

IMF Working Paper

Short- and Long-Term Poverty and Social Policy in a “Snakes and Ladders” Model of Growth

Alfredo Cuevas

IMF Working Paper

Fiscal Affairs Department

**Short- and Long-Term Poverty and Social Policy in a
“Snakes and Ladders” Model of Growth**

Prepared by Alfredo Cuevas*

Authorized for distribution by Adrienne Cheasty

November 2001

Abstract

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Throughout the world, the great popularity of programs to protect those who may *fall into* poverty stands in contrast with the weakness of policies aimed at helping individuals who are already poor to overcome *long-term* poverty. In the paper, an OLG model with persistent poverty and limited social mobility is used to explore some of the reasons for the different success rates of these two types of policies, as well as the gains that can be expected from these and other policies in terms of economic growth. The popularity of social insurance schemes may be due to their relative ex-ante fairness, while the reluctance of societies to support effective policies to reduce long-term poverty may be explained by the redistributive bias of these policies, especially in the short term. However, the failure to attack long-term poverty can reduce long-run growth.

JEL Classification Numbers: I30, O15, O41.

Keywords: poverty, economic growth, human capital, social policy.

Author's E-Mail Address: acuevas@imf.org

* I'm grateful for comments and suggestions I received from Jose Bailen, Giuseppe Bertola, Adrienne Cheasty, Peter Clark, Luis Cubeddu, Avinash Dixit, Hamid Faruquee, Gene Grossman, Sanjeev Gupta, Eduardo Ley, and participants in seminars at the IMF, the Bank of Mexico, and CIDE. Any remaining errors are my responsibility.

Contents	Page
I. Introduction.....	3
II. A “Snakes and Ladders” Model of Growth with Poverty and Social Mobility.....	4
A. The Decision to Invest in Human Capital	4
B. Transmission of Social Status Across Generations.....	7
C. The Evolution of the Income Distribution.....	13
Main specification.....	13
A more general model with a possibility of escaping extreme poverty.....	16
D. Aggregate Production and Economic Growth.....	18
III. Policy Issues.....	20
A. Student Loans and Other Improvements in Financial Markets	20
B. Direct Poverty Reduction	21
C. Social Insurance	23
D. Income Stabilization Schemes for the Poor.....	26
E. Extensions.....	26
Country risk and growth.....	26
Economic crises	27
Mandatory basic education.....	27
IV. Summary and Conclusion.....	27
Figures	
1. Rewards to Study Effort Represented by the Geometric Mean of Earnings under Favorable and Adverse Shocks.	7
2. Intertemporal Decision for an Individual.	10
3. Assumptions A1–A3 Represented by the Crossing of the Horizontal Axis and the Gain in Utility from Studying.....	12
4. The Difference Equation for Income in the Presence of Positive Shocks.....	14

I. INTRODUCTION

One of the most striking contrasts we can find when surveying social policy throughout the world is that offered by the great popularity of mutual support schemes aimed at helping individuals cope with well-specified contingencies, and the weakness of policies aimed at helping traditionally disadvantaged groups overcome poverty.¹ This contrast constitutes one of the stylized facts of development, and understanding its causes and consequences is a challenge facing policymakers involved in development (e.g., Tanzi, Chu, and Gupta, 1999).

In this paper, we develop a theoretical model of economic growth with chronic and temporary poverty. Chronic poverty arises because of a threshold externality that creates a poverty trap, while the temporary poverty results from an adverse realization of a household-specific random shock that takes place late in the life of an individual, after all human capital investment decisions have been made. In this model, social mobility resembles the movements of a token on a board of the game of “snakes and ladders”: as long as productivity shocks are favorable, households headed by skilled workers experience upward social mobility, but a negative shock can produce downward mobility. In the model, the opportunity cost of investing in human capital during youth is forgone labor income. Introducing parental altruism, some degree of lumpiness in educational investment, and borrowing constraints can create a strong inertia in the distribution of income. Long-term growth is modeled as the result of an economy-wide spillover in the form of learning by doing. This externality makes long-run growth dependent on the form of the income distribution. In particular, chronic poverty can slow down steady-state growth.

We use this framework to analyze social policy. Social insurance schemes are modeled as programs designed to protect individuals from the effects of negative shocks during their working lives. By contrast, long-term poverty-reduction policies aim to break the cycle of under-investment in human capital occurring in persistently poor households. We find that the high influence of the groups vulnerable to poverty (relative to that of those already poor) could strengthen a tendency of society to pursue a mix of programs with a small emphasis on poverty reduction, allowing the perpetuation of poverty and possibly missing an opportunity to stimulate long-term growth.

