max}} q(d\theta) = 1$. It is assumed that the population grows at the exogenous rate n and that the population is

[2]Ben-David [1998] studies the case in which the subsistence level is not indexed on technological change. Interestingly, endogenous formation of clubs may arise.

identical in any region[3]. In this context, A is the average labor productivity in the aggregate economy. We note the average value of the $\phi(\theta)'s$ as $\phi \equiv \int_0^{\theta_{\max}} \phi(\theta)q(d\theta)$. Under the condition $\phi \neq 0$, we then define the relative value of $\phi(\theta)$ as $\phi_R(\theta) \equiv \phi(\theta)/\phi$.

Before we proceed to integrate the economy, let us first recall that in regional autarchy the quantity $f\left[K(\theta), \theta A\right] + \phi(\theta)A$ is both the GNP and the GDP of region θ. From this point on, we assume that the economy is integrated through perfect capital mobility. It is now necessary to distinguish between the capital owned by a given region, noted $k(\theta)$, and the capital used by that region, noted $\widehat{k}(\theta)$. The quantity $k(\theta) - \widehat{k}(\theta)$ is accordingly the portion of region $\theta's$ capital installed abroad. The gross rental rate of capital is noted $r + \delta$, with $\delta > 0$ the capital depreciation rate. With an integrated market for capital, $r + \delta$ must be identical across regions. In each country, competitive firms equate the gross marginal productivity of capital to the gross rental rate:

$$f_1\left[\widehat{k}(\theta), \theta\right] = r + \delta. \tag{2}$$

With $f(\cdot)$ homogeneous of degree one, equation (2) implies $\widehat{k}(\theta) = \theta\widehat{k}(1)$. With our normalization, national and average quantities are equal:

$$\widehat{k} \equiv \int \widehat{k}(\theta)q(d\theta) = \int k(\theta)q(d\theta) \equiv k. \tag{3}$$

k is the national capital stock, the average installed capital stock as well as the capital stock installed in the average region. The capital stock installed in θ therefore satisfies:

$$\widehat{k}(\theta) = \theta\widehat{k} = \theta k. \tag{4}$$

We can now write the GNP of region θ as

$$y(\theta) = \theta f(k) + \phi(\theta) + (r + \delta)(k(\theta) - \theta k), \tag{5}$$

where $f(k) \equiv f(k, 1)$.

The real wage in country θ satisfies:

$$w(\theta) = \theta f_2(\hat{k}, \theta) = \theta f_2(k, 1) = \theta w. \tag{6}$$

Hence, in the integrated economy, factor prices ($r(t)$ and $w(t)$) are identical in any region, for all t.

The world output is:

$$y = \int \left(f\left[\widehat{k}(\theta), \theta\right] + \phi(\theta)\right) q(d\theta) = f(k) + \phi.$$

As is well known, the complete integration of the national economy implies instantaneous *conditional* convergence (at the precise moment when

[3]This does not imply any loss of generality, while it simplifies the notation. Our results obtain with any distribution of the total population.

the world economy is integrated) of the gross domestic product (adjusted for $\phi(\theta)$):

$$\widehat{y}(\theta) - \phi(\theta) = \theta(y - \phi), \text{ for all } \theta. \tag{7}$$

In contrast, there is no instantaneous convergence effect affecting the GNP $y(\theta)$. This quantity depends on region $\theta's$ wealth, as is shown in equation (5).

2.1.2 Dynamic structure

Time is continuous. The capital of region θ is driven by:

$$\dot{k}(\theta) = \theta w + rk(\theta) + \phi(\theta) - c(\theta) - (n + x)k(\theta), \quad k(\theta, 0) \text{ given}, \tag{8}$$

with $c(\theta)$ the consumption level of region θ.

We rule out Ponzi games by assuming:

$$\lim_{t \to \infty} e^{-R(0,t)} e^{(x+n)t} k(\theta, t) \geq 0, \tag{9}$$

where $R(0, t) \equiv \int_0^t r(s)ds$.

Aggregating these equations over regions yields the law of motion of the national capital stock:

$$\dot{k} = f(k) + \phi - (\delta + x + n)k - c, \ k(0) \text{ given}. \tag{10}$$

2.2 Households

The representative household of region θ maximizes:

$$U(\theta) = \int_0^\infty \frac{C^{1-\sigma}(\theta, t) - 1}{1 - \sigma} e^{nt} e^{-\rho t} dt, \tag{11}$$

subject to constraints (8-9) and taking the time paths of prices as given. $C(\theta, t)$ is *per capita* consumption in region θ at time t, $\sigma > 0$ the inverse of the constant intertemporal elasticity of substitution, and $\rho > 0$ the utility discount rate. In contrast with autarchy, the assumption that households maximize intertemporal utility – as opposed to the choice of some exogenous saving rates – is central to the convergence results of the integrated model. This point will become clearer later.

With these preferences, it is possible to interpret $\phi(\theta)$ as a measure of intertemporal flexibility: the larger $\phi(\theta)$, the more flexible is region θ in its intertemporal allocation of consumption. With a negative $\phi(\theta)$, region θ will not be able to substitute consumption through time above a certain level. Conversely, a positive $\phi(\theta)$ means that there is at any time a constant source of output which is by definition completely independent of region $\theta's$ time allocation problem and therefore raises its ability to substitute consumption intertemporally.

Note that by simply defining $C_{sg}(\theta) = C(\theta) - \phi(\theta)e^{xt}$ together with technology $y(\theta) = f[k(\theta), \theta]$, it is possible to rewrite the maximization problem above with a standard Stone-Geary intertemporal utility function. This setting is fully equivalent to the one used here, but makes the economic interpretation of a positive $\phi(\theta)$ more difficult.

As pointed out by Caselli and Ventura [2000], these preferences thus make it possible to describe the aggregate economy as a hypothetical representative region endowed with exactly average characteristics. Aggregate paths are found by solving the usual autarchy problem, which leads to:

$$\begin{aligned}\dot{c} &= c\left[\sigma^{-1}\left(f'(k)-\delta-\rho\right)-x\right], &(12)\\ \dot{k} &= f(k)+\phi-c-(n+\delta+x)k, \quad k(0) \ > 0 \text{ given.} &(13)\end{aligned}$$

This system is appended with the transversality condition, that is, inequality (9) taken as an equality. As is well known, we need the following condition on parameters so that this condition always holds in equilibrium:

$$\rho > n + (1-\sigma)x. \tag{14}$$

In the rest of the paper we will find convenient to note the discounted flow of any variable x as: $\tilde{x}(t) \equiv \int_t^\infty x(\tau)e^{-R(t,\tau)}e^{(\tau-t)(x+n)}d\tau$.

Because of the homothetic properties of preferences, the consumption of any region θ can be written as a linear function of its total wealth, for all t[4]:

$$c(\theta,t) = \nu(t)a(\theta,t), \tag{15}$$

where

$$a(\theta,t) = k(\theta,t) + \theta\tilde{w}(t) + \phi(\theta)\tilde{1}(t), \tag{16}$$

and

$$\nu(t) = \left[\int_t^\infty e^{R(t,\tau)(1-\sigma)/\sigma-(\tau-t)(\rho/\sigma-n)}d\tau\right]^{-1}. \tag{17}$$

The key point in (15) is that $\nu(t)$, the propensity to consume out of total wealth, is the same for any region θ, depending only on the aggregate behavior of the national economy. This fact implies that the total wealth of any region grows at a rate given by the aggregate economy alone. A simple expression for that growth rate can be found by taking the time derivative of (16)[5] and substituting from (8):

$$\frac{\dot{a}(\theta,t)}{a(\theta,t)} = r(t) - \nu(t) - n - x. \tag{18}$$

This is a crucial feature of the model. In an economy where factor prices are equated across regions by factor mobility and/or interregional trade, and where preferences are homothetic (in total wealth), all regions accumulate total wealth at precisely the same rate – regardless of relative levels at any

[4] Combine the intertemporal budget constraint $\tilde{c}(t) = a(t)$ with the integral version of (12).

[5] Notice that $\dot{\tilde{x}} = (r-n-x)\tilde{x} - x$.

time. Remember that $a(\theta, t)$ is regional wealth, which is the sort of wealth that matters in an economy where regions may export capital – if we want to be able to interpret regional integration as capital mobility. This obviously is an important departure from the vision of the national economy as a collection of neoclassical closed regional economies. In that sort of economy, poor regions always grow faster that rich ones, both in terms of total wealth and in terms of capital. We will now turn to the growth rates of regional stocks of capital in the integrated economy.

