A Representative Consumer Theory of Distribution

By FRANCESCO CASELLI AND JAUME VENTURA*

This paper introduces various sources of consumer heterogeneity in one-sector representative consumer (RC) growth models and develops tools to study the evolution of the distribution of consumptions, assets, and incomes. These tools are applied to the Ramsey-Cass-Koopmans model of optimal savings and the Arrow-Romer model of productive spillovers. The RC property per se places very few restrictions on the nature of observed distributions, and a wide range of distributive dynamics and income mobility patterns can arise as the equilibrium outcome. An example illustrates how to use these tools to generate quantitative predictions and compare them to the data. (*JEL* E13, O41)

A vast literature uses representative consumer (RC) growth models to characterize the dynamic behavior of aggregate quantities, such as per capita income and consumption. Very rarely, however, does this literature address the distributive properties of these models. In this paper, we show how to introduce various sources of consumer heterogeneity in one-sector RC growth models, and develop a set of tools to study the evolution of the distribution of consumptions, assets, and incomes.

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 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.[1] It is difficult to overemphasize the analytical convenience of this 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. Since keeping track of averages is always simpler than following distribution functions this is a major source of analytical tractability.[2]

The tools we provide allow one to make explicit the distributive dynamics that are "hidden" behind the RC assumption. The RC

* Caselli: Department of Economics, Harvard University, Cambridge, MA 02138, Centre for Economic Policy Research, and National Bureau of Economic Research. Ventura: Department of Economics, Massachusetts Institute of Technology, Cambridge, MA 02139, Centre for Economic Policy Research, and National Bureau of Economic Research. We are grateful to Daron Acemoglu, Robert Barro, Alberto Bisin, Gary Chamberlain, John Coleman, Bill Easterly, Gerardo Esquivel, Aart Kraay, John Leahy, Greg Mankiw, Jonathan Morduch, Steve Pischke, Antonio Rangel, Erik Thorbecke, and two anonymous referees for their useful comments. Caselli thanks the William Ladany Faculty Research Fund and the Graduate School of Business for financial support.

[1] For instance, if one is 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). If one is 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. See Andreu Mas-Colell et al. (1995 Ch. 4) for a review of aggregation results, including a discussion of the concept of a representative consumer.

[2] 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 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 therefore satisfy all the properties of individual demand functions. These properties allow us to derive testable implications from theoretical models.

growth models that we study feature three sources of consumer heterogeneity: tastes, skills, and initial wealth. We examine the behavior of consumers' relative consumption, wealth, and income—that is, 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., tastes, skills, and initial wealth); (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.

Two features of the dynamic evolution of the distribution of consumption, assets, and income play a prominent role in our analysis. The first is the tendency for the distribution of consumptions, assets, or incomes to become more (or less) concentrated around the mean within groups of individuals who have the same skills and tastes. We call this property conditional convergence (or divergence). When all consumers in the population share identical skills and tastes, conditional convergence coincides with a decline in overall inequality. We provide conditions under which the distribution of consumptions, assets, or incomes conditionally converges (diverges) as the economy grows. The second is income mobility. We allow for multiple sources of heterogeneity, and the relative weights of initial assets, skills, and tastes in determining the relative position of an individual in the income distribution change over time. This ensures that any two individuals can, in principle, switch relative ranking in the distribution multiple times over a given period.

To illustrate our algorithm we study in some detail the distributive properties of two popular RC growth models: the Ramsey-Cass-Koopmans model of optimal savings and the Arrow-Romer model of productive spillovers. These examples reveal that the RC property per se places very few restrictions on the nature of observed distributions. In particular, depending on taste and technology parameters, a wide range of distributive dynamics and income mobility patterns can arise as the equilibrium outcome of these models. These results imply that one can accommodate rich distributional dynamics without giving up the analytical convenience of the RC assumption. In another illustration, we show how one could extract quantitative predictions on distributional dynamics from a specific model. In particular, we calibrate a simple RC model of U.S. growth for the period 1970–1990, and show the steps one would take to compare its distributional implications with some aspects of the data.

Our paper is related to Joseph E. Stiglitz (1969) and Satyajit Chatterjee (1994), who study the evolution of the distribution of wealth in heterogeneous agent versions of the Solow and Ramsey models. Our work differs from theirs in three dimensions. First, we introduce two additional sources of consumer heterogeneity: tastes and skills. Without these additional sources of heterogeneity, the RC model is not capable of generating income mobility. Second, our results are more general, since they apply to any one-sector 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. The paper is also related to Giuseppe Bertola (1993), Alberto Alesina and Dani Rodrik (1994), and Torsten Persson and Guido Tabellini (1994), who develop heterogeneous agent models in which, absent policy choice, the RC property holds; the aim of these papers, however, is to characterize the link between income distribution and the choice of policies that affect growth. Once income distribution affects policy choice and the latter affects aggregate behavior, the RC property is violated.

That the RC model can be used to study many distributive issues does not mean that all the interesting questions concerning the distribution of income can be addressed within this framework. For example, there have recently been important papers addressing the question of how alternative initial income distribution profiles map into subsequent income levels and growth rates in the presence of market imperfections. This literature is summarized in Roland Bénabou (1996). Again, because the objective is to explore the channels through which heterogeneity in income or wealth affects aggregate behavior, models used in this line of research do not feature the RC property.

I. Heterogeneity in a RC Model

This section describes an economy with many infinitely lived consumers, indexed by $j = 1, 2, \ldots, J$. Throughout, we assume that J is large and that each consumer is "small" in the sense that his/her choices have negligible effects on aggregate quantities and prices. We allow consumer heterogeneity in tastes, skills, and initial wealth. However, we structure this heterogeneity in such a way that the RC result still applies.

Let $c_j(t) \in \mathfrak{R}_+$ denote the after-tax private consumption of agent j at date t, and $\mathbf{c}(t) = (c_1(t), \ldots, c_J(t)) \in \mathfrak{R}^J_+$ be the economy's vector of consumptions. To simplify notation, we omit the time indexes when this is not confusing and define average variables by omitting the consumer index. For instance, average consumption is defined as $c = (1/J) \sum_j c_j$. The value of any consumption sequence is evaluated by consumer j as follows:

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

where $\beta_j \in \mathfrak{R}_+$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J) \in \mathfrak{R}^J_+$, and $g(t) \in \mathfrak{R}_+$. We interpret $\beta_j g$ as the value of the publicly provided goods that the consumer receives in terms of the private consumption good.[3] The first source of consumer heterogeneity is variation in the value attached to publicly provided goods, as measured by β_j. A high value of β_j indicates that the consumer specially likes publicly provided goods or, perhaps, that the consumer receives a relatively large amount of these goods. Without loss of generality, we normalize the units of the publicly provided good so that $(1/J) \sum_j \beta_j = 1$ and refer to g as the average consumption of publicly provided goods. For later reference, we define x as the growth rate of g.

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 $a_j(t) \in \mathfrak{R}_+$ be the stock of financial assets of consumer j at date t, and let $\mathbf{a}(t) = (a_1(t), \ldots, a_J(t)) \in \mathfrak{R}^J_+$. A third source of consumer heterogeneity is variation in skills. We define $\pi_j \in \mathfrak{R}_+$ as the skill level of agent j, and let $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_J) \in \mathfrak{R}^J_+$. Without loss of generality, we adopt the normalization $(1/J) \sum_j \pi_j = 1$. Let $r(t) \in \mathfrak{R}_+$ and $w(t) \in \mathfrak{R}_+$ denote the after-tax rate of return on assets and the average after-tax wage rate, at date t. Also, define $p \in (0, 1]$ as the ratio between after-tax and before-tax consumption, i.e., $p(t)$ is one minus the proportional consumption tax. For later reference, we define ϕ as the growth rate of p. We assume that r, w, and p 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

$$(2) \qquad \dot{a}_j = r a_j + w \pi_j - c_j / p,$$

stating that increases in financial wealth equal the excess of income over consumption. Finally, we rule out Ponzi games.