This paper draws from the literature on growth reinvigorated, among others, by Lucas’ discussion of human capital and growth (Lucas, 1988), which emphasizes the role of education in development, and owes much to the observation that some countries (or

¹ For example, of the 164 countries surveyed in Social Security Administration 1995, 157 have pension systems, 104 have public health insurance, and 81 have family allowance schemes. Meanwhile, Birdsall and Londoño (1997) have noted that “the emphasis [on poverty reduction] in the international institutions contrasts, however, with disappointing results in the real world. For a sample of 57 countries with comparable data on poverty during the last three decades [. . .] income of the poorest quintiles has increased at slower rates than the income of the richest.”

individuals) may remain trapped in low-income trajectories due to the existence of “threshold externalities” (Azariadis and Drazen, 1990). The relationship between human-capital accumulation and growth—documented, *inter alia*, in Benhabib and Spiegel (1994), Park (1998), and Chapter 12 of Barro and Sala-I-Martin (1995)—has been analyzed in papers studying the organization of the education system in unequal societies (e.g., Glomm and Ravikumar, 1992; Benabou, 1996).

The possibility that poverty traps may prevent the accumulation of human capital with negative effects on equity and growth was discussed by Galor and Zeira (1993) using a model with borrowing constraints and lumpy investment in human capital. We build upon those insights to model an economy with permanent and temporary poverty. The effort to differentiate those who are born into poverty from those who fall into poverty later in life distinguishes this paper from other papers in this area, such as Loury (1981), Owen and Weil (1997), and others (see Chapters 9 and 10 of Aghion and Hewitt, 1998). Further motivation for our work came from the persistence of inequality reported in Levy and Murnane (1992) and Quah (1997), the preservation of social status documented in Becker and Tomes, and the distinction between short- and long-term poverty, developed in Bane and Ellwood (1986) and Stevens (1994).

The next section of the paper develops the main model and Section III uses it to examine social policies. The last section summarizes our main arguments and provides some conclusions.

II. A “SNAKES AND LADDERS” MODEL OF GROWTH WITH POVERTY AND SOCIAL MOBILITY

A. The Decision to Invest in Human Capital

Individuals live and consume for two periods. In period 1 (“youth”), they have an opportunity to invest their time endowments in human capital; but by doing so, they give up the wages they might earn as nonskilled laborers and that, along with an “allowance” they receive from their parents, finance their consumption. We assume that the only form of saving available is investment in human capital. In period 2, individuals supply their labor inelastically, earning a wage determined in part by luck and in part by the amount of human capital they acquired during their youth. These earnings finance the adult’s consumption and the allowance he gives to his offspring.

Concretely, an individual born at t derives utility from his own consumption when young and when old, c_t^y and c_{t+1}^o , and from a gift he gives his child—an allowance, a_{t+1} , where $t+1$ indicates the cohort of the recipient of the gift. This form of altruism prevents individuals

from having infinite planning horizons.² (The parameter λ in the following expression discounts second period utility; the subscript “ i ” identifies the household; in what follows, it will be shown only when strictly necessary.)

$$U_{ti} = \ln(c_{ti}^y) + \lambda \left((1 - \gamma) \ln(c_{t+1,i}^o) + \gamma \ln(a_{t+1,i}) \right). \quad (1)$$

The individual’s problem is to maximize (1) subject to his ability to earn income throughout his life and to the gift received from his parent. A young person can work as an unskilled laborer, earning a wage per unit of time of $w_0 K_t$ —a magnitude proportional to the current state of socially available knowledge, K_t , which is modeled as a nonrival input—for a fraction $1 - s_t$ of his time, and spend the remaining s_t units of his time studying. Young individuals cannot borrow to increase their first-period consumption, and neither can they save any part of their income. Thus, the constraints on the young are described by expression (2):

$$c_{ti}^y = (1 - s_{ti}) w_0 K_t + a_{ti}, \quad 0 \leq s_{ti} \leq 1. \quad (2)$$