3 Transitional convergence and long-run persistence

3.1 Transitional convergence

An interesting form of the law of motion of region $\theta's$ capital obtains by substituting (15) into (8):

$$\dot{k}(\theta,t) = [r(t) - \nu(t) - x - n]\, k(\theta,t) + [w(t) - \nu(t)\widetilde{w}(t)]\, \theta + \left[1 - \nu(t)\widetilde{1}(t)\right] \phi(\theta). \tag{19}$$

Let $h(\theta,t) \equiv k(\theta,t)/k(t)$ denote the relative capital of region θ. Equation (19) and its aggregate version then imply:

$$\dot{h}(\theta,t) = -(\psi_1(t) + \psi_2(t))h(\theta,t) + \psi_1(t)\theta + \psi_2(t)\phi_R(\theta), \tag{20}$$

where $\psi_1(t)$ and $\psi_2(t)$ are two measures of convergence defined as

$$\psi_1(t) \equiv \frac{w(t) - \nu(t)\widetilde{w}(t)}{k(t)} \text{ and } \psi_2(t) \equiv \phi\frac{1 - \nu(t)\widetilde{1}(t)}{k(t)}.$$

Note that both $\psi_1(t)$ and $\psi_2(t)$ are invariant across regions (but not across time). Relationship (20) shows how the distribution of financial wealth changes over time. This expression, which has been emphasized by Caselli and Ventura [2000], is interesting because the three sources of heterogeneity are clearly isolated. What do we learn? First, note that the value of $\psi(t) \equiv \psi_1(t) + \psi_2(t)$ determines the extent of the *conditional convergence effect*. Imagine that regions do not differ in their fundamentals θ and $\phi(\theta)$. Then, from (20), it can be seen that the distribution of financial wealth will shrink if $\psi(t)$ is positive and will expand if $\psi(t)$ is negative. $\psi(t)$ is thus the instantaneous speed of conditional convergence. This speed applies to any region of the national economy.

Second, *absolute convergence* can also be analyzed through equation (20). One can see that the distributions of θ and $\phi_R(\theta)$ are repulsive or attractive depending on the sign of $\psi_1(t)$ and $\psi_2(t)$. For a particular region, the occurrence of absolute catching-up depends on *both* the sign of $\psi_1(t)$ and $\psi_2(t)$. Knowing the time paths of the aggregate values $\psi_1(t)$ and $\psi_2(t)$, one will know how regions move toward their long-run position. In subsection 3.3, we study how these coefficients change in the neighborhood of the aggregate steady-state. The occurrence of local conditional convergence or divergence depends on the parameters describing the fundamentals of the economies. Caselli and Ventura [2000] have shown that the model with a CES production function can display transitional divergence or even successive periods of convergence and divergence, a phenomenon reminiscent of a Kuznets curve.

In fact, in this economy, a key determinant of transitional convergence is the elasticity of substitution between factors. Consider (16) and (18): regardless of the share of financial wealth in its total wealth, any region will find it optimal to accumulate the latter at the same rate. Think of a high elasticity of substitution. As the aggregate economy accumulates capital, the demand for labor will tend to be relatively low because the economy will increasingly substitute capital for labor. So will be the discounted flow of wages, a component of total wealth ("human" wealth). Finally, one sees that to keep total wealth on an optimal path, regions poorly endowed with financial wealth will need to accumulate it at a quicker rate than the financially rich. An alternative expression for ψ illustrates this point. From (18):

$$\psi(t) = \frac{\dot{k}(t)}{k(t)} - \frac{\dot{a}(t)}{a(t)}. \tag{21}$$

This relationship shows that if, *in aggregate terms,* the growth rate of capital is higher than that of total wealth, there will be transitional conditional convergence.

Another point is worth emphasizing here. The important underlying mechanism for transitional convergence is the homothetic property of household preferences, not diminishing returns to capital like in the autarchic model. In this economy, transitional divergence in capital stocks is plausible even though technology is one with diminishing returns.

3.2 Long-run persistence of the cross-section distribution

Integrating (20) yields:

$$h(\theta,t) = \lambda_1(0,t)h(\theta,0) + \lambda_2(0,t)\theta + \lambda_3(0,t)\phi_R(\theta), \tag{22}$$

where

$$\begin{aligned} \lambda_1(0,t) &= \exp[-\int_0^t \psi(s)ds], & (23)\\ \lambda_2(0,t) &= \int_0^t \psi_1(s)\lambda_1(s,t)ds, & (24)\\ \lambda_3(0,t) &= \int_0^t \psi_2(s)\lambda_1(s,t)ds. & (25) \end{aligned}$$

This result has been provided by Caselli and Ventura [2000]. What is interesting in this relationship is that, again, $\lambda_1(0,t)$, $\lambda_2(0,t)$ and $\lambda_3(0,t)$ are independent of θ. This equation therefore provides a very transparent decomposition of the distribution of financial wealth at any point in time, making explicit the respective contributions of the initial financial wealth distribution, the distribution of labor productivity and the distribution of the $\phi(\theta)'s$. Not surprisingly, given the construction of (22), the $\lambda_i's$ sum up to unity, for all (t,t')[6]:

$$\lambda_1(t,t') + \lambda_2(t,t') + \lambda_3(t,t') = 1. \tag{26}$$

The variable $\lambda_1(t,t')$ is a measure of cumulated conditional convergence between t and t', as opposed to $\psi(t')$, which may be viewed as the instantaneous conditional convergence at instant t'. It can be seen by direct examination of (22) taken between t and t' that if $\lambda_1(t,t')$ is less that unity, then there will be cumulated conditional convergence over that period of time.

What does (22) tell us on the asymptotic distribution of financial wealth? In the long run the distribution reads:

$$h(\theta,\infty) = \lambda_1(0,\infty)h(\theta,0) + \lambda_2(0,\infty)\theta + \lambda_3(0,\infty)\phi_R(\theta). \tag{27}$$

There are no particular reasons why $\lambda_1(0,\infty)$ should be equal to zero. Evidently, the distribution of long-run financial wealth will be also influenced by the other two sources of heterogeneity, respectively in θ and $\phi(\theta)$. But the important fact is that we have a situation of long-run persistence of the financial wealth distribution. In other words, there may well be transitional convergence – or divergence –, as was shown in section 3.1, but in all cases these dynamics will come to an halt as the aggregate economy proceeds towards a steady-state. This result is in line with the one obtained by Bliss [1995].

Note that long-run persistence of initial distributions applies in exactly the same way to consumption: just consider (15) and (16). Income distribution dynamics are slightly more complex because the shares of respectively capital and labor incomes in total income may vary over time. It is nonetheless not difficult to show that long-run persistence also applies. Defining the share of capital income in total income $\alpha(t) \equiv r(t)k(t)/y(t)$ and the share of labor income as $\beta(t) \equiv w(t)/y(t)$, an expression for the relative income of country θ, $y_R(\theta,t) \equiv y(\theta,t)/y(t)$, can be found by substitution into (22):

$$y_R(\theta,t') = \gamma_1(t,t')y_R(\theta,t) + \gamma_2(t,t')\theta + \gamma_3(t,t')\phi_R(\theta), \tag{28}$$

where

$$\gamma_1(t,t') = \frac{\alpha(t')}{\alpha(t)}\lambda_1(t,t'), \tag{29}$$

$$\gamma_2(t,t') = \beta(t') + \alpha(t')\lambda_2(t,t') - \frac{\alpha(t')}{\alpha(t)}\beta(t)\lambda_1(t,t'), \tag{30}$$

$$\gamma_3(t,t') = 1 - \gamma_1(t,t') - \gamma_2(t,t'). \tag{31}$$

[6] One way to do this is to sum (22) over the $\theta's$.

A different way of writing equation (28) is to make explicit the income "target" of region θ:

$$y_R(\theta, t') = \gamma_1(t,t')y_R(\theta,t) + (1-\gamma_1(t,t'))y_R^\star(\theta,[t,t']), \tag{32}$$

where the income target $y_R^\star(\theta,[t,t'])$ is defined as:

$$y_R^\star(\theta,[t,t']) = \frac{\gamma_2(t,t')}{\gamma_2(t,t')+\gamma_3(t,t')}\theta + \frac{\gamma_3(t,t')}{\gamma_2(t,t')+\gamma_3(t,t')}\phi_R(\theta). \tag{33}$$

On the one hand, these results are reminiscent of the autarchic model, in the sense that the dynamics is still guided by a gap between the current value of the variables (e.g. relative income $y_R(t)$) and some long run "target". On the other hand, observe that in the integrated economy, this target is time-varying and, more importantly, can never be attained. The model contrasts sharply with the long-run behavior of economies in the autarchic model. One key result of that setting is that any regional economy converges to a unique steady-state level of wealth and income – conditional on structural parameters such the labor productivity and preferences. This process will eventually iron out initial income differences. Here, on the contrary, the effects of initial wealth will be felt forever. Each region will reach its own particular steady-state level of financial wealth and income – even if conditioned on structural parameters. Put differently, there are no stationary wealth and income distributions.

