

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

<http://www.archive.org/details/representativeco00case>

HB31
.M415
no. 96-11



**working paper
department
of economics**

A REPRESENTATIVE CONSUMER THEORY OF DISTRIBUTION

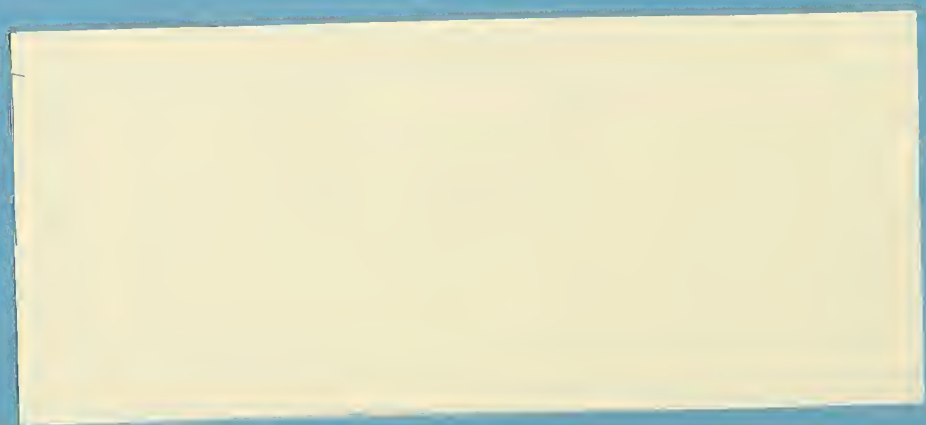
Francesco Caselli
Jaume Ventura

96-11

April. 1996

**massachusetts
institute of
technology**

**50 memorial drive
cambridge, mass. 02139**



A REPRESENTATIVE CONSUMER THEORY OF DISTRIBUTION

**Francesco Caselli
Jaume Ventura**

96-11

April. 1996

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JUL 18 1996

LIBRARIES

A Representative Consumer Theory of Distribution

Francesco Caselli¹
Harvard University

Jaume Ventura¹
M.I.T.

April, 1996

¹We are grateful to Daron Acemoglu, Robert Barro, Alberto Bisin, Gary Chamberlain, Bill Easterly, Gerardo Esquivel, Aart Kraay, David Laibson, John Leahy, Greg Mankiw, Jonathan Morduch, Andrew Newman, Steve Pischke, Antonio Rangel and Eric Thornbecke for their useful comments. Errors are ours. Further comments are welcome. Please write or email to: (Caselli) Harvard University, Department of Economics, Littauer Center, Cambridge, MA 02138, fcaselli@kuznets.fas.harvard.edu; or (Ventura) M.I.T., Department of Economics, 50 Memorial Drive, Cambridge, MA 02139, jaume@mit.edu.

Abstract

This paper shows that growth models featuring the representative consumer (RC) assumption can generate rich dynamics for the cross-sections of consumption, wealth and income. We consider a class of growth models with three sources of consumer heterogeneity: initial wealth, non-acquired skills and taste for consumption-smoothing. Despite this heterogeneity, the models in this class admit a RC and, consequently, exhibit aggregate dynamics that are indistinguishable from those of the standard homogeneous-consumer models. One could therefore interpret our research as an attempt to make explicit the distributive dynamics that underlie the popular RC models.

We examine the behavior of consumers' relative consumption, wealth and income, and derive cross-sectional equations that show how these quantities (at any date) are related to both average variables and the consumer's individual characteristics. Using these equations, we study the distributive dynamics of the Ramsey and the linear growth models. While the former is consistent with the existence of alternating periods of conditional convergence and divergence in a cross-section of individual incomes, the latter is not. Using data from the Panel Study of Income Dynamics (PSID), we find evidence of conditional convergence in the 1970s, and divergence thereafter. This finding is therefore inconsistent with the linear growth model, but not with some versions of the Ramsey model.

1 Introduction

The widespread use of the representative consumer (RC) assumption in the theory of economic growth has lead many to believe this line of research has little to say about distributive issues. This paper argues that quite the opposite is true. We consider a class of growth models with three sources of consumer heterogeneity: initial wealth, non-acquired skills and taste for consumption- smoothing. Despite this heterogeneity, the models in this class admit a RC and, consequently, exhibit aggregate dynamics that are indistinguishable from those of the standard homogeneous-consumer models. Yet this does not mean that the models we study do not have interesting distributive implications. On the contrary, we find that they generate rich dynamics for the cross-sections of consumption, wealth and income. One could therefore interpret our research as an attempt to make explicit the distributive dynamics that underlie the popular RC models.

The RC is a fictional consumer whose utility maximization problem when facing aggregate resource constraints generates the economy's aggregate demand functions. The RC assumption does not necessarily rule out consumer heterogeneity, but only requires that potential sources of consumer heterogeneity have sufficient structure to ensure that the sum of all consumers behaves *as if* it were a single consumer. For instance, if we are not willing to restrict the set of admissible wealth distributions, a necessary and sufficient condition for a model to admit a RC is that all consumers' preferences can be represented by indirect utility functions of the Gorman form (of which special cases are homothetic and quasilinear utility functions).¹ If we are willing to impose enough restrictions on the set of admissible wealth distributions, it is then possible to enlarge the class of preferences that lead to the RC property.

It is difficult to overemphasize the analytical convenience of the RC assumption. In economies with heterogeneous agents the behavior of average quantities depends, in general, on how these averages are distributed across consumers. The RC assumption restricts the behavior of average quantities to depend exclusively on these same averages, and not on their distribution.

¹Let W_j be the wealth of consumer j , and let p be a given price vector. The Gorman indirect utility function is $V(p, W_j) = \alpha_j(p) + \beta(p) \cdot W_j$ (note that all consumers must have the same wealth coefficient). See Mas-Colell, Whinston and Green (1995, chapter 4) for a review of aggregation results, including a discussion of the concept of a representative consumer.

Since keeping track of averages is always simpler than following distribution functions, this first property of RC models constitutes a major source of analytical tractability. Moreover, since disaggregated data sets are relatively scarce, it is convenient to work with theories that deliver testable implications in terms of average variables and not on their whole distribution. But this is not all. We know that, in general, average demand functions can take basically any form (Sonnenschein-Mantel-Debreu theorem). In RC models average demands are generated as a solution of an individual consumer problem and, as a result, they satisfy all the properties of individual demand functions. In applied work these properties are extremely useful, since they allow us to derive testable implications from theoretical models.

Indeed, since much of existing growth theory is based on the RC assumption, it is possible to understand this paper as an attempt to expand the range of applications of existing growth models to include, in addition to the usual predictions regarding average quantities, also predictions regarding how these quantities are distributed in a cross-section of consumers. Our paper is therefore closely related to Stiglitz (1969) and Chatterjee (1994). These papers study the evolution of the distribution of wealth in the Solow and in the Ramsey model, respectively. Our work differs from theirs in three dimensions. First, we introduce additional sources of consumer heterogeneity. Second, our results are more general, since they apply to any RC growth model. Finally and most important, we derive equations that explicitly relate the evolution of the cross-sections of consumption, income and wealth to the evolution of average quantities and prices and the individual characteristics of the consumers.

We study a class of RC growth models with three sources of consumer heterogeneity: initial wealth, non-acquired skills and taste for consumption-smoothing. In particular, we examine the behavior of consumers' relative consumption, wealth and income, i.e. the ratio of a consumer's consumption, wealth and income to the economy's corresponding average. We derive cross-sectional equations that show how these quantities (at any date) are related to both average variables and the consumer's individual characteristics. Since average variables are independent of the distribution of individual characteristics by virtue of the RC assumption, one can use a three step algorithm to extract empirical implications for the cross-sections of consumption, wealth and income: (i) Solve one's preferred RC model and characterize the equilibrium time-path of average quantities and prices; (ii) Make assumptions

regarding the distribution of individual characteristics (i.e. initial wealth, skills and tastes); (iii) With the information obtained in steps (i) and (ii) at hand, use the appropriate cross-sectional equations to obtain testable restrictions regarding the evolution of any initial cross-section of consumptions, wealth and/or real incomes.

We illustrate the use of this algorithm to study distributive dynamics in two popular RC growth models. In our first example we show that, under a standard representation of technology, the Ramsey model generates a Kuznets curve, that is, an inverted-U shaped relationship between inequality measures and the level of development of a country. In the second example, we completely characterize the dynamics of distribution in a linear growth model. This model is extremely simple and permits closed-form solutions that help us develop intuition about the forces that determine the evolution of a wealth distribution profile. Both the Ramsey and linear models are consistent with the existence of conditional convergence and divergence in a cross-section of individual incomes. One empirical feature that distinguishes these models is the possibility of reversals in the patterns of convergence. While the Ramsey model can account for alternating periods of convergence and divergence, the linear model cannot.

Our theoretical results allow us to formulate, for a cross-section of individual incomes, an empirical specification that is analogous to the cross-country or cross-regional “growth regressions” that have been extensively studied in the last decade.² We use data from the Panel Study of Income Dynamics (PSID) to characterize the dynamic behavior of a cross-section of family incomes in the 1970’s and 1980’s. We find evidence of conditional convergence in the former decade, and divergence in the latter. Specifically, our results suggest that in the 1970s the income distribution was converging towards the distribution of skills and/or tastes, while in the 1980s it diverged from the distribution of skills and/or tastes and initial income differences were accentuated. As already mentioned, this reversal in the patterns of convergence is inconsistent with the linear growth model, but not with the Ramsey model.

That the RC model can be used to study many distributive issues does not mean that all the important questions concerning the distribution of

²Examples of cross-country studies are Baumol (1986), De Long (1988), Barro (1991), Mankiw, Romer and Weil (1992), Knight, Loayza and Villanueva (1993) Barro and Sala-i-Martin (1995) and Caselli, Esquivel and Lefort (1995). Examples of cross-regional studies are Barro and Sala-i-Martin (1991, 1992, 1995).

income can be addressed within this framework. For example, there has recently been substantial interest in the question of how, as a result of politico-institutional factors³ and labor⁴ and/or financial⁵ market imperfections, alternative initial income distribution profiles map into subsequent income levels and growth rates. This literature is summarized in Bénabou (1996). By the very nature of the problem these papers analyze (i.e. the effects of lump-sum redistribution policies on income levels and growth rates), the use of RC models is ruled out from the start. The reason is that the RC property implies that lump-sum redistributions do not affect the behavior of average quantities.⁶

The paper is organized as follows: Section two develops the basic theory. Section three applies the theory to the Ramsey and linear growth models. Section four presents some empirical results. Section five concludes.

³See, for instance, Perotti (1993), Alesina and Perotti (1994), Alesina and Rodrik (1994) and Persson and Tabellini (1994).

⁴See, for instance, Acemoglu (1995).