Schooling, a random shock and an intrafamily externality determine income in period 2. These factors will be described shortly; for now, let us denote an adult’s income in period $t + 1$ by $I_{t+1,i} = y_{t+1,i} K_{t+1}$. Due to the form of preferences over adult consumption, $c_{t+1,i}^o$, and gift-giving, $a_{t+1,i}$, adult income is split in the familiar way, yielding adult utility equal to the sum of $\ln(I_{t+1,i})$ and a constant:

$$\begin{aligned} a_{t+1,i}^* &= \gamma I_{t+1,i} \\ c_{t+1,i}^{o*} &= (1 - \gamma) I_{t+1,i}. \end{aligned} \quad (3)$$

As noted above, adult earnings are the product of an individual component, $y_{t+1,i}$, and socially available knowledge, K_{t+1} . The individual component of earnings reflects past investments in human capital and luck. Expressions (4) and (5) indicate how this component is determined:

$$\begin{aligned} y_{t+1,i} &= w_0 & \text{if } s_{ti} = 0; \\ &= w_l & \text{if } s_{ti} > 0, \varepsilon_{t+1,i} = 0; \\ &= w_h h(s_{ti}, y_{ti}) & \text{if } s_{ti} > 0, \varepsilon_{t+1,i} = 1, \end{aligned} \quad (4)$$

where

$$h(s_t, y_t) = \alpha s_t^\beta y_t^\mu, \quad 0 < \beta < 1, \quad 0 < \mu < 1, \quad \beta + \mu \geq 1. \quad (5)$$

² Assuming that parents care about their children’s utility would not alter the main results of this section because a positive correlation between paternal income and size of the gift would be preserved. However, that assumption could alter some of the policy conclusions we can draw from our model.

Expression (4) says that if a young person does not study, his adult income will be that of an unskilled worker, that is, proportional to w_0 . But if a person studies, he will accumulate human capital to be used in production as an adult. The yield of the investment in human capital is, however, random: a household-specific shock, $\varepsilon_{t+1,i}$, affects the usefulness of human capital. We will assume that the shocks $\varepsilon_{t+1,i}$ are independent across households and over time, and have a simple Bernoulli distribution: with probability p , the shock takes on a value of 1, and with probability $1-p$, it takes on a value of zero. The value of the shock becomes known only at the beginning of the second period of life. Then, if the person faces $\varepsilon_{t+1,i}=1$, he will enjoy the full fruits of his investment in human capital, given by $w_h h(s_t, y_t)$, but if $\varepsilon_{t+1,i}=0$, he will have to make do with less income (w_l). Below, we will make the assumptions necessary to ensure that $w_0 < w_l < w_h h(s_t, y_t)$, a chain of inequalities embodying the notion that skilled wages exceed nonskilled wages, and that $\varepsilon_{t+1,i}=0$ represents an adverse shock.

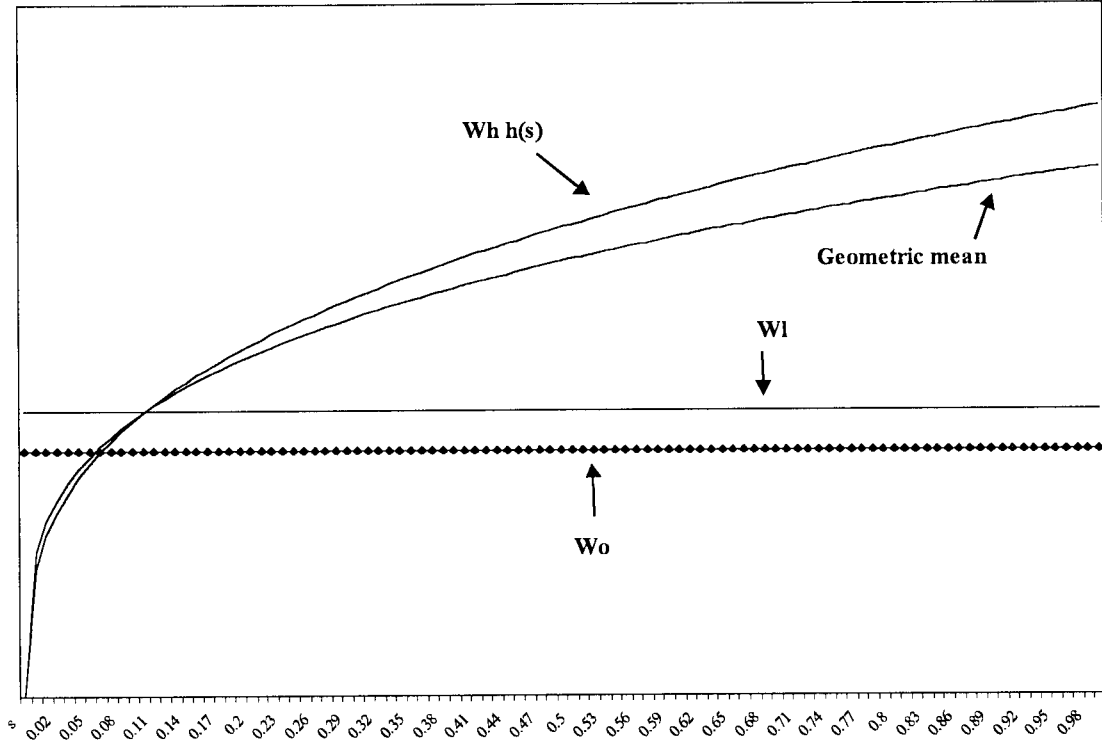
Expression (5) describes human capital as a function of two inputs: an individual's own effort, measured by s_t , and an intrahousehold externality denoted by the personal component of paternal income, y_t , representing the advantages that the children of better-off parents enjoy. Becker and Tomes (1979; 1986) discussed the role of intrahousehold externalities in the reproduction of social status. They observed that income tends to regress to the mean more slowly than talent because of the transfer of resources, including social capital, from one generation to the next. Glomm and Ravikumar (1992) obtained an expression similar to equation (5) by assuming that parents can spend in order to affect the quality of their children's education.