3.3 Characterization around the aggregate steady-state

One way to understand how the distribution of financial wealth changes as the aggregate economy grows towards its steady-state is to linearize equation (22) around the aggregate steady-state. To do this, we need to study the coefficients $\lambda_i(t,t')$, $i=1,2,3$, or alternatively $\psi(t), \psi_1(t)$ and $\psi_2(t)$, around the steady-state.

Note that for this purpose it is easiest "to eliminate time" and to work with aggregate quantities as functions of the capital stock k, not t. A well-known example of this are the policy rules $c(k), a(k), w(k)$,etc. Similarly, we define $\lambda_1(k,k')$ as the value taken by the coefficient λ_1 when the aggregate economy starting from k moves to k'. In the same fashion, we can define $\lambda_i(k,k')$, $i=2,3$, $\psi_i(k)$, $i=1,2$ and $\psi(k)$. Let $k^\star$ denote the steady state aggregate level of capital.

First, these functions satisfy: $\lambda_1(k^\star,k^\star)=1$, $\lambda_2(k^\star,k^\star)=\lambda_3(k^\star,k^\star)=0$, and $\psi_1(k^\star)=\psi_2(k^\star)=\psi(k^\star)=0$. Second, we define $\eta_i^\star$, $i=1,2,3$ as the semi-elasticities of the function $\lambda_i(k,k^\star), i=1,2,3$ with respect to k, evaluated at $k=k^\star$:

$$\eta_i^\star \equiv k^\star \left(\frac{\partial\lambda_i(k,k^\star)}{\partial k}\right)_{k=k^\star}, \quad i=1,2,3. \tag{34}$$

Taking the first order Taylor expansion of $\lambda_i(k,k')$ near the steady-state yields:

$$\lambda_i(k,k^\star) \approx \lambda_i(k^\star,k^\star) + \eta_i^\star \left[\frac{k-k^\star}{k^\star}\right], \; i=1,2,3. \tag{35}$$

Moreover, since the coefficients $\lambda_i, i=1,2,3$ sum up to unity, we have:

$$\eta_1^\star + \eta_2^\star + \eta_3^\star = 0. \tag{36}$$

Assuming that the initial aggregate capital stock $k(0)$ is not too far off $k^\star$, and neglecting second order terms, the long-run distribution of relative financial wealth is given by:

$$h(\theta,\infty) = \left[1-\eta_1^\star\left(\frac{k^\star}{k(0)}-1\right)\right] h(\theta,0) - \eta_2^\star\left(\frac{k^\star}{k(0)}-1\right)\theta - \eta_3^\star\left(\frac{k^\star}{k(0)}-1\right)\phi_R(\theta). \tag{37}$$

This equation can be interpreted as follows. Imagine that the aggregate economy is below its steady-state ($k(0) < k^\star$), so that the economy experiences an episode of growth at a cumulated rate $k^\star/k(0)-1$. Then, the initial distribution of the $h(\theta,0)'s$ shrinks or expands – depending on the sign of $\eta_1^\star$. If $\eta_1^\star > 0$, aggregate growth implies a phenomenon of conditional convergence. Relative productivity level θ has a positive influence on long-run relative wealth as long as $\eta_2^\star < 0$. The influence of $\phi(\theta)$ depends on the sign of $\eta_3^\star$.

In Appendices B and C, we show that the coefficients $\eta_i^\star, i=1,2,3$ can be expressed as functions of the deep parameters of the model:

$$\begin{aligned}
\eta_1^\star &= 1 - \frac{k^\star}{c^\star}\left[\rho + \sigma(x+\mu) - (n+x)\right], && (38)\\
\eta_2^\star &= \frac{k^\star}{c^\star}\sigma\mu - \frac{w^\star}{c^\star}, && (39)\\
\eta_3^\star &= -\frac{\phi}{c^\star}, && (40)
\end{aligned}$$

where $\mu > 0$ is the speed of convergence of the aggregate economy toward its steady-state (see Appendix A).

Moreover, we show in the appendices that:

$$\psi'(k^\star) = -\mu\frac{\eta_1^\star}{k^\star}, \;\; \psi_1'(k^\star) = \mu\frac{\eta_2^\star}{k^\star}, \;\; \psi_2'(k^\star) = \mu\frac{\eta_3^\star}{k^\star}. \tag{41}$$

These equations provide a full characterization of the cross-section dynamics around the steady-state. In particular, we see how the parameters describing preferences and technology influence the distribution of wealth across regions.

One of the most interesting results is the negative relationship between μ, the aggregate speed of convergence, and $\eta_1^\star$, a measure of the conditional convergence effect. We see that a slow adjustment of the world economy translates into a low long-run persistence of the distribution of financial wealth. Ventura [1997] already identified this property of the model, but without quantifying it. Knowing that ϵ, the elasticity of substitution of the production function, and σ, the inverse of the intertemporal elasticity of substitution, both impact negatively on μ, one can conclude that the size of the long-run persistence effect decreases with these parameters. By contrast, this size increases with ϕ, the average level of the $\phi(\theta)$'s.

From (38), one can see that the most realistic configuration of the parameters leads to a conditional convergence effect[7]. A conditional divergence effect would take place only for high values of μ and σ. Also, the influence of $\phi(\theta)$ on the long-run wealth is always positive. The influence of the relative productivity level θ depends on the sign of $\eta_2^\star$. Some kind of substitution effect between financial and human wealths cannot be ruled out. Indeed, $\eta_2^\star > 0$ would mean that regions poorly endowed in terms of θ make up for their low labor productivity by accumulating financial wealth at a faster rate. It could be the case for high values of σ and μ[8].

Starting from (28), we can use the same analysis to study the behavior of the distribution of relative gross income $y_R(\theta)$. Appendix D focuses on the determination of the semi-elasticities $\omega_i, i = 1, 2, 3$ of the coefficients γ_i. The results obtained for the distribution of relative wealth $h(\theta)$ must be amended so as to take into account the fact that respective factor shares in total income may be variable around the aggregate steady state. This exercise clarifies the influence of a high elasticity of substitution on the income distribution. First, a high value of ϵ tends to depress the growth of wages and hence favors conditional convergence of relative wealth, as was noted above. But we have a second effect here. A high value of ϵ also implies a strong decrease of the share of wage in total income when the aggregate economy converges toward its steady-state. This second effect weakens conditional convergence.

4 Strategies for empirical testing

In this section, we give directions along which the above results could be tested. A key preoccupation is to focus on the discriminating predictions of the model – particularly, on what distinguishes the integrated national economy from a collection of autarchic regions. A prediction of the model which clearly possesses this discriminating property is the long-run persistence of initial conditions, in other words, that convergence, when it exists, will never be complete. To simplify our exposition in this section, we assume that θ is the only source of heterogeneity in the fundamentals. More precisely: $\phi(\theta) = 0$ for all θ.

[7]It is important to remember that $\rho + \sigma x$ is the net rate of interest.

[8]Remember that the $\eta_i^\star$ relate to one another by (36). Consequently, conditional divergence implies some substitution effect.

4.1 Integrated economy vs autarchic economies

A natural route to explore the difference between autarchic and integrated economies is to express the distribution of relative incomes as a function of past relative incomes and relative fundamentals. In the integrated economy, relationship (32) is available. By contrast, such expression does not exist in the autarchic case as the dynamics of aggregate income depends on the distribution of capital across regions.

A way to solve this difficulty is to linearize the autarchic dynamics around the steady state, as the distribution of capital among regions is uniquely determined near the steady state. For any country θ, we have:

$$y(\theta, t') - y^\star(\theta) = e^{-\mu(t'-t)} \left(y(\theta, t) - y^\star(\theta) \right), \tag{42}$$

with μ the speed of convergence of a region toward its steady-state and $y^\star(\theta) = \theta y^\star$ the steady-state income of country θ. Note that we used the same symbol μ for the speed of convergence because such speed is exactly equal to the μ characterizing the aggregate dynamics of the integrated economy.

¿From this, we obtain:

$$y_R(\theta, t') = e^{-\mu(t'-t)} y_R(\theta, t) + \left(1 - e^{-\mu(t'-t)}\right) \theta. \tag{43}$$

The integrated-economy analogue to this expression, with $\phi(\theta) = 0$, comes directly from (32):

$$y_R(\theta, t') = \gamma_1(t, t') y_R(\theta, t) + (1 - \gamma_1(t, t'))\theta. \tag{44}$$

The comparison (43) and (44) is straightforward. The term $e^{-\mu(t'-t)}$ will tend to zero as t' grows, whereas its integrated-economy counterpart, the quantity $\gamma_1(t, t')$, will not. Initial conditions have a persistent effect in the integrated economy. This persistence phenomenon does not exist in the autarchic economy and convergence will always translate into complete catching-up.