The consumer's problem of maximizing the objective function (1) subject to the flow budget constraint (2) implies the following Euler equation:

$$(3) \quad \dot{c}_j = \frac{r - \rho + \phi}{\theta} (c_j + \beta_j g) - \beta_j \dot{g},$$

and integrating forward, we find

[3] At the cost of complicating the notation somewhat the model could easily accommodate the case in which the β_j are allowed to be negative.

$$(4)\quad c_j + \beta_j g = \mu\left[a_j + \int_t^\infty \left(\pi_j w + \beta_j g/p\right)\exp(-\int_t^\tau r\,dv)\,d\tau\right],$$

where μ is defined as

$$(5)\quad \mu^{-1} = \int_t^\infty p^{-1}\exp\left\{\int_t^\tau \frac{1}{\theta}[r(1-\theta) - \rho + \phi]\,dv\right\}d\tau.$$

Equation (4) states that total consumption is a fraction μ of total wealth.[4] Total consumption is defined as the sum of private consumption c_j, and the consumption of publicly provided goods $\beta_j g$. Total wealth is the sum of financial wealth a_j, the net present value of wages, $\int_t^\infty \pi_j w \exp(-\int_t^\tau r\,dv)\,d\tau$, and the net present value of publicly provided goods, $\int_t^\infty \beta_j p^{-1} g \exp(-\int_t^\tau r\,dv)\,d\tau$. Since μ is independent of the characteristics of the consumer, equation (4) is linear in a_j, π_j, and β_j. This property of the consumption function, combined with the linearity of the flow budget constraint, is crucial in allowing the nice aggregation properties of this model. Despite this, the model contains sufficient sources of nonhomotheticity to generate interesting distributional dynamics.

A. *The Representative Consumer Result*

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, skills, and taste parameters. Averaging (2) and (4) over all consumers, we find that

$$(6)\qquad \dot{a} = ra + w - c/p$$

and

$$(7)\quad c + g = \mu\left[a + \int_t^\infty \left(w + g/p\right)\exp(-\int_t^\tau r\,dv)\,d\tau\right].$$

These equations are those of a consumer with average characteristics, i.e., $a_j = a$ and $\pi_j = \beta_j = 1$. 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 interested only 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. Even if one is interested in the behavior of whole distributions, it is still useful to solve the model first *as if* all consumers were identical, and characterize the behavior of average quantities and prices. Once this has been achieved, it is then possible to characterize the behavior of whole distributions based on the behavior of average quantities, prices, and the distribution of individual characteristics.

Note that equations (3)–(5) imply that all consumers choose identical rates of growth of total consumption and total wealth, so that the cross sections (to be defined below) of total consumption and total wealth are unchanging. This is a very strong implication of the representative consumer assumption. If we had data on βs and πs (and, of course, a set of common expectations for factor prices and public policy), we could construct total consumption and total wealth and test this restriction, for a cross section of consumers. The representative consumer assumption, however, puts very few

[4] A consumer with the preferences defined by (2) might choose a negative consumption in some periods. Since we cannot find any meaningful economic interpretation for a negative consumption, 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 and taxes (i.e., r, w, and p).

restrictions on the distributional dynamics of the separate subcomponents of total consumption and total wealth. We next show this.

B. *The Cross Section of Consumptions*

Define $c_j^R = c_j/c$ and let $\mathbf{c}^R(t) = (c_1^R(t), \ldots, c_J^R(t))$ be the vector of relative consumptions. We refer to the vector $\mathbf{c}^R$ as the cross section of consumptions. Differentiating c_j^R with respect to time (use the Euler equation), we find that the jth element of $\mathbf{c}^R$ has the following law of motion:

$$(8)\quad \dot{c}_j^R = \left(\frac{r-\rho+\phi}{\theta} - x\right)\frac{g}{c}(\beta_j - c_j^R)$$

or, alternatively,

$$(9)\quad c_j^R(t) = c_j^R(0)\exp\left[-\int_0^t \left(\frac{r-\rho+\phi}{\theta} - x\right)\frac{g}{c}\,d\tau\right] + \beta_j\left\{1 - \exp\left[-\int_0^t \left(\frac{r-\rho+\phi}{\theta} - x\right)\frac{g}{c}\,d\tau\right]\right\}.$$

Equation (8) or its integral version (9) completely characterize the behavior of the cross section of consumptions, given the initial cross section, as a function of average quantities, prices, and the distribution of the taste parameters $\boldsymbol{\beta}$.[5] These equations make clear the role of publicly provided goods in generating cross-sectional dynamics in consumption. If these goods are absent, i.e., $g = 0$, the cross section of consumptions does not change, i.e., $\dot{c}_j^R(t) = 0$ or $c_j^R(t) = c_j^R(0)$ for all t. The intuition for this result is well known and follows from the homotheticity of isoelastic preferences.[6]

To understand the dynamics of a cross section of consumers when there are publicly provided goods, it is useful to assume first that there is no heterogeneity in tastes, i.e., $\beta_j = 1$ for all j. Visual inspection of equations (7) and (8) reveals that periods of convergence (divergence) in the cross section of consumptions are associated with low (high) growth in the provision of publicly provided goods, i.e., $(r - \rho + \phi)/\theta > x$ $[(r - \rho + \phi)/\theta < x]$.[7]

The intuition for this result is as follows: Since preferences are homothetic in total consumption, all consumers choose the same positive growth rate for it, i.e., $(r - \rho + \phi)/\theta$. This growth rate is a weighted average of the growth rates of its two components, private goods and publicly provided goods. The growth rate of the latter is the same for all consumers. If this growth rate is low, private consumption must be growing at a rate that exceeds that of total consumption. But how much faster? It depends on how large is the share of private consumption in total consumption. The poorer the consumer, the lower this share and the higher the rate of growth of private consumption required to sustain the optimal consumption path. This is why poor consumers improve their relative position in the cross section when $(r - \rho + \phi)/\theta > x$. A symmetric argument works for the case in which $(r - \rho + \phi)/\theta > x$.

A similar style of reasoning helps understanding the role of ϕ. When ϕ is high consumption taxes are falling rapidly, and therefore total consumption grows rapidly. Because everybody's consumption of the public good grows at the same rate, the private consumption of relatively poor individuals must grow especially fast. Since it is possible that $(r - \rho + \phi)/\theta - x$ changes signs an arbitrarily large number of times, the model is consistent with alternating periods of convergence and divergence in a cross section of consumptions.

[5] Using equations (4) and (7) it is immediate that relative consumption at time 0, $c_j^R(0)$, is in turn a linear combination of the three individual characteristics (i.e., initial wealth, skills, and tastes).

[6] If g becomes trivial asymptotically—i.e., if it grows at a rate slower than c—the cross section of consumptions approaches a constant.

[7] We say that there is convergence (divergence) in the cross section of consumptions during the interval $[0, T]$ if and only if $|c_j^R(t) - 1|$ is decreasing (increasing) for all $t \in [0, T]$.