The form of expressions (4) and (5) reflects the assumption that investment in human capital has some degree of lumpiness, as noted by Galor and Zeira (1993). These authors required that a fixed amount of resources be spent acquiring an education, making it possible for their model to generate poverty traps. In our model, the nonconvexity is introduced in a different way in order to highlight the interaction between family background and returns to schooling, including any threshold effects. According to (4), if a person studies, he gives up adult unskilled wages in exchange for skilled wages; and according to (5), small amounts of schooling will produce only small rewards—lower, in fact, than unskilled wages for sufficiently small s_t . This point is illustrated in Figure 1 that suggests that negligible amounts of schooling effort are not worth considering.³ We can justify this assumption with the finding that marginal returns to schooling are higher in modern activities than in traditional activities such as agriculture (Rosenzweig, 1995).⁴

³ It is reasonable to assume that if a person studies for only a short amount of time in period 1, he can still work as an unskilled worker in period 2 (i.e., small investments in education should not reduce adult earnings). We retain the formulation in (4) and (5) because, under utility maximization, a person will only choose a positive amount of schooling if it raises his adult earnings.

⁴ This is a controversial issue. For example, Psacharopoulos (1994) argues that both private and social marginal returns to schooling are higher for elementary than for higher education, and Caldwell (1990) cites evidence of
(continued...)

Figure 1. Rewards to Study Effort Represented by the Geometric Mean of Earnings under Favorable and Adverse Shocks



B. Transmission of Social Status Across Generations

We can analyze the problem an individual confronts at the beginning of period 1 in two steps: first, he must decide how much effort to devote to study if he were to become a skilled worker, and then he must decide if he should in fact become skilled.⁵ The first step is to choose s_t to maximize the following expression, which measures expected lifetime utility:

$$V(s_t, y_t) = \ln \left[\left(\frac{\alpha_t}{K_t} + (1 - s_t) w_0 \right) K_t \right] + \lambda(1 - p) \ln(w_l K_{t+1}) + \lambda p \ln(w_h \alpha s_t^\beta y_t^\mu K_{t+1}). \quad (6)$$

the absence of any threshold effects in the negative relationship between maternal education and child mortality in certain developing countries.

⁵ This is just a comparison between an interior solution and a corner solution.