Another way of comparing the respective predictions of the two models is to look at growth rates. To do this, we now have to use the linearized expression of $\gamma_1(t, t')$ (see Appendix D). We therefore assume that the aggregate economy is not too far from a steady state. Let us define :

$$g_R(\theta, t, t') \equiv \frac{y_R(\theta, t') - y_R(\theta, t)}{y_R(\theta, t)}$$

as the growth rate of relative income. From (43), an approximate expression for this growth rate in the autarchic model is:

$$g_R(\theta, t, t') = \left(1 - e^{-\mu(t'-t)}\right) \left[\frac{\theta}{y_R(\theta, t)} - 1 \right]. \tag{45}$$

In the integrated economy, one can use the fact that[9] $1-\gamma_1(t,t') \approx \omega_1^\star\left(\frac{y(t')}{y(t)}-1\right)$, together with (44), results into:

$$g_R(\theta,t,t') = \omega_1^\star\left(\frac{y(t')}{y(t)}-1\right)\left[\frac{\theta}{y_R(\theta,t)}-1\right]. \tag{46}$$

The comparison between (45) and (46) is again transparent. The behavior of an integrated economy differs from its autarchic counterpart in one key aspect: *the intensity of convergence effect depends on the aggregate dynamics.* As the aggregate economy proceeds toward its steady state, the convergence effect dies out.

4.2 A system of modified growth rate regressions

Let us assume that $t'-t=1$ so as to use yearly data and let $g_R(\theta,t)$ be the growth rate of relative income between $t-1$ and t. From this point on, y will no longer be the *intensive* level of income per head but will stand for the level of actual ("extensive") income per head. The two equations (45) and (46) lead to the following reduced form of the relative income dynamics:

$$g_R(\theta,t) = a_1+a_2\frac{y(t)}{y(t-1)}+a_3(\theta)\frac{y(t)}{y(\theta,t-1)}+a_4(\theta)\frac{1}{y_R(\theta,t-1)}+u(\theta,t). \tag{47}$$

In each underlying model, the coefficients $a_1, a_2(\theta), a_3(\theta)$ and a_4 are related to structural parameters in the following way:

Autarchic economy	Integrated economy
$a_1 = -(1-e^{-\mu})$	$a_1 = \omega_1^*$
$a_2 = 0$	$a_2 = \omega_1^*$
$a_3(\theta) = 0$	$a_3(\theta) = \theta\omega_1^* e^{-x}$
$a_4(\theta) = \theta(1-e^{-\mu})$	$a_4(\theta) = -\theta\omega_1^*$

$u(\theta,t)$ is an error term[10]. We allow for some spatial heteroskedasticity.

In addition to (47), one can estimate the following structural form:

$$g_R(\theta,t) = \left(b_1+b_2\frac{y(t)}{y(t-1)}\right)\left[\frac{\theta}{y_R(\theta,t-1)}-1\right]+u(\theta,t). \tag{48}$$

with the following restrictions on parameters:

[9]Alternatively, it is possible to use the log specification $1-\gamma_1(t,t') \approx \omega_1^\star \log\left(\frac{y(t')}{y(t)}\right)$.

[10]We prefer not to refer to "shocks", as the introduction of uncertainty at the agent level would raise complex theoretical issues related to the desire for some insurance against idiosyncratic shocs. The term $u(\theta,t)$ should be viewed as the sum of whatever is not explained by the model in the actual trajectories of the variables. It may account for example for variations known to the agents over their horizon of prevision, which are not modeled and are unobservable to the econometrician (see for example Altug and Labadie [1994, chap. 7]).

Autarchic economy	Integrated economy
$b_1 = 1 - e^{-\mu}$	$b_1 = -\omega_1^*$
$b_2 = 0$	$b_2 = \omega_1^* e^{-x}$

Note that structural parameters in both regimes are exactly identified. These equations may be complemented with the dynamics of the aggregate economy which is identical in the two models:

$$g(t) = \left(1 - e^{-\mu}\right) \left[\frac{e^{x(t-1)} y^*}{y_R(\theta, t-1)} - 1 \right].$$

4.3 σ-convergence

Another approach to testing the model relies on studying the changes over time of the cross-section variance of relative incomes. Since we will study an univariate time series, panel techniques for controlling possible heterogeneity in underlying parameters will not be available any longer. Taking into account heterogeneity in this setting is in theory still possible, but analytically overly complex. What we will analyze in this subsection is therefore *absolute* convergence, as opposed to conditional convergence. Most authors studying convergence in regional data sets have plausibly assumed homogeneity in structural parameters (see for example Barro and Sala-I-Martin [1995, chap. 11]). Formally, applying the variance operator to both sides of equation (44) and assuming that all regions have an identical level of labor productivity $\theta = 1$, we have the following law of motion for the cross-region variance of relatives incomes:

$$\sigma^2(t) = \sigma_u^2 + \gamma_1^2(t-1, t)\sigma^2(t-1) \tag{49}$$

where $\sigma^2(t) = V(y_r(\theta, t))$. We used the fact that the error term $u(\theta, t)$ appended to (48) is not correlated with $y_R(\theta, t-1)$ and assumed that the variance of $u(\theta, t)$, σ_u^2, is time-invariant.

The autarchic equivalent of this expression is:[11]

$$\sigma^2(t) = \sigma_u^2 + e^{-2\mu}\sigma^2(t-1) \tag{50}$$

There is, again, a crucial difference between the two equations. In the autarchic model, the multiplicative coefficient $e^{-2\mu}$ is time-invariant. Its counterpart in the integrated economy, $\gamma_1^2(t-1, t)$, will grow monotonically and is asymptotically equal to unity. Using the fact that γ_1 is not far from a steady state, one gets the approximation:

$$\gamma_1^2(t-1, t) \simeq 2\gamma_1(t-1, t) - 1 \simeq 1 - 2\omega_1^* \left(\frac{y(t)}{y(t-1)} - 1 \right). \tag{51}$$

[11] See for example Barro and Sala-I-Martin [1995], page 384.

By substituting (51) into (50), by introducing the exogenous technical progress x and adding an error term similar to $u(\theta, t)$, we finally get the regression to be tested:

$$\sigma^2(t) = \sigma_u^2 + (1 + 2\omega_1^*)\sigma^2(t-1) - 2\omega_1^* e^{-x} \frac{y(t)}{y(t-1)} \sigma^2(t-1) + \varepsilon(t). \quad (52)$$

5 Some evidence from the US states

The US states data seem to be an most ideal starting point for several reasons. First, as we noted above, there is substantial evidence of the convergence in state incomes slowing down. Second, data on per capital personal income have been compiled by the Bureau of Economic Analysis (BEA) every year from 1929. These are the longest existing series used in the literature on convergence: 69 years (1929-1997), for the 48 contiguous states. Third, as Kim [1997] reports, factors and goods are highly mobile across the US states. The assumption of perfect integration may be plausible in this economy.

The data has been obtained from the BEA's web page. Personal income is provided in nominal terms and it is not possible to deflate the data since state-specific deflators are not available. Consequently, relative income *per capita* are based on nominal levels while the federal consumer price index has been used to deflate aggregate nominal income *per capita*.

We follow the strategy described in the previous section. First, we use panel data techniques to estimate a system of modified growth rates regressions. Second, we study $\sigma-$convergence.

5.1 Growth regressions

We begin by an estimation of the reduced form (47) using weighted least-squares with White's correction to account for both spatial and temporal heteroskedasticities. The detailed results are shown in Tables (1 and 2). Most of the estimated coefficients differ significantly from zero. Moreover, the signs of these coefficients are consistent with the model of integration in which $\omega_1^* > 0$. Several tests are derived from this estimation:

- We test the restrictions on the parameters implied by the autarchic model. The null hypothesis is $H_0 : a_1 = a_3(\theta) = 0$, $\forall\theta$ and the associated Wald statistics is 803,31 which indicates a very strong rejection of the autarchic model.

- We test the restrictions on the parameters imposed by the integrated model. The null hypothesis is $H_0 : a_1/a_4 = a_2(\theta)/a_3(\theta) = 0$, $\forall\theta$ and leads to the structural form (48). The Wald statistics is 78.09 with an associated P-value of 0.0045. This means that the null hypothesis is also rejected. However, this rejection is not so strong as that of the autarchic model. Moreover, the result of this test is very sensitive to the method of estimation. For instance, without White's correction against temporal heteroskedasticity, the Wald statistics is 66.93 with a P-value of 0.045.