It should be clear from this intuition that whether a consumer exhibits a growth rate of consumption above average or not depends on his/her share of private consumption in total consumption. If we now allow for heterogeneity in the taste for publicly provided goods, this share will be larger, the larger the value attached to the latter. A rich consumer that places a high value to publicly provided goods could still have a low share of private consumption, whereas a poor consumer could actually have a high share if he/she does not value much publicly provided goods. In fact, two consumers have the same share if and only if they have the same ratio c_j/β_j. This means that if $(r - \rho + \phi)/\theta > x$ $[(r - \rho + \phi)/\theta < x]$ consumers with a low ratio c_j/β_j will improve (worsen) their position in the cross section of consumptions. We refer to this result as conditional convergence (divergence) in the cross section of consumption since, if we hold constant differences in tastes, poor consumers tend to improve (worsen) their relative position.

In summary, with no heterogeneity in tastes, the relative magnitude of $(r - \rho + \phi)/\theta$ and x determines whether there is absolute convergence or divergence. When there is heterogeneity in tastes the same condition determines whether there is conditional convergence or divergence. Whereas the existence of convergence (divergence) implies that the dispersion in a cross section of consumptions declines (increases) over time, this does not follow from conditional convergence (divergence). The latter concept implies that the cross section of consumptions approaches (moves away from) the distribution of the β_js. If the taste parameters are more dispersed than the cross section of consumptions, conditional convergence (divergence) will be associated with increases (decreases) in the dispersion of consumptions.[8]

[8] For instance, Angus Deaton and Christina Paxson (1994) use cohort data to document absolute divergence of within-age-group consumption in the United States, the United Kingdom, and Taiwan. In a world in which all individuals in one generation have the same tastes, this finding is consistent with equation (8) and (9) if $x > (r - \rho + \phi)/\theta$. With heterogeneity in tastes it could be consistent both with conditional convergence and divergence in consumption.

Finally, note that this model can generate substantial mobility in a cross section of consumptions. In periods of conditional convergence (divergence) consumers that place a high value on publicly provided goods will overtake (be overtaken by) those consumers that place a lower value on those goods. If there are alternating periods of convergence and divergence, we might observe the same two consumers overtaking each other many times. Of course, there is no mobility in total (i.e., including publicly provided) consumption.

C. *The Cross Section of Wealth*

Define $a_j^R = a_j/a$ and let $\mathbf{a}^R(t) = (a_1^R(t), \dots, a_J^R(t))$ be the vector of relative asset holdings or the cross section of wealth. Differentiating a_j^R with respect to time [use equations (2)–(4) and (6)–(7)], one can show that the jth element of $\mathbf{a}^R$ has the following law of motion:

$$\text{(10)} \quad \dot{a}_j^R(t) = \delta_\pi(t)(\pi_j - a_j^R(t)) + \delta_\beta(t)(\beta_j - a_j^R(t)),$$

where the functions δ_π and δ_β are defined as follows:

$$\text{(11)} \quad \delta_\pi(t) = \frac{w}{a} - \frac{\mu}{ap}\int_t^\infty w \exp\left(-\int_t^\tau r\, dv\right) d\tau$$

$$\text{(12)} \quad \delta_\beta(t) = \frac{g}{ap} - \frac{\mu}{ap}\int_t^\infty p^{-1} g \exp\left(-\int_t^\tau r\, dv\right) d\tau$$

and have a clear economic interpretation (remember that π_j and β_j have been normalized to one). The function δ_π measures the net savings out of labor income (as a share of financial assets) of the average consumer. One can also interpret $-\delta_\pi$ as a measure of how fast wages grow, since it compares the current wage with the net present value of future wages. Exploiting the formal analogy and stretching somewhat the concept of savings, δ_β measures the net savings (as a share of financial assets) out of publicly provided goods. Naturally, $-\delta_\beta$ can also be interpreted as a measure of how fast the value of publicly provided goods grows since it compares their current consumption with the net present value of their future consumption. We can integrate (10) to find that

$$a_j^R(t) = a_j^R(0)\exp\left[-\int_0^t (\delta_\pi + \delta_\beta)\, d\tau\right] + \pi_j \int_0^t \delta_\pi \exp\left[\int_\tau^t (\delta_\pi + \delta_\beta)\, dv\right] d\tau + \beta_j \int_0^t \delta_\beta \exp\left[\int_\tau^t (\delta_\pi + \delta_\beta)\, dv\right] d\tau. \tag{13}$$

Equations (10) and (13) provide a complete characterization of the cross section of wealth as a function of average quantities and prices and the distribution of individual characteristics, $\mathbf{a}_j^R(0)$, $\boldsymbol{\pi}_j$, and $\boldsymbol{\beta}$. These equations show that, unlike the case for the cross section of consumptions, the existence of publicly provided goods is not the only force behind the dynamics of a cross section of wealth. Even if these goods did not exist (i.e., $g = 0$), the model generates nontrivial dynamics for the distribution of assets.

Consider first the case in which there is no cross-sectional variation in tastes and skills, i.e., $\pi_j = \beta_j = 1$ for all j. Also, assume that the economy exhibits average asset growth, i.e., $r > \mu^{-1}$. Equations (10) and (13) show that there is convergence (divergence) in a cross section of wealth if and only if $\delta_\pi + \delta_\beta > 0$ ($\delta_\pi + \delta_\beta < 0$).[9] The intuition for this result is as follows: consumers calculate their total wealth and choose the optimal path for their total consumption. Since all consumers have the same spending shares (a property of homothetic preferences), they all spend the same fraction of their total wealth in each date and therefore exhibit identical rates of total wealth accumulation. But the growth rate of total wealth is a weighted average of the growth rates of its three components, the stock of financial assets, the net present value of wages, and the net present value of publicly provided goods. The growth rate of the last two pieces is the same for all consumers. If this growth rate is low, consumers must be accumulating financial assets at a rate that exceeds that of total wealth. But how much faster? It depends on how large is the share of financial assets in a consumer's total wealth. The lower this share, the higher the rate of asset accumulation required to sustain the optimal consumption path. This is why poor consumers improve their relative position in the cross section when $\delta_\pi + \delta_\beta > 0$. A symmetric argument works for the case in which $\delta_\pi + \delta_\beta < 0$. Since it is possible that $\delta_\pi + \delta_\beta - 1$ changes signs an arbitrarily large number of times as a grows, the model is consistent with alternating periods of convergence and divergence in a cross section of wealth. Interestingly, these periods need not (but, of course, might) coincide with alternating periods of convergence and divergence in a cross section of consumptions.

Whether a consumer exhibits a growth rate of financial wealth above average or not depends on his/her share of financial wealth in total wealth. If we allow for heterogeneity in skills, this share will be lower the higher the skills of the worker. If we allow for heterogeneity in tastes, this share will be lower the larger the value that the consumer attaches to publicly provided goods. This is why, in an economy with low (high) average growth in

[9] We say that there is convergence (divergence) in the cross section of wealth during the interval $[0, T]$ if and only if $|a_j^R(t) - 1|$ is decreasing (increasing) for all $t \in [0, T]$.

both wages and publicly provided goods, consumers with low ratios a_j/π_j and a_j/β_j tend to improve (worsen) their relative position in the cross section. We refer to this result as conditional convergence (divergence) in the cross section of wealth since, if we hold constant differences in tastes and skills, poor consumers tend to improve (worsen) their relative position. As in the case of consumption, then, the condition guaranteeing absolute convergence (divergence) absent heterogeneity in skills and tastes is the same that guarantees conditional convergence (divergence) when there is heterogeneity in skills and/or tastes. We repeat the caveat that conditional convergence need not be associated with reductions in the cross-sectional dispersion of wealth holdings.