⁵See, for instance, Loury (1981), Becker and Tomes (1986), Banerjee and Newman (1991 and 1993), Galor and Zeira (1993), Behrman, Pollack and Taubman (1995) and Piketty (1996).

⁶This is not the same as to say that in the RC models any redistribution policy does not affect aggregate variables. This second statement is false.

2 Heterogeneity in the RC Model

This section presents a model with heterogeneous consumers, in which the RC property emerges as an aggregation result. We characterize the behavior of the cross-sections of consumption, wealth and income, as a function of average quantities and prices and the distributions of individual characteristics.

2.1 A RC Model

Consider a one-good economy with heterogeneous consumers, indexed by $j = 1, 2, \dots, J$. Time is continuous and there is no uncertainty. Let $c_j(t) \in \mathbb{R}_+$ denote the consumption of agent j at date t , and $\tilde{c}(t) = (c_1(t), \dots, c_J(t)) \in \mathbb{R}_+^J$ be the economy's vector of consumptions. To simplify notation, we omit the time indexes when this is not confusing. Consumers have preferences that can be represented by a Stone-Geary utility function:

$$\int_0^\infty \frac{(c_j + \beta_j)^{1-\theta} - 1}{1-\theta} \cdot e^{-\rho t} \cdot dt \quad (\theta > 0, \rho > 0) \quad (1)$$

where $\beta_j \in \mathbb{R}$ and $\tilde{\beta} = (\beta_1, \dots, \beta_J) \in \mathbb{R}^J$. We do not impose any restriction on the vector $\tilde{\beta}$. As a result, a first source of consumer heterogeneity is variation in tastes. As we shall see, the Stone-Geary utility function is a specially convenient parametrization of preferences, since it gives rise to consumption demands that are linear in wealth. Despite this simplicity, we are allowing the elasticity of intertemporal substitution to vary both in a cross-section of consumers and in a time series of a single consumer. This elasticity is given by $\frac{c_j + \beta_j}{\theta \cdot c_j}$. If $\beta_j < 0$ ($\beta_j > 0$), then the intertemporal elasticity of substitution is increasing (decreasing) in the level of consumption c_j , and approaches θ^{-1} as $c_j \rightarrow \infty$. If $\beta_j = 0$, the intertemporal elasticity of substitution is constant and equal to θ^{-1} . Also, for a given value of consumption, the intertemporal elasticity of substitution increases with β_j .

Consumers derive income from their ownership of assets and the labor they provide. A second source of consumer heterogeneity is given by their initial asset holdings. Let $a_j(t) \in \mathbb{R}$ be the stock of assets of consumer j at date t , and $\tilde{a}(t) = (a_1(t), \dots, a_J(t)) \in \mathbb{R}^J$ be the economy's vector (distribution) of asset holdings. Throughout the paper we interpret the consumer's

stock of assets in a broad manner, so as to incorporate both financial wealth and acquired skills or education. A third source of consumer heterogeneity is variation in labor productivity or non-acquired skills. We define $\pi_j \in \mathbb{R}_+$ as the labor productivity of agent j , and $\tilde{\pi} = (\pi_1, \dots, \pi_J) \in \mathbb{R}_+^J$ as the economy's vector (distribution) of labor productivity parameters.

Let $r(t) \in \mathbb{R}_+$ and $w(t) \in \mathbb{R}_+$ denote the rate of return on assets and the wage rate measured in efficiency units, at date t . We assume that r and w are smooth functions of time with, at most, a finite number of jumps. Using this notation, we can write consumer j 's flow budget constraint as follows:

$$\dot{a}_j = r \cdot a_j + w \cdot \pi_j - c_j \quad (2)$$

Since we interpret assets holdings in a broad manner, it would be inappropriate to equate the quantity $w \cdot \pi_j$ with the consumer's labor income. The latter also contains the return to human capital which is included in the quantity $r \cdot a_j$.

The consumer's problem of maximizing the objective function (1) subject to the flow budget constraint (2) has solution:⁷

$$c_j = m_1 \cdot a_j + m_2 \cdot \pi_j + m_3 \cdot \beta_j \quad (3)$$

where m_1 , m_2 and m_3 are defined as follows:

$$m_1 = \left(\int_t^\infty e^{\tau} \int_t^\tau \frac{1}{\theta} \cdot (r \cdot (1-\theta) - \rho) \cdot dv \cdot d\tau \right)^{-1}$$

⁷There are two implicit assumptions in this result. First, consumers are free to borrow against future labor income. Second, Ponzi games are not allowed. The first assumption ensures that the flow budget constraint operates at all times. The second assumption implies that the usual transversality condition applies (i.e. $\lim_{t \rightarrow \infty} a_j \cdot (c_j + \beta_j)^{-\theta} \cdot e^{-\rho \cdot t} = 0$). Jointly, these assumptions allow us to integrate the Euler equation:

$$\dot{c}_j = \frac{1}{\theta} \cdot (r - \rho) \cdot (c_j + \beta_j)$$

and the flow budget constraint in (2) and, combining these two integrals, obtain the consumption function in (3). See, for instance, Blanchard and Fisher (1989, chapter 2) and Barro and Sala-i-Martin (1995, chapter 2).

$$m_2 = m_1 \cdot \int_t^\infty w \cdot e^{-\int_t^\tau r \cdot dv} \cdot d\tau$$

$$m_3 = -1 + m_1 \cdot \int_t^\infty e^{-\int_t^\tau r \cdot dv} \cdot d\tau$$

Note that the consumption function is linear in a_j , π_j and β_j .⁸ This property of the chosen representation of preferences, combined with the linearity of the budget constraint, is crucial in allowing the nice aggregation properties of this model.

This economy admits a RC, in the usual sense that the sum of all consumers behaves exactly *as if* the economy contained a single consumer with average asset holdings, labor productivity and taste parameter. To see this, let $c = \frac{1}{J} \cdot \sum_j c_j$ and $a = \frac{1}{J} \cdot \sum_j a_j$ be the average per capita consumption and assets. Also, let the average value of the taste and labor-productivity parameters be $\beta = \frac{1}{J} \cdot \sum_j \beta_j$ and $\pi = \frac{1}{J} \sum_{j=1}^J \pi_j$. Summing (2) and (3) over all consumers and dividing by J , we find that:

$$\dot{a} = r \cdot a + w \cdot \pi - c \quad (4)$$

$$c = m_1 \cdot a + m_2 \cdot \pi + m_3 \cdot \beta \quad (5)$$

These equations are those of the average consumer. This implies that all models satisfying our assumptions have predictions for average quantities and prices that are indistinguishable from those of the homogeneous-consumer models. Within this class of models, if one is only interested in the behavior of averages, there is no loss of generality in assuming that all consumers are identical, and readily proceed to a discussion of firms and technology. However, one should not be misled into believing that RC models do not

⁸A consumer with Stone-Geary preferences might choose a negative consumption in some periods. Since we cannot find any meaningful economic interpretation of a negative consumption rate, we rule out this possibility in what follows. One can read this assumption as a restriction on the distribution of individual characteristics, (i.e. $\tilde{a}(0)$, $\tilde{\beta}$ and/or $\tilde{\pi}$) or, alternatively, as a restriction on the permissible functions for factor prices (i.e. r and w).

have interesting implications for the cross-section of consumers. On the contrary, this class of models is consistent with a rich variety of cross-sectional behavior.

2.2 The Cross-section of Consumption Rates

Define $c_j^R = \frac{c_j}{c}$ and let $\tilde{c}^R(t) = (c_1^R(t), \dots, c_J^R(t))$ be the vector of relative consumptions. We refer to the vector \tilde{c}^R as the cross-section of consumptions.

Equations (3) and (5) imply that the initial cross-section of consumptions is a linear combination of the distribution of all three individual characteristics of the consumer (i.e. initial wealth, skills and tastes). That is, the j th element of $\tilde{c}^R(0)$ is:

$$c_j^R(0) = \frac{m_1 \cdot a_j(0) + m_2 \cdot \pi_j + m_3 \cdot \beta_j}{m_1 \cdot a(0) + m_2 \cdot \pi + m_3 \cdot \beta} \quad (6)$$

Differentiating c_j^R with respect to time (use the Euler equation of each consumer in footnote 7), we find that the j th element of \tilde{c}^R has the following law of motion:

$$\dot{c}_j^R = \frac{r - \rho}{\theta \cdot c} \cdot (\beta_j - \beta \cdot c_j^R) \quad (7)$$

or, alternatively, we have that:

$$\Delta c_j^R(t) = \frac{1 - e^{-\int_0^t \frac{r-\rho}{\theta \cdot c} \cdot d\tau}}{c(0)} \cdot (\beta_j - \beta \cdot c_j^R(0)) \quad (8)$$

where $\Delta c_j^R(t) = c_j^R(t) - c_j^R(0)$.⁹ Equation (7) or its integral version (8) completely characterize the behavior of the cross-section of consumptions

⁹To obtain equation (8), integrate (7) to find that:

$$c_j^R(t) = c_j^R(0) \cdot e^{-\int_0^t \frac{r-\rho}{\theta \cdot c} \cdot \beta \cdot d\tau} + \beta_j \cdot \int_0^t \frac{r-\rho}{\theta \cdot c} \cdot e^{-\int_\tau^t \frac{r-\rho}{\theta \cdot c} \cdot \beta \cdot dv} \cdot d\tau$$

and simplify this expression, using the fact that, integrating the consumer's Euler equation

from any initial cross-section as a function of average quantities, prices and the distribution of taste parameters, $\tilde{\beta}$.