Second, we estimate the structural form (48) by weighted iterative least-squares. The estimated coefficients are shown in Table (2). Note that the table in the appendix provides estimates for both the coefficients b_1, b_2 and the structural parameters ω_1^* and x. The estimated values of b_1 and b_2 confirm that integration matters in the sense that the speed of convergence between regions appear to depend on aggregate dynamics. The structural parameter ω_1^* is 0.62 with a standard error of 0.053. This means that aggregate growth leads to a reduction of disparities between regions. However, the estimated value of the long-run growth rate x is disappointing. The counterfactual negative sign of this coefficient implies that a conditional convergence effect is possible when *per capita* aggregate income actually falls. An interpretation of this is that part of the convergence behavior arises independently of aggregate dynamics.

On balance, these results suggest that the convergence patterns of the US States may be best represented by a model lying in between the integrated and the autarchic models.

5.2 σ-convergence

Before we proceed to estimate the model expressed in terms of variances, equation (52), we need to take care of a problem in the data. When it comes to computing cross-section variances, the respective weights of regions matter. It would be misleading to count each region as one observation in the total population. To make this point clearer, consider the following thought experiment: split California into two states, each with half of California's population and income. The weight of the Californian growth behavior in the cross-section variance of relative incomes would suddenly be twice what it was before – while the economy has not changed a bit. Nor would a simple weighing by populations result in an accurate calculation of variances, because of the quadratic nature of variance. Appendix E details the procedure we propose to handle population weights. Given that the variance is basically a sum of squares, the trick is to weigh each state's relative income with the square root of its population. Figure 2 gives the variances adjusted for population over time.

< Insert figure 2. around here >

An eye examination of the graph does suggest some tendency for the cross-section variance to stabilize at a strictly positive value. But, without any further qualification, this is perfectly compatible with autarchic equation (50). If σ_u^2, the variance of the stochastic innovations of relative incomes, is significantly above zero, the autarchic model will display this sort of behavior[12].

As can be seen from the results of the estimation of (52) in Table1, the data on US states provide some indication that integration matters. Just like in the β-convergence analysis, the model applied to σ-convergence succeeds in reproducing the data quite well. In particular, the coefficient on $\frac{Y(t)}{Y(t-1)}\sigma^2_{t-1}$ is statistically significant. Ljung-Box tests performed on the first three lags do not reject the null hypothesis of no autocorrelation up to the specified lags. The same tests performed on the residuals of equation (51) decisively reject the same null hypotheses.

On the other hand, the values of the estimated structural parameters give more ambiguous information. These values are as follows (standard errors are shown between brackets): $\omega_1^* = 0.379$ [0.0721] and $x = -0.0415$ [0.027]. While $\omega_1^* = 0.379$ is plausible, the value $x = -0.0415$, like in our analysis of β-convergence, suggests that the convergence behavior of the US States may be best represented by some sort of cross of the two models.

Table 1: estimation of (52)

Variable	Coefficient	Std. Error	t-Statistic	p-Value
Intercept	0.000223	0.000541	0.412670	0.6812
σ^2_{t-1}	1.758222	0.144274	12.18669	0.0000
$\frac{y(t)}{y(t-1)}\sigma^2_{t-1}$	-0.790381	0.145032	-5.449707	0.0000
R-squared:	0.990571	Akaike info criterion:	-7.953222	
Adjusted R-squared:	0.990276	Schwarz criterion:	-7.854505	
Lag	Q'_{LB} statistic		p-Value	
1	0.3398		0.560	
2	1.6840		0.431	
3	4.0986		0.251	

6 Conclusion

This model provides a theoretical explanation for the deceleration of convergence among integrated, *i.e.* regional economies. When interactions are taken into account in a Ramsey model of economic growth, a link between aggregate growth and regional speed of convergence appears. Our characterization makes it possible to confront the model with empirical evidence. The results using US states data are mixed. On the one hand, integration appears to matter: there are indications that aggregate dynamics contribute to shaping the pattern of convergence among US states. On the other, part of conditional convergence seems not to be connected with the aggregate economy.

This analysis highlights the importance of interactions among economies for understanding patterns of convergence. We have shown that, in the

[12] The steady state value for σ_t^2 given by equation (50) is $\sigma^2_\infty = \sigma^2_u/(1-e^{-2\mu})$.

long run, the influence of integration may run contrary to the traditional notion of "equalizing exchange". Again, it is important to remember that in this intertemporal utility setting and without any sort of externality, the counterpart of the long-run persistence result is an improvement of dynamic utility for all regions.

Two promising directions for further research on integration and growth can be thought of. First, our empirical results suggest that less than perfect integration should be analyzed, for example through the introduction of non-tradables, adjustment costs or investment irreversibility, as in Vellutini [1997]. In the same direction, liquidity constraints as introduced by Huggett [1997] may potentially affect the persistence result. These extensions could obviously bring additional realism to the model, especially if it is applied to the world economy. Second, the result of long-run GNP persistence could also be modified by mechanisms of technology diffusion. This is clearly a very important extension of the model, as shown by the recent literature on the subject, for example Basu and Weil [1996] who explore the idea of gradual technology diffusion through a mechanisms of "appropriate technology", or Eeckhout and Jovanovic [1998] who show how non-trivial external effects give rise to inequality in productivity.

References

Altug, Sumru and Pamela Labadie, *Dynamic Choice and Asset Markets*, Academic Press, San Diego, 1994.

Barro, Robert and Xavier Sala-I-Martin, *Economic Growth*, McGraw Hill, New York, 1995.

Basu, Susanto and David Weil, *Appropriate Technology and Growth*, Working Paper 5865, NBER December 1996.

Ben-David, Dan, Convergence clubs and subsistence economies, *Journal of Development Economics*, 1998, (55), 155–171.

Bliss, Christopher, *Capital Mobility, Convergence Clubs and Long-Run Economic Growth*, Technical Report 1995. Paper presented at the 10th Economic Association Congress.

Caselli, Francesco and Jaume Ventura, A representative Consumer Theory of Distribution, *American Economic Review*, 2000. Forthcoming.

Eeckhout, Jan and Boyan Jovanovic, *Inequality*, Working Paper, Univ. Pompeu Fabra October 1998.

Fuente, Angel De La, *What kind of regional convergence?*, Mimeo, IAE, U. Autonoma de Barcelona 1998.

Huggett, Mark, The one-sector growth model with idiosyncratic shocks: Steady states and dynamics, *Journal of Monetary Economics*, 1997, (39), 385–403.

Kim, Sukkoo, *Economic Integration and Convergence: U.S. Regions, 1840-1997*, Working Paper 6335, NBER 1997.

Martin, Philippe, *Convergence et politiques régionales en Europe*, La lettre du CEPII 159, CEPII Juillet 1997.

Sherwood-Call, Caroline, The 1980's Divergence in State per Capita Incomes: What does it tell us?, *Economic Review*, 1996.

Vellutini, Charles, Mobilité des capitaux avec irréversibilité de l'investissement dans un modèle néoclassique à deux pays, *Revue d'économie politique*, Novembre 1997, *107* (6), 757–780.

Ventura, Jaume, Growth and Interdependence, *Quarterly Journal of Economics*, 1997, pp. 57–84.

Technical appendices

A The aggregate economy around its steady-state

Around the steady-state $(c^\star, k^\star)$, the aggregate economy dynamics is characterized by the Jacobian matrix:

$$J = \begin{bmatrix} 0 & c^\star \sigma^{-1} f''(k^\star) \\ -1 & f'(k^\star) - (\delta + n + x) \end{bmatrix}. \tag{53}$$

The characteristic polynomial associated to J is:

$$P(\mu) = \mu^2 - (\rho + \sigma x - (n + x))\mu + c^\star \sigma^{-1} f''(k^\star). \tag{54}$$

It is well known that $P(\mu)$ has two roots of opposite signs $\mu_1 > 0$ and $\mu_2 < 0$. Moreover, we know that $\mu_1 + \mu_2 = \mathrm{Tr}(J) = \rho + \sigma x - (n + x)$. $\mu = |\mu_2| > 0$ is the *speed of convergence* of the world economy toward its steady-state. As $\mu_1 \times \mu_2 = \mathrm{Det}(J) = c^\star \sigma^{-1} f''(k^\star)$, we obtain:

$$\mu(\mu + \rho + \sigma x - (n + x)) = -c^\star \sigma^{-1} f''(k^\star). \tag{55}$$

μ can be expressed as a function of the parameters of preferences and technology. Define ϵ and α as representing respectively the elasticity of substitution between factors of the technology and the share of capital income in total income[13]:

$$\epsilon \equiv -\frac{f'(k)(f(k) + \phi - kf'(k))}{k(f(k) + \phi) f''(k)}, \quad \alpha \equiv \frac{kf'(k)}{f(k) + \phi}. \tag{56}$$

Let $r^\star = \rho + \sigma x$ be the (net) interest rate in he steady-state. The aggregate speed of adjustment is then given by:

$$\mu = \frac{1}{2}\left[-(r^\star - (n + x)) + \sqrt{(r^\star - (n + x))^2 + \frac{4}{\epsilon\sigma}(r^\star + \delta)(1 - \alpha)\frac{c^\star}{k^\star}} \right]. \tag{57}$$

Moreover, with our definitions, the ratio $c^\star / k^\star$ satisfies:

$$\frac{c^\star}{k^\star} = \frac{r^\star + \delta}{\alpha} - (\delta + n + x).$$

It is useful to study the local properties of the policy rule $c(k)$ associated to the competitive aggregate dynamics. We eliminate time in the differential system (12,13) so that:

$$c'(k) = \frac{\sigma^{-1}(f'(k) - \delta - \rho)c(k) - xc(k)}{f(k) + \phi - c(k) - (\delta + x + n)k}. \tag{58}$$

[13] Including ϕ in these expressions facilitates the algebrical manipulations below.