Note also that this model can generate substantial mobility in a cross section of wealth. Consider, for instance, an economy with positive growth in financial assets. In periods where wages grow slowly (fast), consumers with high skills will overtake (be overtaken by) consumers with low skills. Similarly, in periods in which publicly provided goods grow slowly (fast) those consumers who value these goods the most will overtake (be overtaken by) those consumers who value these goods less. Even if there are no reversals in the signs of δ_π, δ_β, or their sum, we might observe the same two consumers overtaking each other many times.

D. *The Cross Section of Incomes*

Since available cross-sectional data are mainly on personal income, it is interesting to determine the model's implications for the cross section of incomes, $y_j = ra_j + w\pi_j$. As usual, define $y_j^R = y_j/y$ and let $\mathbf{y}^R(t) = (y_1^R(t), \dots, y_J^R(t)) \in \Re_+^J$ be the vector of relative incomes. A simple algebraic manipulation shows that $y_j^R = \alpha a_j^R + (1-\alpha)\pi_j$, where α is the share of capital income in total income, i.e., $\alpha = ra/y$. It follows from equation (13) that[10]

$$\begin{aligned}(14)\quad & y_j^R(t) \\ &= y_j^R(0)\,\frac{\alpha(t)}{\alpha(0)}\exp\left[-\int_0^t (\delta_\pi+\delta_\beta)\,d\tau\right] \\ &+ \beta_j\alpha(t)\int_0^t \delta_\beta \exp\left[\int_\tau^t (\delta_\pi+\delta_\beta)\,dv\right] \\ &\times d\tau \\ &+ \pi_j\Bigg((1-\alpha(t))\times\Bigg\{1-\frac{\alpha(0)^{-1}-1}{\alpha(t)^{-1}-1} \\ &\qquad\times\exp\left[-\int_0^t(\delta_\pi+\delta_\beta)\,d\tau\right]\Bigg\} \\ &+ \alpha(t)\int_0^t \delta_\pi \\ &\qquad\times\exp\left[\int_\tau^t(\delta_\pi+\delta_\beta)\,dv\right]d\tau\Bigg).\end{aligned}$$

Although somewhat intimidating in appearance, equation (14) is easy to interpret based on the intuition developed for the analysis of the dynamics of a cross section of wealth. In particular, if one is willing to assume that the aggregate share of capital in income does not change much over the relevant period, conditional convergence (divergence) in a cross section of incomes occurs under the same exact conditions as we found for a cross section of wealth. If the share of capital increases (decreases) over time, the conditions for conditional convergence (divergence) in a cross section of incomes are more (less) stringent than those required for conditional convergence (divergence) in a cross section of wealth.

II. Application 1: Two Popular Models

In this section we apply our tools to derive distributive predictions for the transitional dynamics of the Ramsey-Cass-Koopmans model

[10] To see this substitute from equation (13) in the expression for $y_j^R(t)$. Use $y_j^R(0) = \alpha a_j^R(0) + (1-\alpha)\pi_j$ to substitute $a_j^R(0)$ out of the resulting expression.

of optimal savings and the Arrow-Romer model of productive spillovers. These examples will show that the RC assumption places few restrictions on the behavior of the distributions of private consumption and assets. This implies that the observation of complex distributive dynamics does not force us to forgo the analytical convenience of the RC assumption.

Assume there is a large number of identical competitive firms with free access to a technology $f : \mathcal{R}_+^2 \to \mathcal{R}_+$, that maps average stock of capital their workers use and their average skills to production per worker. Let k be the average stock of capital in the economy and A be a measure of average labor productivity. We assume that $f(\cdot, \cdot)$ is homogeneous of degree one in k and A, and that $f_k(k, A) > 0$, $f_{kk}(k, A) < 0$, and $\lim_{k\to\infty} f_k(k, A) = 0$. 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, before-tax factor prices are given by the marginal products of the relevant factor, i.e., $f_k(k, A)$ and $f(k, A) - kf_k(k, A)$. We consider an autarchic economy and, as a result, the average (net) stock of assets must equal the capital stock, $a = k$. With respect to labor productivity, we explore two alternative specifications in what follows.

For simplicity, we assume that all public consumption is financed through the private-consumption tax, so that $v = f_k(k, A)$ and $w = f(k, A) - kf_k(k, A)$. Then the revenue obtained from taxation is $c(1 - p)/p$. (Remember that c is after-tax consumption.) Finally, we impose a balanced-budget constraint and, as a result, publicly provided goods are an increasing function of contemporaneous revenue, i.e., $g = g(c(1 - p)/p)$ and $g'(c(1 - p)/p) \geq 0$.

To simplify the analysis, we assume throughout that $\theta = 1$ and $\phi = 0$, i.e., utility is logarithmic and the proportional consumption tax is constant. This is a very convenient parameter restriction because it implies that $\delta_\pi(t) + \delta_\beta(t) = -(\dot{z}/z)$, where $z \equiv (c + g)/k$.[11] 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 substitute r and w in equations (5) and (6) to obtain the economy's dynamical system in (c, k) space and then rewrite this dynamical system in (z, k) space as follows:

(15)
$$\dot{z} = \left(f_k(k, A) - \rho - \frac{f(k, A) - \varphi(zk)/p}{k}\right) z$$

and

(16) $$\dot{k} = f(k, A) - \varphi(zk)/p,$$

where the function $\varphi(\cdot)$ is obtained by inverting the definition of z, and has the following properties: $\varphi(zk) = c \geq 0$ and $\varphi'(zk) = p/[p + (1 - p)g'] \geq 0$.[12] This system is subject to two boundary conditions, the initial condition for k and a transversality condition for the aggregate economy that is implied by those of each consumer.

A. *The Ramsey-Cass-Koopmans Model*

In the model developed by Frank Ramsey (1928), David Cass (1965), and Tjalling Koopmans (1965), labor productivity is exogenously determined. Under the assumption that A is constant, the dynamical system (14)–(15) has a single steady state that is saddlepath stable.[13]

[11] To see this, note that with $\theta = 1$ and $\phi = 0$, $\mu = \rho p$, and

$$\frac{\dot{z}}{z} = \frac{\dot{c} + \dot{g}}{c + g} - \frac{\dot{k}}{k}$$

$$= \frac{1}{k}\left[\frac{\mu}{p}\int_t^\infty (w + g/p)\exp\left(-\int_t^\tau r\, dv\right) d\tau - w - g/p\right] = -\delta_\pi(t) - \delta_\beta(t)\,.$$

[12] More precisely, the function $\varphi(\cdot)$ is implicitly defined as follows: $\varphi(\zeta) + g\{[(1 - p)/p]\zeta\} = \zeta \;\forall\; \zeta$.

[13] It is often assumed that average labor productivity grows at a constant and strictly positive rate. It is possible to analyze the effects of this assumption in our framework, albeit at the cost of some interesting complications that arise if $g(c(1 - p)/p) \neq 0$ and $g(c(1 - p)/p) \neq \lambda c$, where λ