To understand the intuition behind equations (7) and (8) one should remember the trade-off that consumers face when deciding how much to save. Consider a growing economy, i.e. $r > \rho$. On the one hand, since the rate of return exceeds the rate of time preference consumers would like to postpone consumption and take advantage of high rates of return. On the other hand, since the economy is growing, consumers would like to advance consumption in order to smooth it over time. The strength of this consumption-smoothing desire is measured by the intertemporal elasticity of substitution, i.e. $\frac{1}{\theta} + \frac{\beta_j}{\theta \cdot c_j}$. The larger this elasticity, the less the consumer cares about consumption-smoothing and, as a result, the more the consumer saves and takes advantage of growth opportunities. Since all consumers face the same rate of return and exhibit the same rate of time preference, relative consumption positions improve (worsen) for those consumers that have an elasticity of intertemporal substitution larger (smaller) than average, i.e. $\frac{\beta_j}{c_j} > \frac{\beta}{c}$.¹⁰

Let $\gamma = 1 - \lim_{t \rightarrow \infty} e^{-\int_0^t \frac{r-\rho}{\theta} \cdot d\tau}$. This limit is directly related to the growth path of the economy. For instance, in an economy with positive long-run growth, the rate of return is bounded above the rate of time preference asymptotically, and $\gamma = 1$. In an economy without long-run growth, if the starting point is the steady-state, the rate of return equals the rate of time preference and $\gamma = 0$. If the economy starts below (above) its steady-state, then the rate of return exceeds (falls short of) the rate of time preference, and declines (increases) towards it monotonically. In this case, $0 < \gamma < 1$ ($0 < \gamma$). Using this definition, we can write the limiting distribution or

(see footnote 7):

$$c(t) + \beta = (c(0) + \beta) \cdot e^{-\int_0^t \frac{r-\rho}{\theta} \cdot d\tau}$$

¹⁰Note that in an economy with negative growth just the opposite is true. Consumers with a high elasticity of intertemporal substitution will save very little and, as a result, find that their relative consumption positions worsen.

cross-section of consumptions as follows:

$$\lim_{t \rightarrow \infty} c_j^R(t) = c_j^R(0) \cdot \left(1 - \frac{\beta \cdot \gamma}{c(0)}\right) + \beta_j \cdot \frac{\gamma}{c(0)} \quad (9)$$

The limiting distribution of consumptions is not degenerate, as it is common in models that allow for cross-sectional variation in tastes to be reflected in cross-sectional variation in consumption growth rates. The reason is that asymptotically all consumers choose the same rate of consumption growth. In economies with long-run growth, this is so because the elasticities of intertemporal substitution of all consumers converge asymptotically to θ^{-1} . In economies without long-run growth, this is so because the rate of return converges asymptotically to the rate of time preference. In either case, all consumers choose the same consumption growth rate in the limit, although there might be substantial cross-sectional variation in consumption growth rates during the transition.

2.3 The Cross-section of Wealth

Next define $a_j^R = \frac{a_j}{a}$ and let $\tilde{a}^R(t) = (a_1^R(t), \dots, a_J^R(t))$ be the vector of relative asset holdings or the cross-section of wealth. Using equations (2)-(5), it is easy to show that the j th element of \tilde{a}^R has the following law of motion:

$$\dot{a}_j^R = (d_2 \cdot \pi + d_3 \cdot \beta) \cdot a_j^R - d_2 \cdot \pi_j - d_3 \cdot \beta_j \quad (10)$$

where d_2 and d_3 are defined as follows:

$$d_2 = \frac{m_2 - w}{a} \quad d_3 = \frac{m_3}{a} \quad (11)$$

The integral form of equation (10) is:

$$a_j^R(t) = a_j^R(0) \cdot e^{\int_0^t (d_2 \cdot \pi + d_3 \cdot \beta) \cdot d\tau} - \int_0^t (d_2 \cdot \pi_j + d_3 \cdot \beta_j) \cdot e^{\int_0^\tau (d_2 \cdot \pi + d_3 \cdot \beta) \cdot dv} \cdot d\tau \quad (12)$$

Either equation (10) or its integral version (12) provide a complete characterization of the cross-section of wealth as a function of d_2 and d_3 and

the distribution of individual characteristics, $\bar{a}_j^R(0)$, $\bar{\pi}_j$ and $\bar{\beta}$. Note that the functions d_2 and d_3 depend only on average variables and the latter are independent of their distribution by virtue of the RC result.

There are a number of interesting questions regarding the cross-section of wealth that can be framed in terms of conditions on the functions d_2 and d_3 . For instance, one might want to know whether in a particular time period, holding constant other individual characteristics, poor consumers tend to accumulate faster than rich ones. This question is equivalent to asking whether $\int_0^t (d_2 \cdot \pi + d_3 \cdot \beta) \cdot d\tau$ is negative. The smaller this integral is, the faster is the speed of (conditional) convergence among consumers.

One might also want to know what the model predicts about the cross-country correlation between average growth and changes in the cross-section of wealth. Although the sign of this correlation crucially depends on the specification of firms and technology that one chooses, there is an interesting observation that can be made at this stage. Using the definitions of d_2 and d_3 , one can integrate equation (4) to find that:

$$a(t) = a(0) \cdot e^{\int_0^t (r - m_1 - d_2 \cdot \pi - d_3 \cdot \beta) \cdot d\tau} \quad (13)$$

Equations (12) and (13) show that, if we observe a collection of economies *with the same (path for) r and m_1* ,¹¹ one should find that the higher the average growth rate in a given period, the faster the speed of (conditional) convergence among consumers. If we are willing to assume that the distribution of skill and taste parameters is more concentrated than the distribution of assets, this implies that, *ceteris paribus*, high-growth countries would experience a faster improvement (or slower worsening) in measures of wealth inequality than low-growth countries.¹²

¹¹For instance, a set of open economies with identical values for ρ and θ .

¹²This is true in our model without uncertainty. If consumers are subject to shocks the existence of conditional convergence (divergence) does not necessarily imply that measures of inequality or dispersion improve (worsen). For a fuller discussion of this point, see the distinction between β -convergence and σ -convergence in Barro and Sala-i-Martin (1995, chapters 1 and 11).

2.4 The Cross-section of Income

Since available cross-sectional data is mainly on personal income, it is interesting to determine the model's implications for the cross-section of incomes, $y_j = r \cdot k_j + w \cdot \pi_j$. As usual, define $y_j^R = \frac{y_j}{y}$ and let $\tilde{y}^R(t) = (y_1^R(t), \dots, y_J^R(t)) \in \mathfrak{R}_+^J$ be the vector of relative incomes. A simple algebraic manipulation shows that:

$$y_j^R = \alpha \cdot a_j^R + (1 - \alpha) \cdot \frac{\pi_j}{\pi} \quad (14)$$

where α is the share of capital income in total income, i.e. $\alpha = \frac{r \cdot a}{y}$. Therefore, the only extra piece of information required to derive theoretical restrictions for the cross-section of real incomes is the function α .

3 Distribution in Two Popular Models

Much of existing growth theory takes the RC model and characterizes the behavior of average quantities under alternative assumptions about firms and technology. The reinterpretation of the RC model developed above allows us to expand the range of these theoretical exercises to include a characterization of the whole distribution of the quantities of interest. More precisely, we can use the following algorithm to extract testable restrictions regarding the cross-sections of consumption, wealth and income from any RC model:

1. Solve your preferred RC model and characterize the equilibrium time-path of average quantities, prices and the functions d_2 , d_3 and α ;
2. Make assumptions about the distribution of individual characteristics, $\tilde{a}(0)$, $\tilde{\pi}$ and $\tilde{\beta}$;
3. With the information obtained in the previous steps at hand, use the appropriate cross-sectional equations to obtain testable restrictions regarding the evolution of any initial cross-section of consumptions, wealth and/or real incomes.

Of course, the usefulness of this algorithm depends on whether the class of models defined by our assumptions includes interesting specifications. By way of example, this section shows that this is the case.

3.1 The Ramsey Model and the Kuznets Curve

Assume that there is a large number of identical competitive firms with free access to the following technology $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$, that maps the stock of capital that a worker uses and its skill parameter to production per worker. Let k be the average stock of capital (physical and human) in the economy. We assume that f is homogeneous of degree one and that $f_k > 0$, $f_{kk} < 0$ and $\lim_{k \rightarrow \infty} f_k = 0$. Since all firms are identical, they will use the same proportions of capital and skills, and these proportions have to be the average ones. Moreover, since firms are competitive, factor prices are given by the marginal products of the relevant factor, i.e. $r = f_k(k, \pi)$ and $w = f(k, \pi) - k \cdot f_k(k, \pi)$. Moreover, in autarky the average (net) stock of assets must equal the capital stock, $a = k$.

Average variables in this economy behave exactly as in the well-known Ramsey-Cass-Koopmans model. A complete characterization of the distributive dynamics that are possible in this model is outside the scope of this paper. However, we have been able to establish a number of interesting properties for the logarithmic case, i.e. $\theta = 1$. This is a very convenient parameter restriction because it implies that $d_2 \cdot \pi + d_3 \cdot \beta = \frac{\dot{z}}{z}$, where $z = \frac{c + \beta}{k}$.

¹³ Accordingly, a cross-section of consumers exhibits conditional convergence (divergence) in wealth whenever z is falling (rising).

To study this model, it is therefore useful to rewrite the standard Ramsey dynamical system in (z, k) space as follows:

$$\dot{z} = \left(z - \rho + \frac{f(k, \pi) + \beta}{k} - f_k(k, \pi) \right) \cdot z \quad (15)$$

$$\dot{k} = f(k, \pi) + \beta - z \cdot k \quad (16)$$

As is well-known, this system has a single steady state that is saddle-path stable. The steady-state values of z and k are implicitly defined by: $f_k(k^*, \pi) = \rho$ and $z^* \cdot k^* = f(k^*, \pi) + \beta$. If the economy's initial capital is k^* , we have that the cross-section of wealth does not vary over time. ¹⁴ This is also true for the cross-section of consumptions, since $\gamma = 0$.

If the economy starts from below the steady-state, we know that k approaches monotonically its steady-state value. However, the behavior of z depends crucially on the properties of $f(\cdot, \cdot)$. For instance, consider first the popular Cobb-Douglas technology: $f(k) = k^\alpha \cdot \pi^{1-\alpha}$. Figure 1 shows the phase diagram of this example. As the economy travels along the stable arm, z declines monotonically towards its steady-state value. As a result, the Ramsey model with logarithmic utility and Cobb-Douglas technology predicts conditional convergence during the transition (from below) towards the steady-state.

To caution the reader against drawing quick conclusions from the Cobb-Douglas example, consider this CES technology: $f(k) = (k^b + \pi^b)^{\frac{1}{b}}$ with

¹³To see this, note that $\frac{\dot{z}}{z} = \frac{\dot{c}}{c + \beta} - \frac{\dot{k}}{k} = \frac{1}{k} \cdot ((m_2 - w) \cdot \pi + m_3 \cdot \beta) = d_2 \cdot \pi + d_3 \cdot \beta$.

¹⁴To see this, note that $\dot{z} = 0$ and, as a result, $d_2 \cdot \pi + d_3 \cdot \beta = 0$. Since the latter is true for any value of π and β , it follows that $d_2 = d_3 = 0$. Therefore, equation (12) implies that $k_j^R(t) = k_j^R(0)$.

$b < 0$. The elasticity of substitution between capital and labor is less than one, i.e. $(1 - b)^{-1} < 1$. Figure 2 shows the phase diagram for this economy, assuming that ρ is high enough (so that the steady-state is located in the upward-sloping region of the $\dot{z} = 0$ line) and $\beta = 0$. In this economy, z is always increasing and, as a result there is always conditional divergence. Figure 3 shows the case in which ρ is low. If the economy starts at a low enough level of capital, z is initially increasing and then declines. At low levels of development, the economy exhibits conditional divergence, while at high levels of development it is characterized by conditional convergence.