The interior solution to this problem is given by the following expression, where we have already substituted $\gamma y_t K_t$ for a_t , since adults (in this case the parent of the individual whose decision we are examining) give a fraction γ of their income to their children, as we saw in expression (3).

$$s_t = s(y_t) = \left(\frac{\beta \lambda p}{1 + \beta \lambda p} \right) \frac{w_0 + \gamma y_t}{w_0}. \quad (7)$$

Equation (7) indicates that investment in human capital increases with paternal income. The next step is to compare the utility earned by choosing s_t according to (7) against that obtained by setting $s_t = 0$. The quantity $\Delta V(y_t)$ given in the following expression is defined as the additional utility one may obtain by studying over what one would obtain by remaining unskilled for life.

$$\Delta V(y_t) = \ln \left(\frac{(\beta \lambda p)^{\beta \lambda p}}{(1 + \beta \lambda p)^{1 + \beta \lambda p}} \right) + \lambda \ln \left(\frac{(\alpha w_h y_t^\mu)^p w_l^{1-p}}{w_0} \right) + \beta \lambda p \ln \left(\frac{w_0 + \gamma y_t}{w_0} \right). \quad (8)$$

It is easy to verify that lower β implies a higher productivity of schooling, given that schooling can only take up a fraction of the unit time endowment; higher λ makes the reward to schooling, coming late in life, more desirable; and higher p makes it more likely that a person will experience the full benefits of schooling. A higher reward to schooling, measured by $w_h^p w_l^{1-p}$, will increase $\Delta V(y_t)$. By contrast, a higher unskilled wage, w_0 , reduces the value of $\Delta V(y_t)$ because it increases the opportunity cost of studying while reducing the relative pay of skilled labor.

More important for us, expression (8) says that the advantage a person may gain by studying depends on his parent's income. This works through two channels. First, the *productivity* of a person's investment in schooling is magnified by the intrahousehold externality, as reflected in the middle term of equation (8). And second, the *opportunity cost* of study—additional consumption when young—is less important for children who receive generous allowances at home. This is captured in the third element of equation (8). Thus, the higher up a person starts out in the socioeconomic ladder, the more attractive it is for him to invest in human capital, as can be verified by noting the positive sign of the derivative of $\Delta V(y_t)$ with respect to y_t .

$$\frac{\partial \Delta V(y_t)}{\partial y_t} = \frac{\gamma \beta \lambda p}{w_0 + \gamma y_t} + \frac{\lambda p \mu}{y_t} > 0. \quad (9)$$

For the remainder of the paper, we will assume that the parameters of the model are such that a person whose parents are unskilled will choose to remain unskilled. This will make possible the existence of poverty traps.

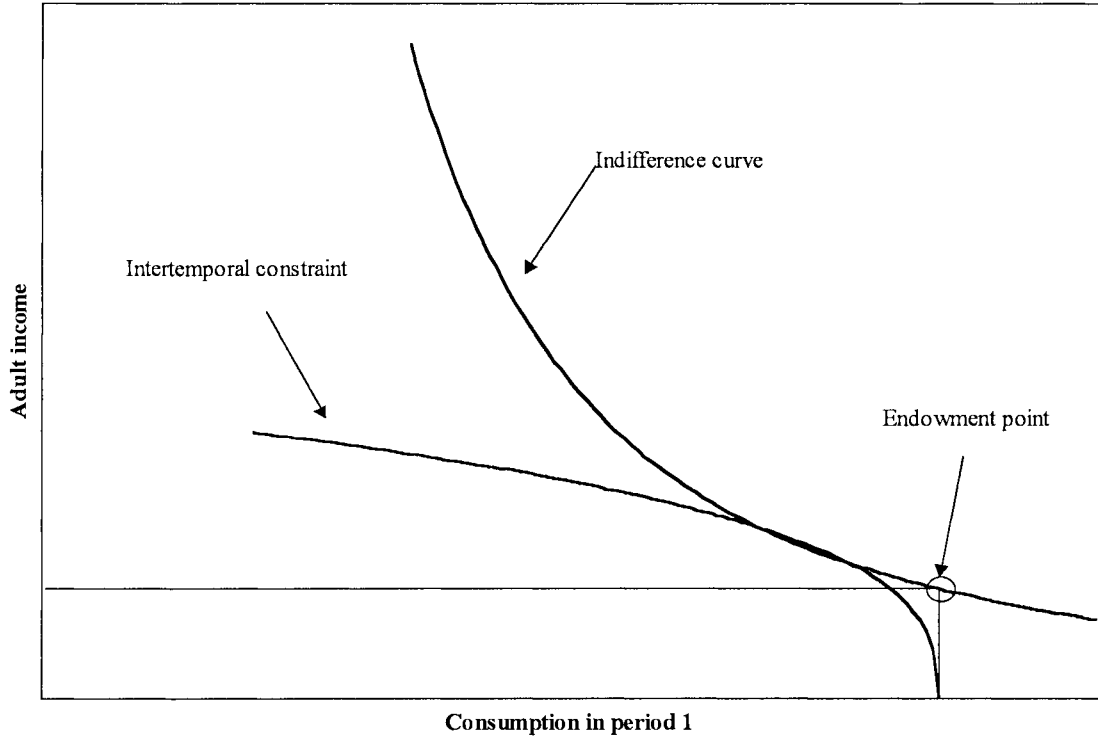
$$A1: \Delta V(w_0) = \ln \left(\frac{(\beta \lambda p)^{\beta \lambda p}}{(1 + \beta \lambda p)^{1 + \beta \lambda p}} \right) + \lambda \ln \left(\frac{(\alpha w_h w_0^\mu)^p w_l^{1-p}}{w_0} \right) + \beta \lambda p \ln(1 + \gamma) < 0. \quad (10)$$