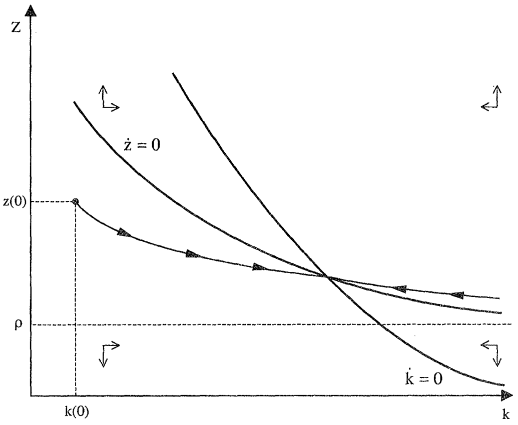

FIGURE 1. COBB-DOUGLAS TECHNOLOGY, $P = 1$

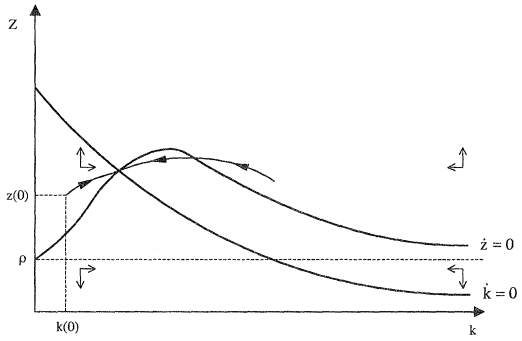

FIGURE 2. CES TECHNOLOGY, LARGE ρ, $P = 1$

The steady-state values of z and k are implicitly defined by $f_k(k^*, A) = \rho$ and $pf(k^*, A) = \varphi(z^*k^*)$. If the economy's initial capital is k^*, we have that the cross sections of both consumptions and wealth do not vary over time.[14] If the economy starts from below the steady state, we know that k monotonically approaches its steady-state value. However, the behavior of z depends crucially on the properties of $f(\cdot, \cdot)$. To show this, we consider six examples.

In the first three examples, we assume that there is no government (i.e., $p = 1$), which implies that $\varphi(zk) = zk$. We know already that this means that there are no dynamics in the cross section of consumptions.[15] It is still possible, however, to observe interesting dynamics in the cross section of wealth. Figure 1 shows the phase diagram of (14)–(15) for the case of Cobb-Douglas technology: $f(k) = k^\alpha A^{1-\alpha}$. As the economy travels along the stable arm, z declines monotonically if the initial capital stock is below the steady-state value, and it increases monotonically if the initial capital stock exceeds it. As a result, the Ramsey model with logarithmic utility and Cobb-Douglas technology predicts conditional convergence (divergence) during the transition from below (above) toward the steady state.

To caution the reader against drawing quick conclusions from the Cobb-Douglas example, consider this constant-elasticity-of-substitution (CES) technology: $f(k) = (k^b + A^b)^{(1/b)}$ with $b < 0$. The elasticity of substitution between capital and labor is less than 1, 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). In this economy, z is always increasing and, as a result, there is always conditional divergence, if the approach to the steady state is from below. Instead, if the transitional dynamics take place from above, and the initial stock of capital is sufficiently high, there is a phase of divergence followed by a period of convergence. 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 its conditional divergence, whereas at high levels of development it is characterized by conditional convergence. If $k(0)$ exceeds k^*, instead, there is monotonic divergence during the transition.

We consider next three examples in which

is a positive constant. We do not discuss these complications here since they would lead us too far afield.

[14] Regarding the cross section of consumptions, note that $r = \rho$. Regarding the cross section of wealth, note simply that (i) w and c are constant and (ii) $\mu = r = \rho$, which jointly imply that $\delta_\pi = \delta_\beta = 0$.

[15] Actually, if we assume that public goods are produced proportional to the tax revenue, i.e., $g(c(1 - p)/p) = \lambda c(1 - p)/p$, we find that $\varphi(zk) = \{p/[p + \lambda(1 - p)]\}zk$, and the results in this paragraph apply as well to this case. These two cases are formally identical because in both instances the cross section of consumption rates is constant, i.e., $r - \rho = x$.

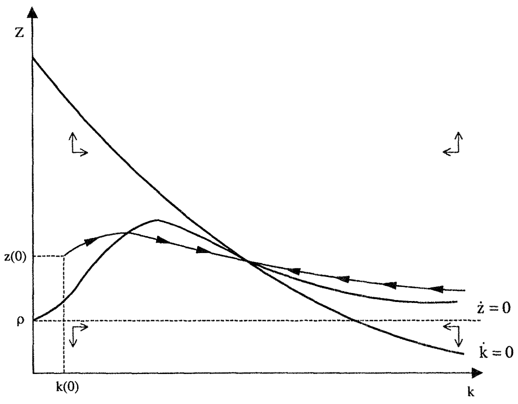

FIGURE 3. CES TECHNOLOGY, SMALL ρ, $P = 1$

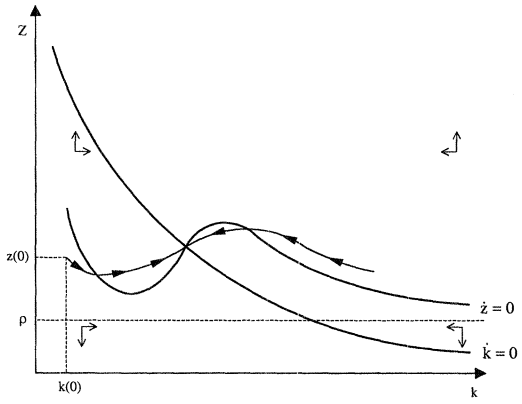

FIGURE 4. CES TECHNOLOGY, LARGE ρ, $P < 1$

$g(c(1 - p)/p) = c(1 - p)/p + \lambda$, where λ is a positive (but not too large) constant. This formulation embeds the idea that there are two types of publicly provided goods. The first term captures goods such as infrastructure, law enforcement, and the output of public enterprises, characterized by: (i) production functions that are similar to those of private goods and (ii) nonsatiation. The second term captures goods such as money, diplomacy, and regulations, characterized by: (i) close-to-zero costs of production and (ii) satiation. Under these assumptions, $\varphi(zk) = p(zk - \lambda)$ and $x < r - \rho$.[16] As a result, the cross section of consumptions exhibits conditional convergence. Once again, the dynamics of the cross section of wealth depend on the nature of technology. Figure 4 shows the CES economy again assuming a high ρ. If the economy starts at a low enough level of capital, z is initially decreasing and then increases. At low levels of development, the economy exhibits conditional convergence in wealth, whereas at high levels of development it is characterized by conditional divergence in wealth. Figure 5 presents the phase diagram for an economy with low ρ. From below, this economy exhibits two periods of convergence and one period of divergence. From above, there is always divergence. Finally, consider Figure 6, in which ρ is high enough that the steady state lies on the first downward-sloping portion of the $\dot{z} = 0$ locus. Here we have monotonic convergence from below, but convergence followed by divergence from above.

As the six examples here show, the dynamics of the cross sections of wealth in the Ramsey-Cass-Koopmans model depend crucially on the properties of the technology and preferences that one is willing to assume. Figures 1–6 show cases in which there is conditional convergence always, never, and sometimes. It is not difficult to generate conditional convergence and divergence cycles by appropriately choosing the production function.

The distributive dynamics in Figures 3 and 5 are 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 Simon Kuznets (1955) who first hypothesized its existence and gave some evidence in its support. Without taking a stand on whether the Kuznets curve is a fact or a myth, we simply note that our CES examples imply that the Ramsey-Cass-Koopmans model is consistent with it. To show this, we use the coefficient of variation of k_j as our measure of inequality: $CV^k = \sqrt{\text{Var}\{k_j^R\}}$; and assume that there is no cross-sectional variation in tastes and skills, i.e., $\pi_j = \beta_j = 1$ for all j. Simple algebra shows that

[16] To see this, note first that total consumption always grows at the rate $r - \rho$ and this rate is nothing but an average of the growth rates of private consumption and the consumption of publicly provided goods. But it follows from the postulated function $g(\cdot)$ that the growth rate of publicly provided goods is always lower than the growth rate of private consumption.