Since 0/0 is indeterminate, (58) is uninformative on the value of $c'(k^\star)$. However, by applying L'Hôpital's rule, it is easy to show that $c'(k^\star)$ solves $P(\mu) = 0$. Consequently, there are two candidates for $c'(k^\star)$: the two eigenvalues μ_1 and μ_2, corresponding respectively to the stable and unstable arms of the phase diagram. It is easy to see that the stable arm has a positive slope in the (k, c) plan so that:

$$c'(k^\star) = \mu_1 = \mu + \rho + \sigma x - (n + x), \tag{59}$$

is the solution of interest.

B Properties of $\lambda_1(k, k^\star)$ and $\psi(k)$ around the steady-state

$\lambda_1(t, t')$ can alternatively be expressed in terms of the aggregate capital stocks $k(t)$ and $k(t')$. Substituting (21) into (23) provides a convenient expression:

$$\lambda_1(k, k') = \frac{a(k')/a(k)}{k'/k}. \tag{60}$$

¿From this, we have

$$\eta_1^\star = 1 - (k^\star/a^\star)a'(k^\star). \tag{61}$$

We need to characterize the policy rule $a(k)$ near the steady-state $k^\star$. Knowing that $\dot{a} = (r - n - x)a - c$ and using the expression of the aggregate dynamic system, we have:

$$a'(k) = \frac{(f'(k) - \delta - n - x)a(k) - c(k)}{f(k) + \phi - c(k) - (\delta + n + x)k}.$$

Again, by applying L'Hôpital's rule to this ratio of functions, we see that $a'(k^\star)$ satisfies:

$$-c'(k^\star)a'(k^\star) - a^\star f''(k^\star) + c'(k^\star) = 0,$$

so that:

$$a'(k^\star) = 1 - a^\star \frac{f''(k^\star)}{c'(k^\star)}.$$

Further, knowing that $c^\star/a^\star = \rho + \sigma x - (n + x) = \nu^\star$, we get:

$$a'(k^\star) = 1 + \frac{\sigma\mu}{\rho + \sigma x - (n + x)}. \tag{62}$$

¿From this last expression, we conclude that:

$$\eta_1^\star = 1 - \frac{k^\star}{c^\star}\left[\rho + \sigma(x + \mu) - (n + x)\right]. \tag{63}$$

The instantaneous speed of conditional convergence $\psi(k)$ relates to the function $\lambda_1(k, k')$ through the expression:

$$\frac{\partial \log \lambda_1(k, k')}{\partial k} = \frac{\psi(k)}{f(k) + \phi - c(k) - (n + \delta + x)k}.$$

At $k = k^\star$, we apply the L'Hôpital's rule once more, using (59) to show:

$$\left(\frac{\partial \log \lambda_1(k, k^\star)}{\partial k}\right)_{k=k^\star} = \frac{\eta_1(k^\star)}{k^\star} = \frac{\psi'(k^\star)}{-\mu}.$$

Finally:

$$\psi'(k^\star) = -\mu \frac{\eta_1(k^\star)}{k^\star}. \tag{64}$$

C Properties of $\lambda_3(k, k^\star)$ and $\psi_2(k)$ around the steady-state

Remember that:

$$\lambda_3(t, t') = \int_t^{t'} \psi_2(s)\lambda_1(s, t')ds.$$

Making the change of variable $t \to k$, we obtain:

$$\lambda_3(k, k') = \int_k^{k'} \frac{\psi_2(u)\lambda_1(u, k')}{f(u) + \phi - c(u) - (\delta + n + x)u} du.$$

By differentiating this equality with respect to k:

$$\lambda_3'(k) \equiv \frac{\partial \lambda_3(k, k^\star)}{\partial k} = \frac{-\psi_2(k)\lambda_1(k, k^\star)}{f(k) + \phi - c(k) - (\delta + n + x)k}.$$

Knowing that $\lambda_1(k^\star, k^\star) = 1$ and $\psi_2(k^\star) = 0$, we apply the L'Hôpital rule to show that:

$$\lambda_3'(k^\star) = \frac{\psi_2'(k^\star)}{\mu}. \tag{65}$$

We now turn to the evaluation of the derivative $\psi_2'(k^\star)$. The function $\psi_2(k)$ satisfies $\psi_2(k) = \phi[1 - \nu(k)\widetilde{1}(k)]/k$. Differentiating this expression with respect to k yields:

$$\psi_2'(k) = -\phi \frac{k\left[\nu'(k)\widetilde{1}(k) + \nu(k)\widetilde{1}'(k)\right] + (1 - \nu(k)\widetilde{1}(k))}{k^2}.$$

This expression has to be evaluated at $k = k^\star$. First, we need $\nu'(k^\star)$. By definition, $\nu(k) = c(k)/a(k)$ so that:

$$\frac{\nu'(k)}{\nu(k)} = \frac{c'(k)}{c(k)} - \frac{a'(k)}{a(k)}.$$

It is easy to check that $\nu^\star = \nu(k^\star) = \rho + \sigma x - (x+n)$. Using (59) and (62), one gets:

$$\nu'(k^\star) = -\frac{1}{a^\star}\mu(1-\sigma), \tag{66}$$

Second, we need $\widetilde{1}'(k^\star)$. Using the expression:

$$\widetilde{1}(t) = \int_t^\infty e^{-\int_t^\tau (r(s)-n-x)ds} d\tau,$$

By making the change of variable $t \to k$:

$$\widetilde{1}'(k) = \frac{(f'(k)-n-x-\delta)\widetilde{1}(k)-1}{f(k)+\phi-c(k)-(\delta+x+n)k}.$$

For $k = k^\star$, we apply the L'Hôpital rule so that:

$$\widetilde{1}'(k^\star) = \frac{f''(k^\star)\widetilde{1}(k^\star)+\widetilde{1}'(k^\star)(f'(k^\star)-\delta-x-n)}{f'(k^\star)-c'(k^\star)-(\delta+x+n)k}.$$

Knowing that $\widetilde{1}(k^\star) = (\rho+\sigma x-(x+n))^{-1}$ and that the denominator is equal to μ, one can see that:

$$\widetilde{1}'(k^\star) = \frac{\sigma\mu}{c^\star(\rho+\sigma x-(n+x))}. \tag{67}$$

Using equations (65) (66) and (67), we see that:

$$\psi_2'(k^\star) = -\frac{\mu\phi}{k^\star c^\star}. \tag{68}$$

Knowing that $\lambda_3'(k^\star) = \psi_2'(k^\star)/\mu$ and $\eta_3^\star = k^\star\lambda_3'(k^\star)$, we finally obtain:

$$\eta_3^\star = -\phi/c^\star. \tag{69}$$

D Properties of $\gamma_i(k,k')$

Let $\gamma_1(k,k')$, $\gamma_2(k,k')$ and $\gamma_3(k,k')$ be functions of capital stock, corresponding to the coefficients $\gamma_i, i=1,2,3$ of equation (28). The objective of this appendix is to find closed-form expressions for the semi-elasticities $\omega_i^\star, i=1,2,3$ of these functions around the aggregate steady-state.

From (29) we have:

$$\omega_1^\star = \eta_1^\star - \frac{k^\star\alpha'(k^\star)}{\alpha^\star}, \tag{70}$$

with $\alpha(k) = kf'(k)/(f(k)+\phi)$. The elasticity of α around the steady-state can be computed using definitions (56). We obtain:

$$\frac{k^\star\alpha'(k^\star)}{\alpha^\star} = (1-\alpha^\star)(1-1/\epsilon^\star).$$

This equation, together with (63), implies:

$$\omega_1^\star = 1 - \frac{k^\star}{c^\star}\left[\rho + \sigma(x+\mu) - (n+x)\right] - (1-\alpha^\star)(1-(1/\epsilon^\star)). \tag{71}$$

Note that when $\epsilon^\star = 1$, which corresponds to the Cobb-Douglas case with $\phi = 0$, $\omega_1^\star$ is equal to $\eta_1^\star$.