Assumption A1 is likely to hold when γ and λ are low and w_0 is high, that is, when young people perceive labor earnings during period 1 as very valuable, perhaps because parents are not generous, and consider the future relatively unimportant. That assumption, along with the positive sign of the derivative of excess utility, $\Delta V(y)$, with respect to paternal income—expression (9)—ensures that we can solve the equation $\Delta V(y)=0$ and use its solution to define y_{min} as the level of paternal income that makes a child indifferent between remaining unskilled and studying. Although we can only solve $\Delta V(y)=0$ explicitly for some parameter configurations, we can identify some of the properties of y_{min} . It exceeds w_0 ; using the implicit function theorem we can establish that y_{min} is increasing in β , but decreasing in the relative wage $w_h^p w_l^{1-p}/w_0$, and in the parameters p , γ , λ and μ . These properties are reasonable: poorer people will consider schooling when its relative private reward is high and when their parents are more generous.

Figure 2, drawn assuming that $p=1$ for the sake of simplicity, illustrates some of the implications of assumption A1. The “endowment point” represents the amounts of current consumption and future income a person has guaranteed just as a result of having been born in a certain household: its abscissa is equal to $w_0 + \gamma y_{t,i}$, and its ordinate is w_0 . We have drawn the indifference curve passing through the endowment point, which represents the utility from being a full-time worker when young. The intertemporal constraint is the curve transforming cuts in current consumption into future income. It is simply a mirror image of Figure 1 (the human capital production function) re-scaled to take account of the fact that the horizontal axis in Figure 2 is measured in monetary units.

In Figure 2, the endowment point shifts horizontally to the right with increases in paternal income, flattening the indifference curve passing through the endowment point. A similar flattening can be obtained by raising γ . The intertemporal constraints of richer individuals are steeper than those of their poorer contemporaries on account of the stronger intrahousehold externalities they enjoy. A similar effect can be obtained increasing the productivity of study for all individuals by increasing β . Assumption A1 says that we will work with a range of parameter values that prevent us from drawing a picture such as Figure 2 for the children of unskilled workers. If the endowment point is given by the pair $([1+\gamma] w_0, w_0)$, representing the situation of the child of an unskilled worker, we should draw the intertemporal constraint completely to the left and below the indifference curve passing through the endowment point. (The individual represented in Figure 2, who appears rather indifferent to schooling, must be the child of a person earning about y_{min} .) Assumption A1, thus, makes it possible for the lumpy human capital-production process to generate poverty traps. If A1 were violated, all

Figure 2. Intertemporal Decision for an Individual



individuals would study and there would be no permanent poverty in this model, although we would still have random temporary poverty. This would make it closer to Loury (1981) than to Galor and Zeira (1993).

Maintaining assumption A1, and after some manipulation, we can write the identity $\Delta V(y_{min})=0$ in terms of the human capital accumulated by an indifferent individual:

$$\lambda p \ln \left(\frac{w_h h(s_{min}, y_{min})}{w_l} \right) = \ln \left(\frac{w_0^\lambda (1 + \beta \lambda p)}{w_l^\lambda} \right). \quad (11)$$

Using expression (11), we can write the condition needed to ensure that the state defined by $\varepsilon_{t+1}=0$ is worse than that defined by $\varepsilon_{t+1}=1$ for those who study, given that A1 holds. This condition, amounting to the requirement that the left-hand side of expression (11) be positive is labeled A2:

$$A2: \quad \frac{w_l}{w_0} < (1 + \beta \lambda p)^{\frac{1}{\lambda}}. \quad (12)$$