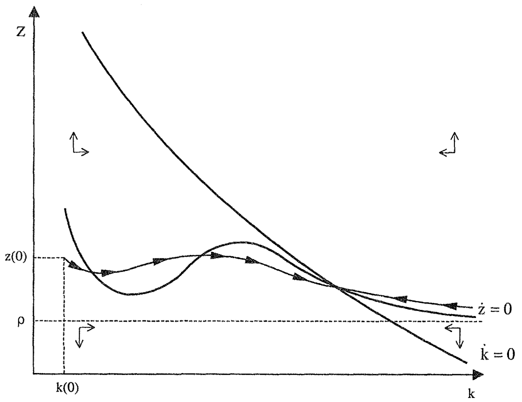

FIGURE 5. CES TECHNOLOGY, SMALL ρ, $P < 1$

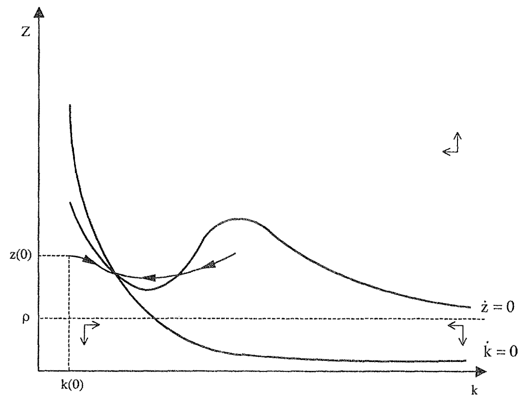

FIGURE 6. CES TECHNOLOGY, VERY LARGE ρ, $P < 1$

$$(17)\quad CV^k(t) = CV^k(0)\exp\left[-\int_0^t (\delta_\pi + \delta_\beta)\, d\tau\right].$$

That is, inequality increases (decreases) in periods where z is increasing (decreasing), i.e., $\delta_\pi + \delta_\beta < 0$ ($\delta_\pi + \delta_\beta > 0$). It follows that, in the Ramsey economies depicted in Figures 3 and 5, 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.

Still assuming that there is no cross-sectional variation in tastes and skills, it is easy to see that the model is consistent with the existence of a Kuznets curve in the cross section of consumptions. The key assumption for this to happen is that the growth rate of publicly provided goods is high at low levels of income, i.e., $x > r - \rho + \phi$ if $c < \bar{c}$, and declines afterward, i.e., $x < r - \rho + \phi$ if $c > \bar{c}$. To see this note that, if we use again the coefficient of variation in c_j as our measure of inequality: $CV^c = \sqrt{\mathrm{Var}\{c_j^R\}}$, its dynamics are given by

$$(18)\quad CV^c(t) = CV^c(0) \times \exp\left[-\int_0^t \left(\frac{r-\rho+\phi}{\theta} - x\right)\frac{g}{c}\, d\tau\right].$$

We want to stress finally that, in this discussion of Kuznets curves in wealth/consumption, we have been able to equate increasing (decreasing) inequality to the concepts of convergence (divergence) because we have assumed that there is no dispersion in tastes and skills. If we assumed, for instance, that the distributions of π_j and β_j were more dispersed than the distributions of wealth and consumption, conditional convergence (divergence) would be associated with decreasing (increasing) inequality instead. Under this new assumption, the Kuznets curve in wealth arises in Figure 4 but not in Figure 3. Also the conditions for a Kuznets curve in consumption are reversed. Now the sufficient condition is that the growth rate of publicly provided goods is low at low levels of income, i.e., $x < r - \rho + \phi$ if $c < \bar{c}$, and increases afterward, i.e., $x > r - \rho + \phi$ if $c > \bar{c}$.

B. *The Arrow-Romer Model*

In the models of Kenneth Arrow (1962) and Paul Romer (1986), it is assumed that labor productivity positively depends on the aggregate stock of capital, i.e., $A(k)$ with $A'(k) > 0$. In the spirit of Romer (1986), we assume that $k/A(k)$ is bounded in such a way that the rate of return asymptotically remains above the rate of time preference, i.e., $\lim_{k\to\infty} f_k(k, A(k)) > \rho$.

Let's start again by considering an economy without government (i.e., $p = 1$), which implies that $\varphi(zk) = zk$. Remember that in this case there are no dynamics in the cross section of consumers. The canonical special case of the

Arrow-Romer model is that in which $A(k) = k$. In this case, the economy does not exhibit transitional dynamics but instead features a constant growth rate for k equal to $f_k(1, 1) - \rho$, whereas z is constant and equal to $f(1, 1) + \rho - f_k(1, 1)$. As a result, the cross section of wealth exhibits no dynamics either.[17]

Consider next the case in which $g(c(1 - p)/p) = c(1 - p)/p + \lambda$. As in the Ramsey-Cass-Koopmans model, this specification generates conditional convergence in the cross section of consumptions. But it also generates conditional convergence in the cross section of wealth. To see this, one can use the dynamical system (14)–(15) and the appropriate boundary conditions to show that $\dot{z}$ is always negative and that z asymptotically approaches $f(1, 1) + \rho - f_k(1, 1)$ from above. In this version of the Arrow-Romer model, the growth rate of k is increasing over time and asymptotically approaches $f_k(1, 1) - \rho$.

It should be apparent by now that, by choosing the "right" specification of $A(\cdot)$ and $g(\cdot)$, it is possible to generate many types of dynamics in the cross sections of private wealth and private consumption.

III. Application 2: The Productivity Slowdown

In Section I we have developed an algorithm to track distributional dynamics in RC models, and in Section II we have showed how one would apply this algorithm to characterize the qualitative behavior of cross sections of consumption and wealth in two specific RC models. In this section we show how one could extract quantitative predictions on distributional dynamics from a specific model. For the purposes of this illustration we write down a simple RC model of U.S. growth for the period 1970–1990, and look at how its distributional implications compare with some aspects of the data. We stress the fact that the purpose of this exercise is merely to illustrate the steps one would take in a quantitative application of our methodology, and not to test the model or to provide a convincing recount of the 1970–1990 period.[18]

Since our purpose is mainly illustrative, we chose to focus on a particularly simple model. We therefore assume that the production function is Cobb-Douglas, i.e., $f(k) = k^{\alpha}A^{1-\alpha}$, productivity grows at an exogenous rate γ, i.e., $\dot{A}/A = \gamma$; the consumption tax rate is exogenous and constant, i.e., $\phi = 0$; and the production function for the public good is linear in tax revenues: $g = c(1 - p)/p$. Under these assumptions, and with competitive factor and good markets, the economy's dynamic path is characterized by an initial condition on the capital stock, a transversality condition, and the differential equations

$$(19) \qquad \frac{d\hat{c}/dt}{\hat{c}} = \frac{1}{\theta}\left[\alpha\hat{k}^{\alpha-1} - \delta - \rho - \theta\gamma\right]$$

and

$$(20) \qquad d\hat{k}/dt = \hat{k}^{a} - \frac{\hat{c}}{p} - (\gamma + \delta)\hat{k},$$

where $\hat{c} = ce^{-\gamma t}$ ($\hat{k} = ke^{-\gamma t}$) is consumption (capital) per effective unit of labor. The steady state for this economy features constant levels of consumption and capital per effective worker. In steady state, consumption and capital per capita grow at the constant rate γ.

The foremost development over the 1970–1990 period is that, following several decades of roughly constant growth of total factor productivity, and per capita income, capital, and consumption, all these quantities started growing at considerably lower rates. Our stylized way of interpreting these facts is that the economy was on a steady-state growth path in 1970, with productivity growth γ_0, and that in 1970

[17] This property of the Arrow-Romer model was noted first by Bertola (1993).