This last result is reminiscent of one of the most famous ideas in economic development: the existence of an inverted-U shaped relationship between measures of income inequality and the level of income of a country. This curve is named the Kuznets curve, since it was Kuznets (1955) who first hypothesized its existence and gave some evidence in its support. Since then a large number of empirical studies have piled up either to further document or to deny the existence of the Kuznets curve. At the same time, a substantial body of research has tried to generate Kuznets curves in theoretical models. Without taking a stand on whether the Kuznets curve is a fact or just an hypothesis, we simply note that our CES example implies that the Ramsey model is consistent with the Kuznets curve hypothesis. To show this, we use the coefficient of variation of k_j as our measure of inequality: $CV = \sqrt{\text{Var} \{k_j^R\}}$; and assume, for simplicity, that there is no cross-sectional variation in tastes and skills, i.e. $\beta_j = \beta$ and $\pi_j = \pi$ for all j . Simple algebra shows that:

$$CV(t) = CV(0) \cdot e^{\int_0^t (d_2 \cdot \pi + d_3 \cdot \beta) \cdot d\tau} \quad (17)$$

That is, inequality increases (decreases) in periods where z is increasing (decreasing), i.e. $d_2 \cdot \pi + d_3 \cdot \beta > 0$ ($d_2 \cdot \pi + d_3 \cdot \beta < 0$). It follows that, in a Ramsey economy with CES technology as depicted in figure 2, at low levels of development inequality is rising, while at sufficiently high levels of development inequality is falling. This is nothing but the Kuznets curve.¹⁵ Some intuition for this result can be obtained by noting that, if $\beta = 0$:

¹⁵Note, however that in the Ramsey economy there is no Kuznets curve for consumption. The cross-section of consumption rates will always exhibit convergence (if $\beta > 0$) or divergence (if $\beta < 0$).

$$d_2 \cdot \pi + d_3 \cdot \beta = \frac{\rho \cdot \int_t^\infty w \cdot \pi \cdot e^{-\int_t^\tau r \cdot dv} \cdot d\tau - w \cdot \pi}{k}$$

Ceteris paribus, if the growth rate of wages is low, ϕ is negative and consumers that have a low stock of capital relative to their labor productivity parameter, accumulate capital at a higher rate. The intuition for this result follows from the assumption that the economy is populated by permanent-income consumers. These consumers calculate the net present value of their income and choose their optimal consumption path. Since all consumers have the same spending shares (a property of homothetic preferences), they all spend the same fraction of their wealth in each date and therefore exhibit identical rates of wealth accumulation. But the growth rate of wealth is a weighted average of the growth rates of its two components, the stock of capital and the net present value of wages. The growth rate of the latter is the same for all consumers. If this growth is low, consumers must be accumulating capital at a rate that exceeds that of total wealth. But how much? It depends on how large is the share of capital in a consumer's total wealth. The lower the share of capital is, the higher is the rate of capital accumulation that is required to sustain the optimal consumption path. This is why consumers that have a low stock of capital relative to their labor productivity parameter tend to accumulate capital faster when the growth rate of wages is low. A symmetric argument works for the case in which the net present value of wages grows at a high rate. In our Kuznets-curve example, wage growth accelerates as the economy accumulates capital. If ρ is low enough, this leads to a reversal in the sign of $d_2 \cdot \pi + d_3 \cdot \beta$.¹⁶

As the examples here show, the distributive dynamics of the Ramsey model depend crucially on the properties of the technology that one assumes. Figures 1 to 3 show examples in which there is conditional convergence always, never and sometimes. More general dynamics are also possible. In particular, it is not difficult to generate conditional convergence and divergence cycles by appropriately choosing the production function. A necessary (but not sufficient) condition for having p alternating periods of conditional convergence and divergence is that the first derivative of $g(k) = f_k(k, \pi) - \frac{f(k, \pi) + \beta}{k}$, change signs $p - 1$ times as k increases.

¹⁶See Ventura (1996) for a similar argument in the context of international growth comparisons.

3.2 Distribution in a Linear Growth Model

Assume that there is a large number of competitive firms with free access to the following linear technology: $f(k) = A \cdot k + \pi_j$; where k is the average stock of capital (physical and human) and the parameter A satisfies the following restrictions $A > \rho > A \cdot (1 - \theta)$. The first inequality generates positive growth rates, while the second inequality ensures that utility is bounded. To rule out negative consumption, we assume that $k(0) > \max \left\{ -\frac{\beta + \pi}{A}, \frac{A - \rho}{\rho + (\theta - 1) \cdot A} \cdot \frac{\beta}{A} - \frac{\pi}{A} \right\}$.

Since firms are competitive, factor prices are given by the marginal products of the relevant factor, i.e. $r = A$ and $w = 1$. Also, we have that the average stock of assets must equal the capital stock, $a = k$. It follows immediately from equation (5) that average consumption is given by:

$$c = \frac{\rho + (\theta - 1) \cdot A}{\theta} \cdot \left(k + \frac{\pi}{A} \right) - \frac{A - \rho}{\theta} \cdot \frac{\beta}{A} \quad (18)$$

Substituting this consumption function in (4) and integrating, we find that:

$$k(t) = \left(k(0) + \frac{\beta + \pi}{A} \right) \cdot e^{\frac{A - \rho}{\theta} \cdot t} - \frac{\beta + \pi}{A} \quad (19)$$

If $\beta + \pi \geq 0$ ($\beta + \pi \leq 0$) both average consumption and the capital stock grow at a rate that is positive, non-increasing (non-decreasing) and asymptotically approaches $\theta^{-1} \cdot (A - \rho)$.

Next we want to determine the cross-section of wealth.¹⁷ Using equation (12) we find that:

$$\Delta k_j^R(t) = \frac{\left(1 - e^{-\frac{A - \rho}{\theta} \cdot t} \right) \cdot \left(\beta_j + \pi_j - (\beta + \pi) \cdot k_j^R(0) \right)}{A \cdot k(0) + (\beta + \pi) \cdot \left(1 - e^{-\frac{A - \rho}{\theta} \cdot t} \right)} \quad (20)$$

¹⁷Remember that the information required for this purpose is summarized by two functions: d_2 and d_3 . It follows from (18) and (19) that:

$$d_2 = d_3 = \frac{\rho - A}{\theta} \cdot \left[(A \cdot k(0) + \beta + \pi) \cdot e^{\frac{A - \rho}{\theta} \cdot t} - (\beta + \pi) \right]^{-1}$$

This equation shows that the relative-wealth position of a consumer increases if and only if $\beta_j + \pi_j > (\beta + \pi) \cdot k_j^R(0)$ (since, by assumption, $A \cdot k(0) > -(\beta + \pi)$). It follows from this equation that the linear economy exhibits conditional convergence (divergence) if and only if $\beta + \pi \geq 0$ ($\beta + \pi \leq 0$). Interestingly, in the linear model there cannot be alternating periods of conditional convergence and divergence, as in the Ramsey model.

To understand the intuition behind equation (20) it is useful to distinguish between the effects of β_j and π_j in determining changes in the distribution of wealth. Consider first the taste parameter. As mentioned in the previous section, if a consumer satisfies the following condition: $\beta_j > \beta \cdot k_j^R(0)$, her elasticity of intertemporal substitution exceeds that of the average consumer. *Ceteris paribus*, this consumer saves more than average and her relative wealth position improves. Not surprisingly, we find that, in a growing economy, consumers that are flexible in the timing of their consumption will be moving upwards in the wealth distribution. Consider next the skills parameter. If a consumer satisfies the following condition: $\pi_j > \pi \cdot k_j^R(0)$, she has a larger share of wage income in wealth. In a growing economy, the marginal propensity to consume out of wage income is less than one.¹⁸ *Ceteris paribus*, the larger is the share of wage income in wealth, the larger is the addition to the stock of assets in every period. As a result, we find that consumers with high share of wage income in wealth save more and move upward in the wealth distribution.

Finally, one can also use equation (20) to determine the long-run or steady-state distribution of wealth:

$$\lim_{t \rightarrow \infty} k_j^R(t) = k_j^R(0) \cdot \frac{A \cdot k(0)}{A \cdot k(0) + \beta + \pi} + (\beta_j + \pi_j) \cdot \frac{1}{A \cdot k(0) + \beta + \pi} \quad (21)$$

As this expression shows, the long-run distribution is a weighted average of the initial wealth distribution and the distributions of taste and skill parameters. If $\beta + \pi \geq 0$, the limiting distribution of wealth moves away from the initial distribution of wealth and towards the distribution of skills and tastes. If $\beta + \pi \leq 0$, the limiting distribution of wealth accentuates the initial distribution of wealth. This is nothing but another way of looking at the convergence properties of the model, already discussed above.

¹⁸In this linear economy, this marginal propensity is $1 - \theta^{-1} \cdot \left(1 - \frac{\rho}{A}\right)$.

4 Economic Growth in a Cross-section of Consumers

In this section we estimate a number of “growth regressions” for a representative sample of American families. We find evidence of conditional convergence throughout the 1970’s and divergence in the 1980’s. Since, unlike the Ramsey model, the linear model does not admit reversals in the pattern of convergence, this can be seen as an empirical failure of the simplest of growth models.

4.1 Objective and Motivation

In the last ten years a substantial literature has studied the dynamic properties of per capita income in cross-sections of countries and regions. A common finding is that, once a set of control variables is held constant, there is a negative correlation between initial GDP and subsequent growth. This finding is usually referred to as conditional convergence, because it implies that, conditional on the controls, economies that have relatively high levels of income at the beginning of the period will grow at a slower rate than initially poor countries.

The theoretical results of this paper lead us naturally to investigate an analogous empirical specification for individual incomes. Using equations (12) and (14), some algebra allows us to write:

$$\Delta y_j^R(t+h) = (\lambda_1(t, h) - 1) \cdot y_j^R(t) + \lambda_2(t, h) \cdot \pi_j + \lambda_3(t, h) \cdot \beta_j \quad (22)$$

If $\lambda_1(t, h) < 1$, there is a negative (partial) correlation between initial relative income and the subsequent change.¹⁹ This is akin to conditional convergence: it implies that, conditional on π_j and β_j , lower income individuals have gained ground relative to higher income ones. On the other hand, if $\lambda_1(t, h) > 1$ those who started the period with high relative incomes find their relative

¹⁹The coefficients in equation (21) are given by:

$$\begin{aligned} \lambda_1(t, h) &= \frac{\alpha(t+h)}{\alpha(t)} \cdot e^{\int_t^{t+h} (d_2 \cdot \pi + d_3 \cdot \beta) \cdot d\tau} \\ \lambda_2(t, h) &= -\frac{\Delta \alpha(t+h)}{\pi} - \frac{\lambda_1(t, h) + 1}{\pi} \cdot (1 - \alpha(t)) - \end{aligned}$$

position further improved, an outcome we will describe as conditional divergence.