We need one last definition to describe the evolution of household income. Using expression (7), we can find the lowest level of paternal income that leads a young person

to become a full-time student. If paternal income equals y_f , defined by $s(y_f)=I$, the constraint $s \leq I$ is just binding. It is easy to verify that $y_f = w_0/(\gamma\beta\lambda p)$. With this definition, and under assumptions A1–A2, the evolution of the personal component of the incomes of the successive cohorts of a household can be described by the following first order stochastic difference equation:

$$\begin{aligned}
 y_{t+1} &= f(y_t, \varepsilon_{t+1}) = w_0 && \text{if } y_t \leq y_{\min}; \\
 &= w_l && \text{if } y_t > y_{\min}, \varepsilon_{t+1} = 0; \\
 &= \alpha w_h \left(\frac{\beta p}{1 + \beta p} \frac{w_0 + \gamma y_t}{w_0} \right)^\beta y_t^\mu && \text{if } y_{\min} < y_t \leq y_f, \varepsilon_{t+1} = I; \\
 &= \alpha w_h y_t^\mu && \text{if } y_t > y_f, \varepsilon_{t+1} = I.
 \end{aligned} \tag{13}$$

So far, our assumptions ensure that an unskilled worker has children who remain unskilled. To generate socioeconomic segregation, we need to ensure that the children of skilled workers will choose to become skilled, even if their parents are passing through a spell of bad luck. This assumption is very important; if it were violated, a negative shock would doom a household to permanent poverty, and, in time, the economy would consist only of unskilled workers.

$$A3: \quad w_l > y_{\min}. \tag{14}$$

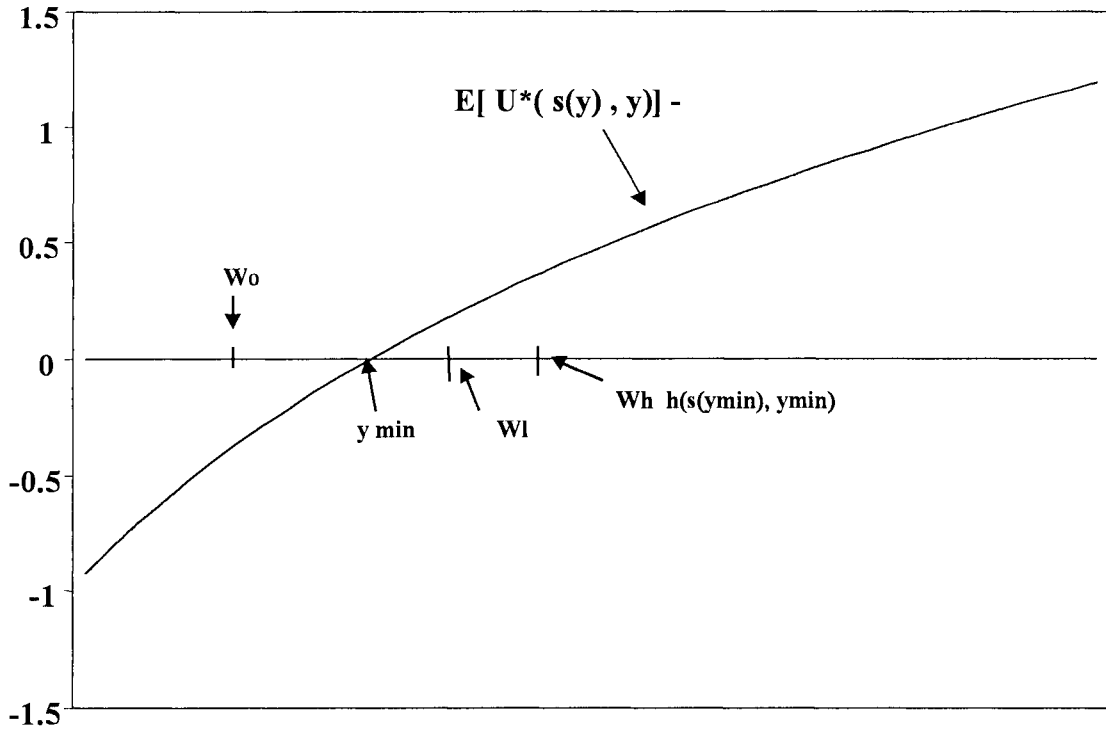
Assumptions A1–A3 produce the chain of inequalities $w_0 < y_{\min} < w_l < w_h h(s_{\min}, y_{\min})$, as shown in Figure 3. With these assumptions, the child of an unskilled worker will prefer to remain unskilled. We refer to this situation as *chronic poverty*. These households are poor because they are unskilled, and their poverty makes the sacrifice necessary to become skilled unacceptable. By contrast, the child of a skilled worker will always choose to acquire some human capital. The children of individuals currently experiencing a negative shock study for the shortest positive time, $s(w_l)$. These households experience *temporary poverty*. By construction, temporary poverty is less severe than chronic poverty, and is transitory because the schooling decision of the child holds the possibility of upward social mobility as soon as the household's luck improves. Once good luck returns, household income will rise to $w_h h(s(w_l), w_l) > w_l$, and subsequent generations will stand a chance of being even better off.