[18] For example, our illustrative model does not take into account the role of skill-biased technical change, which, according to many authors, has played an important role in distributive dynamics since the 1970's. See, e.g., Lawrence Katz and Kevin M. Murphy (1992), Chinhui Juhn et al. (1993), and David Autor et al. (1998) for the facts and, e.g., Oded Galor and Daniel Tsiddon (1997), Jeremy Greenwood and Mehmet Yorukoglu (1997), Daron Acemoglu (1998), and Caselli (1999) for the theories.

the rate of productivity growth fell permanently to the level $\gamma_1 < \gamma_0$.[19] Using a phase diagram it is immediate to characterize the dynamic adjustment of the economy to this change. The economy converges to a new steady state with higher consumption and capital per effective worker: $\hat{c}^*(\gamma_1) > \hat{c}^*(\gamma_0)$, $\hat{k}^*(\gamma_1) > \hat{k}^*(\gamma_0)$. In the transition, consumption per effective worker initially jumps discretely so as to put the economy on the new saddlepath, and thereafter consumption and capital per effective worker grow monotonically toward the new steady state. Given choices for the parameters characterizing technology and preferences, standard numerical techniques allow us to generate time series for all the aggregate variables for all dates following 1970.

Let us now rewrite equation (14) as

$$(21)\quad y_j^R(1990) = y_j^R(1970)\lambda_y(1970, 1990) + \pi_j\lambda_\pi(1970, 1990) + \beta_j\lambda_\beta(1970, 1990).$$

where $\lambda_y(1970, 1990)$, $\lambda_\beta(1970, 1990)$, and $\lambda_\pi(1970, 1990)$ stand for the coefficients on y_j^R, π_j, and β_j.[20] Note that λ_y, λ_β, and λ_π are functions of the time series of wages, interest rates, and the capital stock following 1970, and can therefore be computed from a numerical simulation of the model. At the same time, these coefficients can, in principle, be estimated from actual data and compared with their theoretical counterparts.

Table 1 reports values of $\lambda_y(1970, 1990)$ obtained from numerical simulations of the simple model described earlier under alternative specifications of the technological parameter α and the preference parameter θ.[21] The table

TABLE 1—SIMULATED VALUES OF λ_y (1970, 1990)

	θ				
α	.5	1	3	5	10
0.3	0.817	0.951	2.7	8.12	70.2
0.4	0.828	0.954	2.07	3.78	10.1
0.5	0.838	0.96	1.64	2.16	3.42
0.6	0.851	0.969	1.33	1.46	1.79
0.7	0.878	0.983	1.12	1.14	1.22

Note: Simulated values of λ_y in equation (21) under different values for α and θ.

shows that λ_y is quite sensitive to variations in θ and, when θ is large, in α.

The relation between λ_y and θ is monotonically increasing, a fact that is essentially driven by the positive dependence of μ on θ [equation (5)], the negative dependence of $\delta_\pi + \delta_\beta$ on μ [equations (12) and (13)], and the negative dependence of λ_y and $\delta_\pi + \delta_\beta$ [equation (14)]. The intuition for this is transparent if one recalls our discussion in Section II, subsection C. The larger θ, the less willing consumers are to substitute intertemporally, and the slower the rate of total wealth accumulation. This slow rate of total wealth accumulation is identical across individuals, and so is the rate of growth of the labor income component of total wealth (recall that λ_y holds skills and tastes constant). Hence, slower rates of total wealth accumulation translate into slower rates of financial wealth accumulation, the more so the smaller the weight of financial wealth in total wealth. This disproportionately affects those individuals for whom labor income is a relatively large *fraction* of income. Therefore, holding the *amount* of labor and publicly provided consumption constant across individuals, a large θ allows rich individuals to make large gains relative to poor ones.

The dependence of λ_y on α is more complex—increasing at low levels of θ but decreasing for large θ. To understand this pattern notice that a large α implies a large interest rate, which clearly favors high-income individuals because for them financial wealth is a relatively large source of income. Hence, this effect leads to a positive relation between λ_y and α. On the other hand, a

[19] The more accurate statement of this assumption is that agents believed the shock to be permanent throughout the period.

[20] It is easy to check that $\lambda_y + \lambda_\beta + \lambda_\pi = 1$.

[21] The other parameters are as follows: $\delta = 0.05$ and $\rho = 0.02$ (from Robert J. Barro and Xavier Sala-i-Martin, 1995), $p = 0.9$ (from Andrew Abel and Ben Bernanke, 1998), $(1 - \alpha)\gamma_0 = 0.0193$, and $(1 - \alpha)\gamma_1 = 0.0043$ (from Robert Gordon, 1996). We view these numbers as fairly uncontroversial.

large α also implies that convergence to the steady state is slow (since the marginal productivity of capital declines slowly), with the consequence that wages grow fairly slowly during the transition. By an argument analogous to the one given previously, since everyone's total wealth grows at the same rate, slow growth in the labor income component of total wealth requires fast growth of financial wealth, the more so the smaller the contribution of financial wealth to total wealth. Hence, this effect leads to a negative relation between λ_y and α. As θ increases the second effect becomes more important, since a low elasticity of substitution further slows down convergence to the steady state.

Suppose we took seriously our illustrative model of U.S. growth: how could we assess which (if any) of the parameter configurations in Table 1 is consistent with the income distribution data? The simplest approach would be to compare values of λ_y obtained from numerical simulation of the model with the values obtained through an empirical estimate of equation (21). We can obtain this estimate of $\lambda_y(1970, 1990)$ by treating it as an exogenous parameter to be estimated with cross-sectional techniques. Specifically, we use data from the Panel Study of Income Dynamics (PSID) to run cross-family regressions of the form

$$(22) \quad y_j^R(1990) = \delta y_j^R(1970) + \gamma' \mathbf{X}_j + \varepsilon_j(1990),$$

where y_j^R is total family money income (divided by the population average), and includes all market incomes earned by all the members of the family, before taxes and transfers; $\mathbf{X}_j$ is a vector containing a constant and a number of individual characteristics of the head of family j (age, age squared, sex, race, and education); and $\varepsilon_j(1990)$ is a zero-mean stochastic error term.[22] Equation (22) is an empirical version of equation (21), in which we interpret $\gamma' \mathbf{X}_j$ as the linear predictor of the quantity $\lambda_\pi \pi_j + \lambda_\beta \beta_j$, given $\mathbf{X}_j$. This method allows us to interpret δ as an estimate of $\lambda_y(1970, 1990)$.[23]

The Appendix provides detailed information on our data and our estimation technique (which involves the use of instrumental variables). For our purposes, however, the key point is the empirical estimate of λ_y, which in our preferred specification is 1.276 with a standard error of 0.191. Hence, our estimates are extremely imprecise, and several of the simulated values of λ_y in Table 1 fall within the 95-percent confidence interval (which goes from 0.894 to 1.658). Nevertheless, it is apparent that values of θ much less than 1—conditional on our illustrative model being "right"—could not have generated the data. Furthermore, the larger the value of θ, the larger the value of α required to match the data. For example, if $\theta = 10$, the capital share must be in excess of 0.6.

IV. Conclusion

In this paper, we have studied the effects of introducing three sources of heterogeneity in a class of widely used RC growth models, and developed a set of tools to study the evolution of the distribution of consumptions, wealth, and incomes. We find that the RC property per se places few restrictions on the dynamics of observed distributions. For example, a wide range of distributive dynamics and income mobility patterns can arise as the equilibrium outcome of two popular models: the Ramsey-Cass-Koopmans model of optimal savings and the Arrow-Romer model of productive spillovers. This does not mean that we have an "almost anything goes" type of result. The theory links distributional dynamics to measurable variables, such as the evolution of factor prices, attitudes toward risk, labor productivity, taxes and government expenditure, and so on. This should eventually permit other researchers to calibrate models and perform quantitative analyses of distributional changes generated by

[22] Inspection of this data reveals that they are consistent with a wealth of recent empirical studies documenting an increase in income inequality throughout the 1970's and the 1980's. More details on how we put together our data are given in the Appendix.