The examples we develop in Section 3 indicate that different models have different implications for the expected magnitude of λ_1 . In the linear technology model we have shown that $\partial \Delta k_j^R(t) / \partial k_j^R(0) < 0$ (and hence $\lambda_1 < 1$) if and only if $\beta + \pi > 0$. Therefore, both convergence and divergence are possibilities, according to whether the equalizing effect of labor income is stronger or weaker than the effect of intertemporal substitution (that goes in the opposite direction if $\beta < 0$). However, there is no possibility of sign reversal ($\beta + \pi$ being a constant): in the linear technology model, if there is one period characterized by conditional convergence all other periods must display convergence too. On the other hand, our Kuznets curve example shows that, in the Ramsey model, it is possible to have reversals in the direction of distributive dynamics.

A closely related interpretation of equation (22) is as a decomposition of relative income in its determinants: initial income, skills and tastes. Rewrite (22) in levels (assuming $\beta \neq 0$):

$$y_j^R(t+h) = \lambda_1(t, h) \cdot y_j^R(t) + \pi \cdot \lambda_2(t, h) \cdot \pi_j^R + \beta \cdot \lambda_3(t, h) \cdot \beta_j^R \quad (23)$$

where $\pi_j^R = \frac{\pi_j}{\pi}$ and $\beta_j^R = \frac{\beta_j}{\beta}$. Clearly, $\lambda_1(t, 0) = 1$, and $\lambda_2(t, 0) = \lambda_3(t, 0) = 0$. Now suppose that, for $h > 0$, $\lambda_1(t, h) < 1$. This implies that $\pi \cdot \lambda_2(t, h) + \beta \cdot \lambda_3(t, h) > 1$ (see footnote 18). Hence, conditional convergence implies that the contribution of initial income has fallen over the period, while either or both of the weights of skills and tastes have risen. In other words, the distribution of income is moving away from the distribution of initial income, and in the direction of a combination of the distributions of tastes and skills. An opposite conclusion applies if $\lambda_1(t, h) > 1$.

$$\lambda_3(t, h) = -\alpha(t+h) \cdot \int_t^{t+h} d_3 \cdot e^{\int_t^{\tau} (d_2 \cdot \pi + d_3 \cdot \beta) \cdot dv} \cdot d\tau$$

These coefficients satisfy the restriction $\lambda_1(t, h) + \lambda_2(t, h) \cdot \pi + \lambda_3(t, h) \cdot \beta = 1$.

Our objective is to provide estimates of $\lambda_1(t, h)$ for a cross-section of American families for the period 1970-1990 (and its subperiods). As is well known, this is an epoch of increased dispersion in individual incomes. The foregoing discussion implies that whether we find evidence of conditional convergence or divergence has far reaching implications for the nature of the recent increase in inequality.²⁰ If there is convergence, increased dispersion implies that market forces have increasingly rewarded skills and flexibility in consumption, and that these characteristics are widely dispersed among the population. If there is divergence, inequality has increased merely because the rich has gotten richer, regardless of the distribution of skills and tastes.²¹

4.2 Data and Method

We work with data from the Panel Study of Income Dynamics (PSID). The PSID contains comprehensive information on the incomes of a large number of families and their members, and the panel – initiated in 1968 – covers a relatively long time span. We extract annual cross-sections of total family money income between 1970 and 1990. This variable includes income of all sources (labor, interest, transfers,...) earned by all the members of the family. Equation (22) require us to work with relative incomes. We thus divide size-adjusted family income by the population average.²² Table 1 reports some descriptive statistics of these cross-sections of family incomes for three selected years: 1970, 1980 and 1990. In every interview year a “head” is identified for each family. In order to compare reasonably homogeneous age groups, in Table 1 we focus, for each year, on families whose head is between 20 and 50 years of age.

The story told by Table 1 is consistent with a wealth of recent empirical

²⁰Increasing income dispersion is not inconsistent with conditional convergence: if the skill and taste coefficients are more widely dispersed than initial income, conditional convergence can take place at the same time as absolute divergence.

²¹Note that equation (23) imposes two strong restrictions on the dynamics of individual (relative) incomes. First, for a given individual, current income depends linearly on initial income. We refer to this as the linearity property. Second, the coefficient in this linear relationship is the same across individuals. That is, the quantity $\lambda_1(t, h)$ is a function of aggregate variables only, and does not depend on any individual characteristic of the consumer (i.e. π_j and β_j). We call this an invariance property. In the appendix we present a joint test of these linearity and invariance properties.

²²More details on how we put together our data are given in Appendix 1.

Table 1: Family Income Distribution in 1970, 1980 and 1990

Year	Std. Dev.	10 _{th} Perc.	25 _{th} Perc.	75 _{th} Perc.	90 _{th} Perc.	Obs	Avg. age
1970	0.639	0.355	0.564	1.281	1.705	3160	35.173
1980	0.863	0.279	0.499	1.283	1.827	4510	33.451
1990	0.904	0.244	0.469	1.299	1.846	5691	36.642

Note: Summary statistics of the cross-sections of 1970, 1980 and 1990 family income relative to the population average. For each year, only families whose head is between 20 and 50 years of age are included.

studies documenting an increase in income inequality throughout the 70s and the 80s. The coefficient of variation of the family income distribution has increased rather dramatically over the sample period (recall that this measure of inequality coincides with the standard error of our income variable, and is thus reported in Table 1). In addition, both the 90_{th}-10_{th} and the 75_{th}-25_{th} percentile ranges have increased in both sub-decades. Hence, in the lexicon of the recent empirical growth literature, Table 1 confirms that the last 20 years are characterized by σ -divergence of the family income distribution. It remains to be seen whether there is β -divergence or convergence.

We treat $\lambda_1(t, h)$ as an exogenous parameter to be estimated with cross-sectional techniques. Specifically, for various t and h , we run cross-family regressions of the form:

$$y_j^R(t + h) = \gamma' X_j + \delta \cdot y_j^R(t) + \epsilon_j(t + h) \quad (24)$$

where X_j is a vector containing a constant and a number of individual characteristics of the head of family j (age, age squared, education, sex and race) and $\epsilon_j(t + h)$ is a zero-mean stochastic error term. Hence, we are treating $\gamma' X_j$ as the linear predictor of the quantity $\lambda_2(t, h) \cdot \pi_j + \lambda_3(t, h) \cdot \beta_j$, given X_j . This method allows us to interpret δ as an estimate of $\lambda_1(t, h)$ in equation (23).

Estimation of (24) requires an instrumental variable for $y_j^R(t)$. First, the error term contains the deviation of $\lambda_2(t, h) \cdot \pi_j + \lambda_3(t, h) \cdot \beta_j$ from its linear predictor. It is unlikely that this deviation will be incorrelated with $y_j^R(t)$. Hence, there is a potential endogeneity (or omitted variable) problem.

Second, measurement error is a pervasive problem in the PSID. Measurement error would induce downward bias in an ordinary least squares estimate of (24), while the correlated, omitted individual effect induces upward bias if $\lambda_1(t, h) < 1$, and downward bias if $\lambda_1(t, h) > 1$.

We use as instruments *aggregate* measures of income in the state of residence of individual j . More precisely, we use aggregate data from the *Statistical Abstract of the United States* to construct the instrumental variables

$$MN_j = \frac{Y_j^{MN}(t)}{Y_{US}^{MN}(t)} \text{ and } MD_j = \frac{Y_j^{MD}(t)}{Y_{US}^{MD}(t)}$$

where $Y_j^{MN}(t)$ ($Y_j^{MD}(t)$) is per capita income (median family income) in the state where family j was resident in year t , and $Y_{US}^{MN}(t)$ ($Y_{US}^{MD}(t)$) is the average (median) U.S. level. The basic idea is that the skill and taste characteristics of a specific family are unlikely to have an impact on state-wide summary statistics, such as per capita income or median family income. Hence, our instruments should be uncorrelated with the individual effect in the error term.²³ On the other hand, the income of a family is likely to be affected by the general level of activity in the area where the family is economically active. Therefore, our instruments will have some degree of correlation with the variable to be instrumented.²⁴

Although our empirical exercise is close in spirit to the recent cross-country empirical growth literature, our estimation strategy is quite different. Because the error term (potentially) contains an unobservable idiosyncratic effect (innate skills, tastes) we cannot estimate a cross-sectional regression using ordinary-least-squares. This problem is also present in cross-country growth regressions if there are unobserved differences in, say, the aggregate

²³One possible problem is that there might be a “state effect” in the labor productivity parameter π_j . We would argue, however, that the state effect enters mainly through systematic differences in average educational achievement. Because we control for education the part of π_j that is left in the error term should be largely purged from state effects. This intuition is confirmed by statistical testing: we perform Hansen’s tests for the validity of the overidentifying restrictions for all the regressions we present, and we always fail to reject.

²⁴In principle, this correlation may be tenuous. Hence, for each instrument and subsample we have checked the robustness of the instrument by computing the F statistic for a “first stage” ordinary least squares regression of our explanatory variable on a constant and the instrumental variables themselves. MD and MN are the two candidate instruments, of an initially larger set, for which such F tests are systematically and comfortably high.

production function. Some recent empirical growth papers have tried to solve the problem using panel data techniques, such as differencing away the country-specific effect. In our case, however, the individual effects multiply time-varying coefficients ($\lambda_2(t, h)$ and $\lambda_3(t, h)$ in equation 23), so that differencing would not eliminate the problem.²⁵

4.3 Results

We start by studying conditional convergence in the two ten-year sub-periods: the 1970's and the 1980's. Hence, we first estimate equation (24) with $t = 1970$ and $h = 10$, and then with $t = 1980$ and $h = 10$. Table 2 presents a variety of results. Each row corresponds to a different combination of sample selection criteria, normalization of the data, list of controls included from vector X_j , and age group of the household head at time t . Hence, the first row uses all families whose head was between 20 and 50 years of age at time t , and the whole set of controls. In the second row of Table 2 we drop all controls (except for age: we want to be sure our results are not driven by life-cycle effects). Rows (3) and (4) drop the very rich and the very poor, and rows (5) and (6) drop families whose head is out of the labor force in either of the initial and the final periods. Rows (7) and (8) add the change in family size as a control, and rows (9) and (10) normalize total family money income by family size. Finally, rows (11)-(13) further explore the robustness of the results to compositional effects, by estimating the growth regressions for relatively homogeneous age groups.

The point estimates appear quite stable across specifications, and several of them are statistically different from 1. All the estimates of $\lambda_1(70, 10)$ are less than 1; all but one (and all of the significant) estimates of $\lambda_1(80, 10)$ are greater than 1. Further variations of the basic specification do not lead to significant changes in the results of Table 2. Hence, we tend to regard the qualitative result that the convergence coefficient is less than one in the

²⁵Besides being analogous to cross-country and cross-regional growth regressions, Equation (23) is also similar in spirit to the inter-generational income mobility regressions estimated, among many others, by Solon (1992), Zimmermann (1992) and Mulligan (1993). The main difference is, of course, that they regress childrens' income on parents' income, while we look at mean reversion (conditional on π_j and β_j) within the life span of an individual.