Despite its simplicity, an inconvenient of equation (71) is that it is not homogeneous in the parameters describing the aggregate fundamentals. For instance, μ depends on $\epsilon^\star$. One way to circumvent this difficulty is to use (55) to show that:

$$\frac{k^\star \alpha'(k^\star)}{\alpha^\star} = 1 - \frac{k^\star(\rho + \sigma x + \delta)}{c^\star + (n+x+\delta)k^\star} - \frac{k^\star}{c^\star}\frac{\sigma\mu(\mu + \rho + \sigma x - (n+x))}{\rho + \sigma x + \delta}.$$

By combining this expression with (63), we obtain a homogenous expression for $\omega_1^\star$. However, this expression is complex. It simplifies in the case $\delta = n = x = 0$:

$$\omega_1^\star = \frac{k^\star}{c^\star}\frac{\sigma\mu^2}{\rho}, \tag{72}$$

which is always positive. This means that in the absence of depreciation, demographic growth and technological progress there is always conditional convergence in gross income around the steady state.

The expression of $\omega_3^\star$ can be obtained from (31). It is easy to show that:

$$\omega_3^\star = \alpha^\star \frac{\phi}{c^\star} - (1-\alpha^\star-\beta^\star)\left[\omega_1^\star - \alpha^\star\right], \tag{73}$$

where $\beta^\star = w^\star/y^\star$ is the share of wages in total income.

Knowing that the $\gamma_i' s$ sum up to unity, we have $\omega_2^\star = -\omega_1^\star - \omega_3^\star$. Hence:

$$\omega_2^\star = -\alpha^\star \frac{\phi}{c^\star} + (\alpha^\star + \beta^\star)\omega_1^\star - \alpha^\star(1-\alpha^\star-\beta^\star). \tag{74}$$

For $\phi = 0$, this reduces to $\omega_2^\star = -\omega_1^\star$.

E Adjusting for population size

The model assumes that all regions have identical populations, although this is grossly untrue in reality. When computing the cross-section variance of relative *per capita* incomes, we need to adjust for population size. The following explains how to do this with the data at hand.

Suppose that the country consists of P regions indexed by $i = 1, ..., P$, each being populated with a single inhabitant. Define Y as the country total income, $y(i)$ as $i's$ *per capita* income (also equal to its total income), $Y(i) \equiv Y(i)/Y$ as $i's$ relative income and $y_R(i) \equiv Y(i)/(Y/P)$ its relative *per capita* income. Symmetric notation is used for regions. Suppose that there exists a partition of the country population into n regions indexed by $j = 1, ..., n$ such that

$$\text{for all } j \text{ and for all } i \in I_j, Y(i) = Y(j)/P(j), \tag{75}$$

where I_j is the set of $P(j)$ regions belonging to region j and $Y(i)$ is $i's$ total income.

Let us compute the variance of the $y_R(i)$, which is the variable we are primarily interested in:

$$V(y_R(i)) = \frac{1}{P}\sum_{i=0}^{P} y_R^2(i) - \left(\frac{1}{P}\sum_{i=0}^{P} y_R(i)\right)^2 = \frac{1}{P}\sum_{i=0}^{P} y_R^2(i) - 1 \tag{76}$$

We leave it to the reader to verify that the variance of *per capita* income across regions, $V(y_R(j))$, will usually give a very different result.

Let us now compute the quantity $\sum_{j=1}^{n}\left(\frac{Y(j)}{Y}P^{-1/2}(j)\right)^2$, which will prove useful to solve our problem. First note that (75) implies: $Y^2(j) = P(j)\sum_{i\in I_j} Y^2(i)$. Substituting:

$$\sum_{j=1}^{n}\left(\frac{Y(j)}{Y}P^{-1/2}(j)\right)^2 = \frac{1}{Y^2}\sum_{j=1}^{n}\sum_{i\in I_j} Y^2(i) = \frac{1}{Y^2}\sum_{i=0}^{P} Y^2(i) = \frac{1}{P^2}\sum_{i=0}^{P} y_R^2(i)$$

Finally:

$$V(y_R(i)) = P\sum_{j=1}^{n}\left(\frac{Y(j)}{Y}P^{-1/2}(j)\right)^2 - 1 \tag{77}$$

Variances of relative *per capita* income have been calculated using this formula.

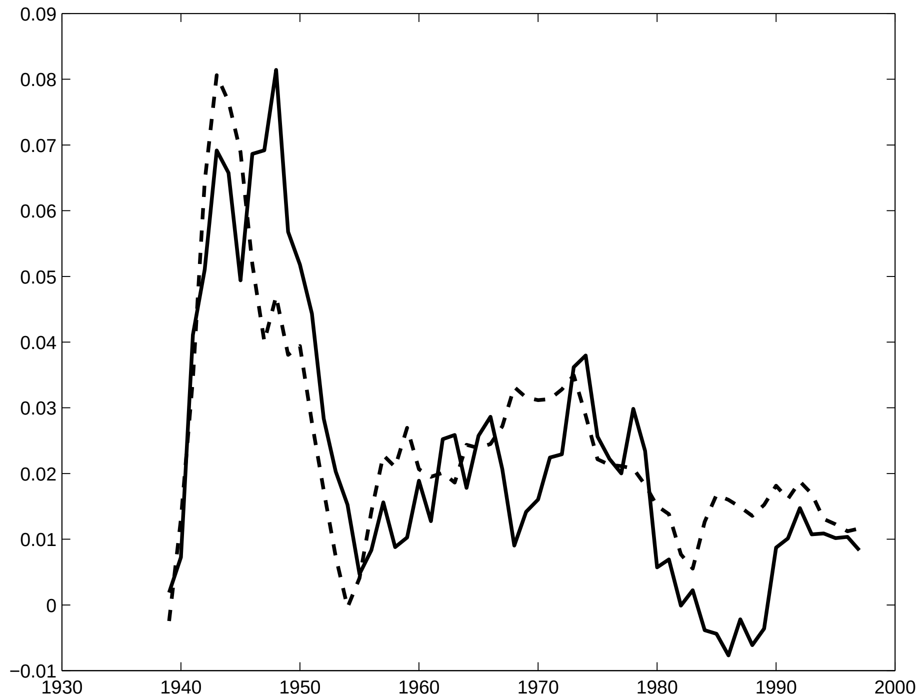

Figure 1: Aggregate annual growth rate (over decades) and associated speed of regional convergence for US states (solid line), from 1939 to 1997.

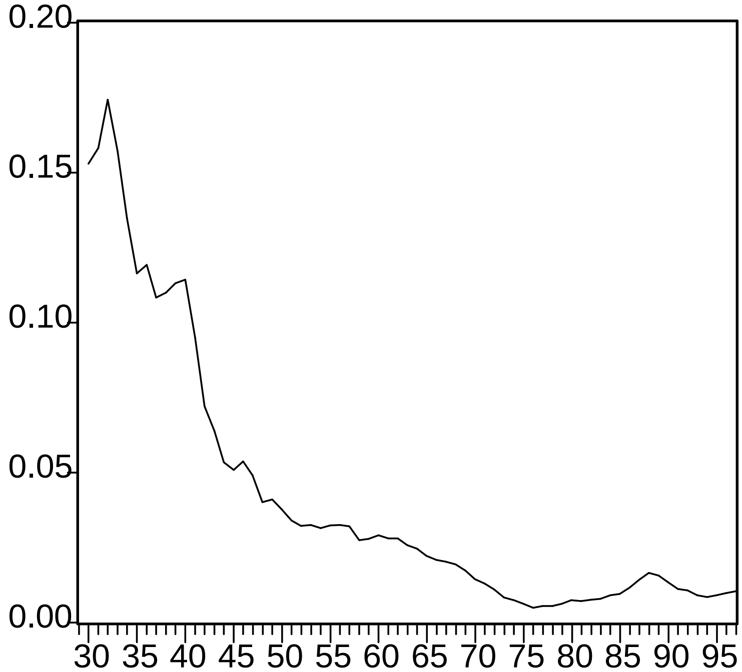

Figure 2: Evolution of cross section variance among US states

Method: GLS (Cross Section Weights)
White Heteroskedasticity-Consistent Standard Errors & Covariance
Sample: 1930 1997