[23] The advantage of this method is that it does not require estimates of or assumptions on the distribution of π_j and β_j. A more ambitious method would be to compare the whole distribution of $y_j^R(1990)$ as obtained from equation (21) with the actual one.

specific events such as a technology shock, a trade liberalization, or a change in fiscal policy. In fact, adopting the RC assumption substantially simplifies the task of deriving distributive predictions from any specific growth model. We used a simple version of the Ramsey-Cass-Koopmans model applied to U.S. data to illustrate the steps one would take to perform such an exercise.

There are a couple of interesting extensions to the Representative Consumer Theory of Distribution that could be developed. First, it would be interesting to determine how labor supply choices interact with consumption choices in shaping the dynamics of the cross sections of consumption, wealth, and income. This can be done by introducing in the theory subutility functions of the Gorman form in consumption and leisure. These preferences preserve the RC property but permit the study of nontrivial labor supply choices. Second, it would also be interesting to let luck play a role in shaping the dynamics of the cross sections of consumption, wealth, and income. One could reinterpret our model as one in which there is no aggregate risk and idiosyncratic risks are perfectly pooled across consumers. Under these conditions, only consumer heterogeneity plays a role in the evolution of these cross sections. We conjecture that it is possible to extend the Representative Consumer Theory of Distribution to the case in which consumers receive nondiversifiable income shocks that are proportional to total wealth.[24] It would be interesting to work out this case and determine how luck (modeled by these shocks) interacts with consumer heterogeneity to determine distributive dynamics.

Appendix: Data and Methods

From the 1971 and 1991 cross-sectional family data files we extract the variables family sampling *weight,* family *identity number, family size, taxable income of head and wife,* and *taxable income of others,* whose sum is our main measure of family income, *total family money income, state of residence,* and the head's *age, occupation, industry,* and *race.* To compute the mean of the family income distribution in a given year (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 for that particular regression.

Estimation of (22) requires an instrumental variable for $y_j^R(1970)$. First, the term $\gamma'\mathbf{X}_j$ will probably not fully account for heterogeneity in skills and tastes. Therefore, the error term contains the deviation of $\lambda_\pi \pi_j + \lambda_\beta \beta_j$ from its linear predictor. It is unlikely that this deviation will be uncorrelated with $y_j^R(1970)$. Because the error term (potentially) contains such unobservable idiosyncratic effects there is a possible omitted variable problem. Hence, we cannot use ordinary least squares.[25] A second reason for our choice of an instrumental variable approach is measurement error, which is a pervasive problem in the PSID. Measurement error would induce downward bias in an ordinary least-squares estimate of (22), whereas the correlated, omitted individual effect induces upward bias if $\lambda_y < 1$, and downward bias if $\lambda_y > 1$.

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

[24] Actually, if one is willing to restrict the analysis to the case of logarithmic preferences, we conjecture that it is also possible to treat shocks to the skill and taste parameters within the RC framework.

[25] This problem is also present in cross-country growth regressions if there are unobserved differences in, say, the aggregate 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 [λ_π and λ_β in equation (21)], so that differencing would not eliminate the problem. Therefore, although our empirical exercise is close in spirit to the recent cross-country empirical growth literature, our estimation strategy is quite different.

$$MN_j = \frac{Y_j^{MN}(1970)}{Y_{US}^{MN}(1970)} \quad \text{and}$$

$$MD_j = \frac{Y_j^{MD}(1970)}{Y_{US}^{MD}(1970)}$$

where Y_j^{MN} (Y_j^{MD}) is per capita income (median family income) in the state where family j was resident in year t, and Y_{US}^{MN} (Y_{US}^{MD}) is the per capita (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. 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. With this technique we estimate (22) with the result reported in the text. A variety of robustness checks (the results of which are available from the authors) show little sensitivity to alternative specifications.

REFERENCES

Abel, Andrew and Bernanke, Ben. *Macroeconomics.* Menlo Park, CA: Addison-Wesley, 1998.

Acemoglu, Daron. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics*, November 1998, *113*(4), pp. 1055–89.

Alesina, Alberto and Rodrik, Dani. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics*, May 1994, *109*(2), pp. 465–90.

Arrow, Kenneth. "The Economic Implications of Learning by Doing." *Review of Economic Studies*, June 1962, *29*, pp. 155–73.

Autor, David; Katz, Lawrence and Krueger, Alan. "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, November 1998, *113*(4), pp. 1169–213.

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

Bénabou, Roland. "Inequality and Growth," in Ben Bernanke and Julio J. Rotemberg, eds., *NBER macroeconomics annual 1996.* Cambridge, MA: MIT Press, 1996, pp. 11–74.

Bertola, Giuseppe. "Factor Shares and Savings in Endogenous Growth." *American Economic Review*, December 1993, *83*(5), pp. 1184–98.

Caselli, Francesco. "Technological Revolutions." *American Economic Review*, March 1999, *89*(1), pp. 78–102.

Cass, David. "Optimum Growth in an Aggregative Model of Capital Accumulation." *Review of Economic Studies*, July 1965, *32*(3), pp. 233–40.

Chatterjee, Satyajit. "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model." *Journal of Public Economics*, May 1994, *54*(1), pp. 97–119.

Deaton, Angus and Paxson, Christina. "Intertemporal Choice and Inequality." *Journal of Political Economy*, June 1994, *102*(3), pp. 437–67.

Galor, Oded and Tsiddon, Daniel. "Technological Progress, Mobility, and Economic Growth." *American Economic Review*, June 1997, *87*(3), pp. 363–82.

Gordon, Robert. "Comment," in Ben S. Bernanke and Julio J. Rotemberg, eds., *NBER macroeconomics annual 1996.* Cambridge, MA: MIT Press, 1996, pp. 259–67.

Greenwood, Jeremy and Yorukoglu, Mehmet. "1974." *Carnegie-Rochester Conference Series on Public Policy*, June 1997, *46*, pp. 49–95.

Juhn, Chinhui; Murphy, Kevin M. and Pierce, Brooks. "Wage Inequality and the Rise in the Return to Skill." *Journal of Political Economy*, June 1993, *101*(3), pp. 401–42.

Katz, Lawrence and Murphy, Kevin M. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics*, February 1992, *107*(1), pp. 35–78.

Koopmans, Tjalling. "On the Concept of Optimal Economic Growth," in *The econometric approach to development planning.* Amsterdam: North-Holland, 1965.

Kuznets, Simon. "Economic Growth and Income Inequality," in Deepak Lal, ed., *Development* *economics*, Vol. 2. Brookfield, VT: Ashgate (distributor) [Elgar], 1992, pp. 189–216.

Mas-Colell, Andreu; Whinston, Michael D. and Green, Jerry R. *Microeconomic theory*. New York: Oxford University Press, 1995.

Persson, Torsten and Tabellini, Guido. "Is Inequality Harmful for Growth?" *American Economic Review*, June 1994, *84*(3), pp. 600–21.

Ramsey, Frank. "A Mathematical Theory of Saving." *Economic Journal*, December 1928, *38*, pp. 543–59.

Romer, Paul. "Endogenous Technological Change." *Journal of Political Economy*, October 1990, *98*(5), pp. S71–S102.

Stiglitz, Joseph E. "Distribution of Income and Wealth Among Individuals." *Econometrica*, July 1969, *37*(3), pp. 382–97.