Table 2: Distributional Dynamics in the 1970s and 1980s

	1 $\lambda_1(70, 10)$	2 $\lambda_1(80, 10)$	3 sample	4 controls	5 init. age	6 obs ₇₀	7 obs ₈₀
1	0.699 (0.237)	1.300 (0.203)	a	all	20-50	2205	2994
2	0.791 (0.163)	1.323* (0.130)	a	age	20-50	2229	3005
3	0.779* (0.105)	1.315 (0.195)	b	all	20-50	2101	2800
4	0.819* (0.078)	1.345* (0.129)	b	age	20-50	2124	2809
5	0.831 (0.104)	1.386% (0.212)	c	all	20-50	1629	1691
6	0.875% (0.071)	1.321* (0.140)	c	age	20-50	1645	1698
7	0.689 (0.239)	1.273 (0.197)	a	all, Δ size	20-50	2205	2994
8	0.785 (0.165)	1.277* (0.126)	a	age, Δ size	20-50	2229	3005
9	0.686* (0.159)	0.930 (0.175)	d	all	20-50	2079	2872
10	0.782 (0.120)	1.024 (0.106)	d	age	20-50	2103	2882
11	0.731 (0.224)	1.375 (0.341)	a	all	20-30	824	1395
12	0.600* (0.171)	1.289 (0.338)	a	all	30-40	741	1140
13	0.362 (0.514)	1.053 (0.313)	a	all	40-50	796	674

Note. Column 1 (2): estimate of δ in equation 23, with $t = 1970$ (1980) and $h = 10$. Standard errors in parenthesis. The symbol * (%) indicates that the 95% (90%) confidence interval does not include 1. Column 3: a: all families; b: drops families with less than 1/10 and more than 10 times average income; c: further drops families whose head is not employed in either t or $t + h$; d: all families, but income is divided by family size. Column 4: all: age, age², sex, race, education; Δ size: change in family size over estimation period. Column 5: age group of family head at t . Column 6 (7): sample size for column 1 (2) regression. Estimation method: optimally weighted, heteroscedasticity consistent two stage least squares.

1970's, and greater than one in the 1980's as reasonably robust.²⁶

The results presented in Table 2 indicate that there has been a reversal in distributive dynamics during the 1970-1990 period. The 1970s emerge as a period of conditional convergence: after controlling for heterogeneity in skills and tastes, we find that individuals who started the decade with relatively high incomes have subsequently lost part of their advantage. In the 1980s instead, individual incomes have conditionally diverged. The appearance of cross-individual conditional divergence in the 1980's stands in sharp contrast with the cross-regional evidence, that tends to be characterized by convergence.²⁷

It is also useful to read the results in terms of a decomposition of relative incomes. In terms of equation (23), the fact that λ_1 has fallen below 1 in the 70's corresponds to an increased weight of skills and tastes in determining the relative position of an individual in the distribution of income. In other words, in the 70's individual success was mainly driven by one's skills and flexibility. The rise of λ_1 above 1 in the 1980's instead, represents a fall in the contribution of skills and tastes, and an increase in the weight of initial wealth. The 1980's are therefore a period that rewards initial wealth and, as a result, they are characterized by a sharp increase in inequality.

Until now we have estimated (24) keeping h fixed and varying t . We can further explore the reversal in distributive dynamics by keeping t constant and letting h vary. Figure 4 plots the point estimates (and confidence bands) of $\lambda(70, h)$ against h , for $h = 1, 2, \dots, 20$. In order to have an homogeneous sample, we include only families whose head has been the same individual in all the 20 years. Apart from that, the specification is the same as in Row 1 of Table 2. The point estimates in Figure 4 show a fairly smooth time path, and confirm the basic insight from Table 2: conditional convergence throughout the 1970's, a turnaround, and divergence thereafter. Unfortunately, the estimates are too imprecise to identify with any confidence the exact date of the turnaround.

²⁶Ordinary least square estimates of the basic specification in row 1 of Table 2 lead to point estimates of 0.634 for $\lambda_1(70, 10)$ and 0.513 for $\lambda_1(80, 10)$. Both are smaller (the second dramatically so) than those reported above: as predicted, ordinary least squares suffer from severe measurement error and/or endogeneity bias.

²⁷See the study of convergence across the U.S. states by Barro and Sala-i-Martin (1995). It is interesting to note, however, that they report much lower estimates of the speed of convergence for the 1980s than they do for the previous decades.

No matter how we look at it, however, it seems clear to us that there has been a reversal in the patterns of convergence around the middle of our sample period. This evidence is inconsistent with the linear growth model which, as we have shown above, does not admit reversals. Without any further restriction on technology, the Ramsey model is consistent with this evidence. One should however note that simple formulations such as logarithmic utility and Cobb-Douglas technology are not capable to generate the pattern in the data. Therefore, we can also use this evidence to narrow the set of empirically admissible technologies that can be associated with the Ramsey model.

5 Concluding Remarks

This paper shows that the RC model gives us ample room to study the forces that determine the evolution of the interpersonal distributions of consumption, wealth and income. We analyze a class of growth models with three sources of heterogeneity: initial wealth, innate skills, and taste for consumption smoothing. Despite this heterogeneity, these economies admit a representative consumer and, as a result, exhibit aggregate dynamics that are indistinguishable from those of the homogenous-consumer models. Yet the existence of heterogeneity allows us to derive equations describing the dynamics of the cross-sections of consumption, wealth and income. We find that the class of models analyzed here is capable of generating a rich variety of distributive dynamics, while retaining the analytical tractability of the homogeneous-consumer models.

We think that a study of the distributive implications of alternative growth models might have a substantial payoff, since a model's ability to generate distributive outcomes that are consistent with the data constitutes a further criterion against which to evaluate its empirical relevance. In this spirit, we study the distributive implications of two popular growth models: the Ramsey and linear growth models. We find that both of them are consistent with conditional convergence and divergence in a cross-section of consumers. However, while the Ramsey model admits the possibility of alternating periods of divergence and convergence, the linear growth model does not. Using data from the PSID, we uncover that there has been a reversal in the pattern of convergence of family incomes at the end of the 1970s or beginning of the 1980s. In particular, the evidence shows that the decade of the 1970s was characterized by conditional convergence while the decade of the 1980s exhibited conditional divergence. This finding is inconsistent with the linear growth model, but not with some versions of the Ramsey model.

References

- [1] Acemoglu, D. (1996), "Matching, Heterogeneity and the Evolution of Income Distribution", CEPR Discussion Paper No. 1345.
- [2] Alesina, A. and R. Perotti (1994), *Income Distribution, Political Instability and Investment*, Columbia University Working Paper 751, October.
- [3] Alesina, A. and D. Rodrik (1994), "Distributive Politics and Economic Growth", *Quarterly Journal of Economics*, 109, 465-490.
- [4] Banerjee, A. and A. Newman (1991), "Risk Bearing and the Theory of Income Distribution", *Review of Economic Studies*, 58, 211-235.
- [5] Banerjee, A. and A. Newman (1993), "Occupational Choice and the Process of Development", *Journal of Political Economy*, 101, 274-297.
- [6] Barro, R.J. (1991), "Economic Growth in a Cross-Section of Countries", *Quarterly Journal of Economics*, 106, 407-443.
- [7] Barro, R.J. and X. Sala-i-Martin (1991) , "Convergence across States and Regions", *Brookings Papers on Economic Activity*, 1, 107-182.
- [8] Barro, R.J. and X. Sala-i-Martin (1992), "Convergence", *Journal of Political Economy*, 100, 223-251.
- [9] Barro, R.J. and X. Sala-i-Martin (1995), *Economic Growth*, McGraw Hill.
- [10] Baumol, W. (1986), "Productivity Growth, Convergence and Welfare: What the Long-Run Data Show", *American Economic Review*, 76, 1072-1085.
- [11] Becker, G. and N. Tomes (1986), "Human Capital and the Rise and Fall of Families", *Journal of Labor Economics*, 4, S1- S39.
- [12] Behrman, J., Pollak, R. and P. Taubman (1995), *From Parent to Child*, University of Chicago Press.

- [13] Bénabou, R. (1996), *Inequality and Growth*, mimeo, New York University.
- [14] Blanchard, O.J. and S. Fischer (1989), *Lectures on Macroeconomics*, MIT Press.
- [15] Caselli, F., G. Esquivel and F. Lefort (1995), "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics", *Journal of Economic Growth*, forthcoming.
- [16] Chamberlain, G. (1983), "Panel Data", in Z. Griliches and M. Intriligator (eds), *Handbook of Econometrics*, North-Holland, .
- [17] Chatterjee, S. (1994), "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model", *Journal of Public Economics*, 54, 97-119.
- [18] De Long (1988), "Productivity Growth, Convergence and Welfare: Comment", *American Economic Review*, 78, 1138-1154.
- [19] Galor, O. and J. Zeira (1993), "Income Distribution and Macroeconomics", *Review of Economic Studies*, 60, 35-52.
- [20] Knight, M., N. Loayza and D. Villanueva (1993), "Testing the Neoclassical Growth Model", *IMF Staff Papers*, 40, 512-541.
- [21] Kuznets, S. (1955), "Economic Growth and Income Inequality", *American Economic Review*, 65, 1-28.
- [22] Loury, G. (1981), "Intergenerational Transfers and the Distribution of Earnings", *Econometrica*, 49, 843-867.
- [23] Mankiw, N.G., D. Romer and D. Weil (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, 107, 407-437.
- [24] Mas-Colell, A., M.D. Whinston and J.R. Green (1995), *Microeconomic Theory*, Oxford University Press.

- [25] Mulligan, C. (1993), "On Intergenerational Altruism, Fertility, and the Persistence of Economic Status", University of Chicago, Ph.D. Dissertation.
- [26] Perotti, R. (1993), "Political Equilibrium, Income Distribution and Growth", *Review of Economic Studies*, 60, 755-776.
- [27] Persson, T. and G. Tabellini (1994), "Is Inequality Harmful for Growth?", *American Economic Review*, 84, 600-621.
- [28] Piketty, T. (1996), "The Dynamics of the Wealth Distribution and Interest Rate with Credit Rationing", *Review of Economic Studies*, forthcoming.
- [29] Solon, G. (1992), "Intergenerational Income Mobility in the United States", *American Economic Review*, 82, 393-407.
- [30] Stiglitz, J.E. (1969), "Distribution of Income and Wealth Among Individuals", *Econometrica*, 37, 382-397.
- [31] Ventura, J. (1996), *Growth and Interdependence*, mimeo, MIT.
- [32] Zimmerman, D. (1992), "Regression Toward Mediocrity in Economic Stature", *American Economic Review*, 408-429.