Coefficient	estimated	t-Statistic	Coefficient	estimated	t-Statistic
a1	0.597619	5.440780			
a2	-0.643826	-5.902049			
a3(ALA)	0.480879	7.581268	a4(ALA)	-0.447988	-7.054486
a3(ARI)	0.632111	5.609858	a4(ARI)	-0.590856	-5.218866
a3(ARK)	0.435637	4.664617	a4(ARK)	-0.403794	-4.182896
a3(CAL)	0.770569	5.037655	a4(CAL)	-0.719198	-4.662324
a3(COL)	0.519441	3.917197	a4(COL)	-0.467334	-3.505049
a3(CON)	0.846989	5.112692	a4(CON)	-0.785355	-4.705064
a3(DEL)	0.836341	3.833641	a4(DEL)	-0.783921	-3.553896
a3(DCO)	-0.061751	-0.271465	a4(DCO)	0.141642	0.612222
a3(FLO)	0.573614	5.674636	a4(FLO)	-0.527241	-5.167499
a3(GEO)	0.443211	6.627201	a4(GEO)	-0.402323	-5.986030
a3(IDA)	0.855205	1.513705	a4(IDA)	-0.815726	-1.446400
a3(ILL)	0.848975	6.135464	a4(ILL)	-0.800868	-5.736485
a3(IND)	0.899161	8.794022	a4(IND)	-0.860037	-8.354711
a3(IOW)	0.720656	1.717696	a4(IOW)	-0.675108	-1.619606
a3(KAN)	0.756320	4.566151	a4(KAN)	-0.714863	-4.279037
a3(KEN)	0.463233	6.291491	a4(KEN)	-0.428204	-5.749152
a3(LOU)	0.421534	5.305934	a4(LOU)	-0.382810	-4.767320
a3(MAI)	0.530411	4.468744	a4(MAI)	-0.489463	-4.084849
a3(MAR)	0.569608	4.341993	a4(MAR)	-0.513922	-3.886095
a3(MAS)	0.495153	3.523578	a4(MAS)	-0.437127	-3.084803
a3(MIC)	1.125598	7.962658	a4(MIC)	-1.086419	-7.609309
a3(MIN)	0.494765	4.338739	a4(MIN)	-0.444537	-3.868579
a3(MIS)	0.465977	4.943786	a4(MIS)	-0.438873	-4.563622
a3(MI)	0.536703	5.351204	a4(MI)	-0.491624	-4.863741
a3(MON)	0.513169	1.646378	a4(MON)	-0.468880	-1.468911
a3(NEB)	0.679113	1.770262	a4(NEB)	-0.634911	-1.674038
a3(NEV)	0.695879	1.262192	a4(NEV)	-0.637716	-1.144664
a3(NHA)	0.413882	3.675984	a4(NHA)	-0.361636	-3.186787
a3(NJ)	0.713879	4.882395	a4(NJ)	-0.655028	-4.442119
a3(NEM)	0.504533	6.705502	a4(NEM)	-0.467290	-6.141164
a3(NYO)	0.633524	4.037920	a4(NYO)	-0.577869	-3.649714
a3(NCA)	0.413758	5.807360	a4(NCA)	-0.372635	-5.166459
a3(NDA)	0.509475	0.617782	a4(NDA)	-0.460926	-0.541863
a3(OHI)	0.922230	7.708824	a4(OHI)	-0.881892	-7.315065
a3(OKL)	0.494265	4.443157	a4(OKL)	-0.455509	-4.059441
a3(ORE)	0.806509	5.854084	a4(ORE)	-0.762698	-5.496431
a3(PEN)	0.676252	5.716457	a4(PEN)	-0.630460	-5.287191
a3(RHO)	0.454663	3.044060	a4(RHO)	-0.403836	-2.680050
a3(SCA)	0.374609	6.162632	a4(SCA)	-0.336340	-5.492981
a3(SDA)	0.728954	1.270689	a4(SDA)	-0.688271	-1.185450
a3(TEN)	0.525486	7.577217	a4(TEN)	-0.489085	-7.000460
a3(TEX)	0.499744	5.759757	a4(TEX)	-0.455048	-5.216213
a3(UTA)	0.548159	2.801829	a4(UTA)	-0.508277	-2.561244
a3(VER)	0.496641	5.039425	a4(VER)	-0.455141	-4.582686
a3(VIR)	0.411776	4.108234	a4(VIR)	-0.361119	-3.560310
a3(WAS)	0.858042	6.423250	a4(WAS)	-0.810976	-6.020307
a3(WVI)	0.435207	5.151321	a4(WVI)	-0.398825	-4.665768
a3(WIS)	0.658965	6.256072	a4(WIS)	-0.615231	-5.798839
a3(WYO)	0.542769	4.097084	a4(WYO)	-0.494519	-3.700324

Weighted Statistics			
R-squared	0.186944	Mean dependent var	0.002811
Adjusted R-squared	0.162039	S.D. dependent var	0.045047
S.E. of regression	0.041236	Sum squared resid	5.495713
Log likelihood	11151.98	Durbin-Watson stat	2.096958

Unweighted Statistics			
R-squared	0.151211	Mean dependent var	0.002783
Adjusted R-squared	0.125212	S.D. dependent var	0.044437
S.E. of regression	0.041562	Sum squared resid	5.582945
Durbin-Watson stat	2.556498		

Table 1: This Table reports estimates associated with the following equation $g_R(\theta,t) = a_1 + a_2\frac{y(t)}{y(t-1)} + a_3(\theta)\frac{y(t)}{y(\theta,t-1)} + a_4(\theta)\frac{1}{y_R(\theta,t-1)} + u(\theta,t)$

Estimation Method: Iterative Weighted Least Squares
Sample: 1929 1997
Sequential weighting matrix & coefficient iteration

Coefficient	Estimated	Std. Error	t-Statistic
omega	0.628980	0.053699	11.71318
x	-0.049842	0.006502	-7.665259
b(1)	-0.627522	0.0535291	-11.72300
b(2)	0.659912	0.054494	12.11031
theta(ALA)	0.745964	0.035904	20.77634
theta(ARI)	0.939660	0.054036	17.38947
theta(ARK)	0.696278	0.049065	14.19091
theta(CAL)	1.133195	0.048407	23.40986
theta(COL)	0.973826	0.068613	14.19303
theta(CON)	1.326624	0.067555	19.63770
theta(DEL)	1.196326	0.105938	11.29268
theta(DCO)	0.955052	0.128860	7.411556
theta(FLO)	0.956919	0.043825	21.83508
theta(GEO)	0.789708	0.035235	22.41231
theta(IDA)	1.114787	0.141607	7.872404
theta(ILL)	1.163701	0.048402	24.04221
theta(IND)	1.126427	0.050746	22.19747
theta(IOW)	1.075787	0.147798	7.278771
theta(KAN)	1.058404	0.077064	13.73414
theta(KEN)	0.747560	0.035162	21.26027
theta(LOU)	0.749997	0.044007	17.04275
theta(MAI)	0.859514	0.053845	15.96271
theta(MAR)	1.055838	0.047313	22.31596
theta(MAS)	1.031795	0.052770	19.55276
theta(MIC)	1.278563	0.076761	16.65636
theta(MIN)	0.934110	0.062768	14.88187
theta(MIS)	0.680818	0.052642	12.93287
theta(MI)	0.907625	0.031133	29.15323
theta(MON)	0.882066	0.103000	8.563761
theta(NEB)	1.025329	0.127441	8.045517
theta(NEV)	1.180946	0.162857	7.251444
theta(NHA)	0.898854	0.047512	18.91846
theta(NJ)	1.196127	0.046304	25.83219
theta(NEM)	0.802558	0.035964	22.31568
theta(NYO)	1.092052	0.046771	23.34914
theta(NCA)	0.767834	0.044970	17.07440
theta(NDA)	0.932230	0.170901	5.454803
theta(OHI)	1.131824	0.043125	26.24505
theta(OKL)	0.808585	0.058339	13.86013
theta(ORE)	1.096303	0.063764	17.19308
theta(PEN)	1.016283	0.029109	34.91335
theta(RHO)	0.919784	0.063503	14.48419
theta(SCA)	0.704040	0.042756	16.46641
theta(SDA)	1.059616	0.152036	6.969522
theta(TEN)	0.817900	0.035764	22.86950
theta(TEX)	0.875644	0.043313	20.21687
theta(UTA)	0.862385	0.078632	10.96737
theta(VER)	0.838074	0.037344	22.44230
theta(VIR)	0.870980	0.052967	16.44373
theta(WAS)	1.171933	0.061373	19.09537
theta(WVI)	0.736010	0.039444	18.65950
theta(WIS)	0.985408	0.035972	27.39378
theta(WYO)	0.945191	0.069298	13.63957

Table 2: This table reports Estimates associated with the following equation $g_R(\theta,t) = \left(b_1 + b_2 \frac{y(t)}{y(t-1)}\right)\left[\frac{\theta}{y_R(\theta,t-1)} - 1\right] + u(\theta,t)$, for $t = 2, ..., T$