APPENDIX 1: DATA SELECTION CRITERIA

We extract data from 21 cross-sectional family data files, starting from interview year 1971 (which contains data on income in 1970) through 1991. From each of these we extract the variables total family money income, the family sampling weight, and the family identity number for that year. From the 1971, 1981 and 1991 files we also extract family size and the variable occupation of head, that allows us to determine whether the household head was in the labor force during the sample period (values of 0 or missing are treated as out of the labor force). From the 1971 and 1981 files we further extract the age and construct the sex (0 male, 1 female), race (0 white, 1 nonwhite) and education (9 dummies corresponding to 9 levels of educational achievement, from illiterate to graduate school) dummies that we then include in vector X . To compute the mean of the family income distribution in a given year (in order to construct the relative income measure) we use PSID sample weights. To link families across time, we use the principle that the “successor” family at date $t + h$ of a family at date t must have the same individual as household head. This is implemented by extracting the variables relationship to head and identification number from the cross-year individual file, and combining them with the family data. If a family does not have the same head in t and $t + h$ it is dropped from the sample.

APPENDIX 2: A TEST OF LINEARITY AND INVARIANCE

Equation (23) imposes two strong restrictions on the dynamics of individual (relative) incomes. First, for a given individual, current income depends linearly on initial income. We refer to this as the linearity property. Second, the coefficient in this linear relationship is the same across individuals. That is, the quantity $\lambda_1(t, h)$ is a function of aggregate variables only, and does not depend on any individual characteristic of the consumer (i.e. π_j and β_j). We call this an invariance property. In this appendix we present a joint test of these linearity and invariance properties. To understand the logic of our testing strategy, suppose that we divided the total population of an economy into several income groups, say by quartiles. Next, imagine that we obtain separate estimates of $\lambda_1(t, h)$ for each of these groups. The linearity and invariance properties jointly imply that the estimated coefficients should be the same for all groups. Hence, this logic leads us to perform a piecewise

linear regression of equation (24), and to test the restriction that the slope coefficient is the same across “pieces”. Unfortunately, as the sample size falls the standard errors of our instrumental variable regressions increase dramatically. Suppose that we estimated equation (24) separately for four income brackets. Due to high standard errors a test that $\lambda_1(t, h)$ is the same across these four groups would have very low power.²⁸ In other words, we need an approach involving more precisely estimated parameters, so as to give the data a chance of rejecting the null hypothesis of invariance and linearity.

A feasible strategy is available for the case in which $\beta_j = 0$ for every j . Consider three dates: t , $t + h$ and $t + 2 \cdot h$. Use equation (23) (with $\beta_j = 0$ and adding our stochastic error term $\epsilon_j(t + h)$) to solve for π_j . Substitute the resulting expression for π_j in the $t + 2 \cdot h$ version of equation (23). The result is

$$y_j^R(t + 2 \cdot h) = \alpha_1 \cdot y_j^R(t + h) + \alpha_2 \cdot y_j^R(t) + e_j(t + 2 \cdot h) \quad (A1)$$

where

$$\alpha_1 = \frac{1 - \lambda_1(t, 2 \cdot h)}{1 - \lambda_1(t, h)} \quad \alpha_2 = \lambda_1(t, 2 \cdot h) - \lambda_1(t, h) \cdot \frac{1 - \lambda_1(t, 2 \cdot h)}{1 - \lambda_1(t, h)}$$

and

$$e_j(t + 2 \cdot h) = \epsilon_j(t + 2 \cdot h) - \frac{1 - \lambda_1(t, 2 \cdot h)}{1 - \lambda_1(t, h)} \cdot \epsilon_j(t + h)$$

We thus have a reduced form which is entirely in terms of observed variables (income in three time periods). In addition, the two reduced-form parameters α_1 and α_2 depend uniquely on the structural-form parameters $\lambda(t, 2 \cdot h)$ and $\lambda(t, h)$. Hence, holding t , $t + h$ and $t + 2 \cdot h$ constant, α_1 and α_2 can be estimated with standard cross-sectional techniques. We implement this with $t_0 = 1970$, $t_1 = 1980$ and $t_2 = 1990$.²⁹

²⁸Indeed, when we perform the piecewise linear regression exercise described above we can never reject the hypothesis of linearity and invariance.

²⁹Estimation of equation (A1) also requires the use of instrumental variables. The reason is that the error term $e_j(t + 2 \cdot h)$ is - by construction - correlated with the regressor $y_j^R(t + h)$ (income in 1980). We propose to use the age, sex, race and education of the household head as instruments. Clearly in this case the correlation between instruments and explanatory variables is much stronger than in the previous subsection. Hence, this approach delivers much more tightly estimated parameters.

Unfortunately, the system of two equations in two unknowns that defines the structural-form parameters in terms of the reduced-form ones is linearly dependent, and cannot be solved for $\lambda_1(1970, 20)$ and $\lambda_1(1970, 10)$. Hence, we cannot apply the linearity and invariance test directly to these structural parameters. However, it is clear that if the structural-form parameters are the same for every individual, so are the reduced-form ones. Hence, a test of the linearity and invariance properties can be performed by estimating α_1 and α_2 in piecewise linear fashion, and asking whether these coefficients are stable across groups.

Table 3: Estimates of Equation (A1)

Sample	α_1	α_2	Observ.
all sample	1.268 (.061)	-.249 (.054)	1955
group 1	1.349 (.269)	-.418 (.469)	489
group 2	1.877 (.370)	-1.096 (.523)	486
group 3	1.612 (.276)	-.719 (.328)	493
group 4	1.183 (.124)	-.176 (.094)	487

Note: α_1 and α_2 are generalized instrumental variable estimates of equation (A1) (standard errors in parenthesis). Dependent variable is y_{90} . Explanatory variables are y_{80} and y_{70} . Instruments are y_{70} , sex, race, educ70 and age70. (See Table 1 for data definitions). Group 1 has families with income below the first quartile. Group 2 between first quartile and median. Group 3 between median and third quartile. Group 4 above third quartile.

Table 3 reports the estimation results for the sample as a whole (first block), and then for a piecewise linear regression over four income groups (second to fifth block): families with 1970 income below the first quartile, between the first quartile and the median, between the median and the third quartile, above the third quartile. The point estimates appear fairly stable and close to those for the whole sample. A possible exception is group 2, which, however, is also the one for which the estimates are least precise.

In any case, no systematic pattern of variation of the estimates among the groups is discernible.

The formal test of the hypothesis that α_1 and α_2 are equal across the four income groups is reported in Table 4. For each pair of income groups, we have computed a Wald-test for the hypothesis that the estimated coefficients are the same.³⁰ The table reports the corresponding p-values. Not once the null hypothesis of invariance is rejected, at any standard confidence level.

Table 4: Wald test for income groups (p-values)

.	0.3544	0.7062	0.8487
0.3544	.	0.8207	0.2047
0.7062	0.8207	.	0.2476
0.8487	0.2047	0.2476	.

Note: Element i,j is the p-value associated to the Wald-test statistic for the coefficients of income groups i and j .

How robust are these results to other subdivisions of the sample? When we perform the test on two groups (below and above the median in 1970) we cannot reject at any standard confidence interval. When we use three groups we cannot reject for any pairwise comparison at the 1% confidence level, but we reject at the 5% level for one of the three comparisons. Finally, when we use five groups we again can never reject our null of linearity and invariance, at standard confidence levels. For each of these subdivisions we have also performed a test of invariance over all groups simultaneously, as opposed to separate pairwise tests (the two approaches coincide when there are two groups only). The lowest p-value obtains in the case of three groups, and it is 0.6876. The reader can draw his/her own conclusions.

³⁰The test statistic is $[\alpha^i - \alpha^j]'Q^{-1}[\alpha^i - \alpha^j]$, where $\alpha^z = [\alpha_1 \ \alpha_2]'$ as estimated for income group z and Q is an estimate of the variance covariance matrix of $\alpha^i - \alpha^j$. This test has the advantage that it does not require homoscedasticity across income groups (hence it is better suited to our purposes than a Chow test). Notice, however, that this test is known for over-rejecting. Hence, in selecting it we set the odds against our null.

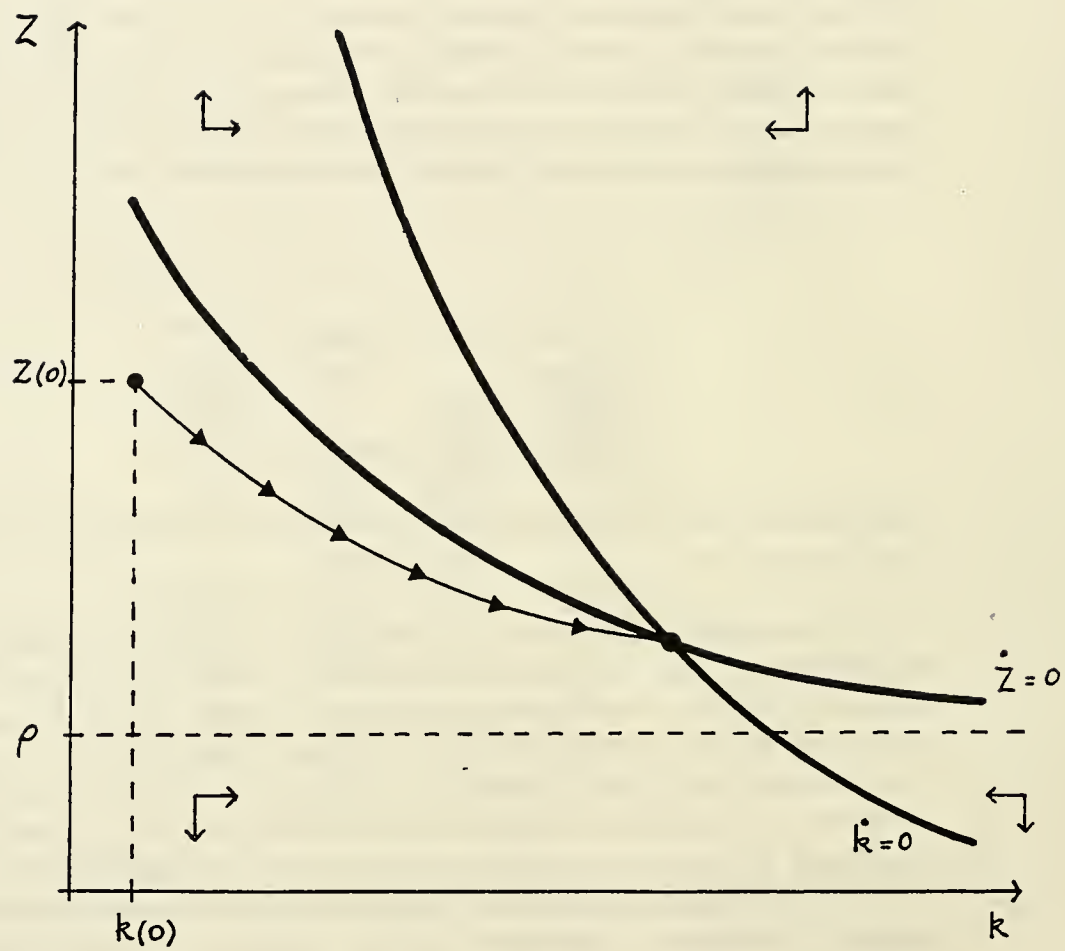


FIGURE 1

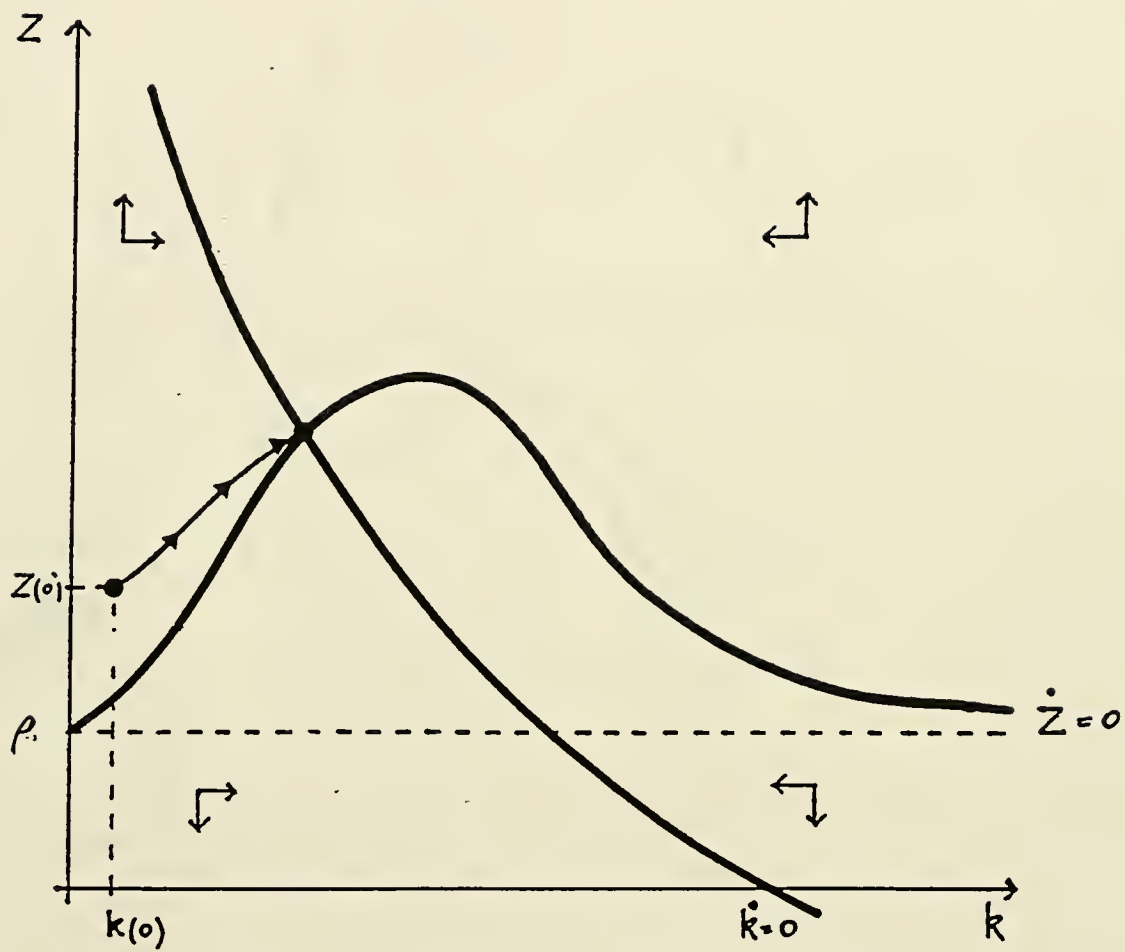


FIGURE 2

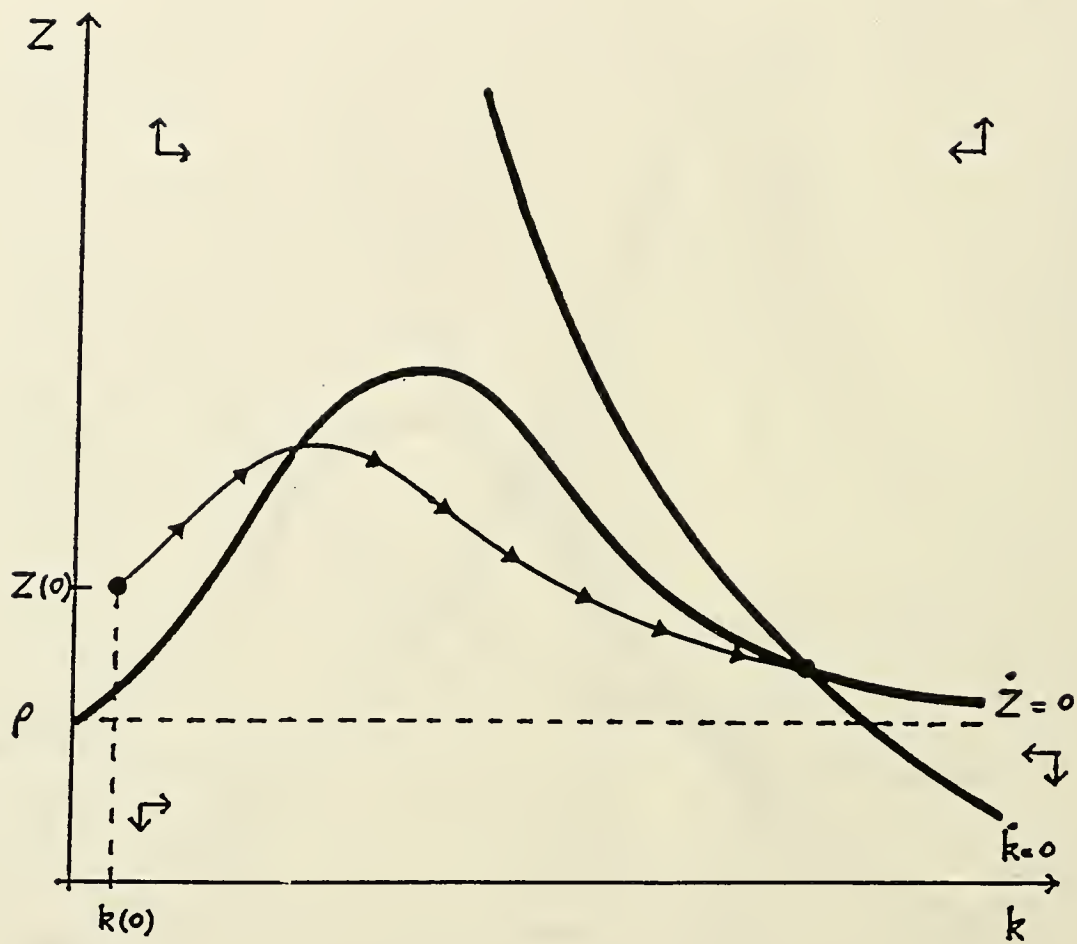
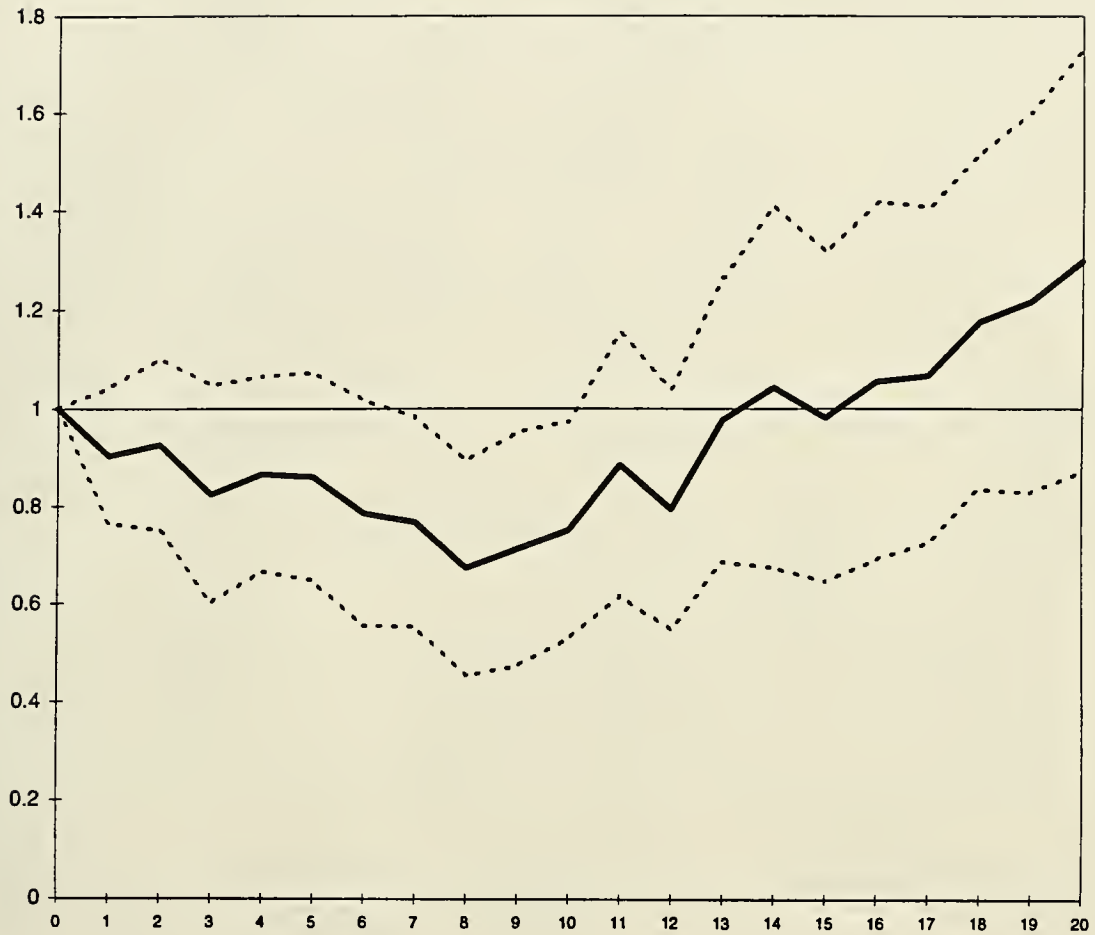


FIGURE 3

Figure 4: Conditional Convergence, 1970-1990



Y-axis: point estimates and 95% confidence bounds of LAMBDA(1970,h). X-axis: h.

Date Due

Lib-26-67