Notice that the assumptions we have made so far are not sufficient to guarantee that a skilled worker's expected lifetime income will be higher than it would have been had he remained unskilled, although this is more likely to be true for the children of better-off parents. To see this, notice that, for anyone who chooses to study, $\Delta V(y_l) > 0$. This fact, and the concavity of the logarithmic function, allows us to write the following for children of parents with incomes $y' \geq y_{\min}$:

$$0 \leq \Delta V(y') < \lambda \ln \left[\frac{p(w_h h[s(y')]) + (1-p)w_l}{w_0} \right] + \ln \left(\frac{\gamma y' + [1-s(y')]w_0}{\gamma y' + w_0} \right) \quad (15)$$

$$\Leftrightarrow w_0(\gamma y' + w_0) < (\gamma y' + [1-s(y')]w_0) [p(w_h h[s(y')]) + (1-p)w_l].$$

Figure 3. Assumptions A1–A3 Represented by the Crossing of the Horizontal Axis and the Gain in Utility from Studying



The second line of (15) is just a restatement of the first one after taking antilogs on both sides of the first inequality, using the fact that $\lambda \leq 1$, and rearranging. Expression (15) says that the product of adult income and child consumption is higher as a result of studying for those who study. This may reflect higher lifetime income, a smoother consumption path during the two periods of life, or both. It is conceivable that a person may choose to study, even at some cost in terms of lifetime income, just in the interest of achieving some consumption smoothing. This possibility arises because education is the only vehicle for the intertemporal transfer of resources in the model, and the “endowment point” has more resources in period 1 than in period 2. This type of effect is absent from models where consumption is concentrated in one period, such as those found in Glomm and Ravikumar (1992), Galor and Zeira (1993), and Owen and Weil (1997), and it would disappear if financial savings were allowed (we develop this issue in Section III.A). Other things equal, the higher the value of γ and λ , the stronger the influence on the studying decision exerted by the desire to smooth consumption. Also

notice that the left-hand side of (15) is linear in y' , but the right-hand side of (15) rises faster (in fact, if $\beta + \mu \geq 1$, it increases at least with the square of y'). Thus, for large y' , studying will be more likely to result in higher, and not just smoother, lifetime income.

Let us turn now to nonpoor households. Assume for the moment that there are individuals who decide to be full-time students. If their descendants enjoy an arbitrarily long run of good luck, their incomes may continue to vary on account of the externality factor included in the human-capital accumulation equation, y_t^μ . The resting value of y_t during this run of good luck should be considered the maximum level of income because no higher value can persist, given $\mu < 1$. This level of income is defined by the equation $y_{max} = \alpha w_h y_{max}^\mu$, and it equals $(\alpha w_h)^{1/(1-\mu)}$. Now, we started by assuming that some individuals choose to be full-time students. The following inequality ensures that y_{max} , as defined, exceeds y_f , a condition for the existence of full time students in the long-run equilibrium of the model. In what follows, however, we will not insist on making this assumption.

$$A4: \quad (\alpha w_h)^{\frac{1}{1-\mu}} > \frac{w_0}{\beta \gamma p}. \quad (16)$$

Based on expression (13), Figure 4 shows what the difference equation for family income during runs of good luck, $f(y_b, I)$, looks like when A1-A3 hold but A4 does not, and there is only one fixed point with $y_t > w_0$.⁶ In general, $f(y_b, I)$ could have more than one fixed point to the right of w_0 . But the number of individuals with incomes above the smallest of those fixed points would approach zero over time because the family's income falls to w_l when an adverse realization of the shock obtains. From there, it can reach only the smallest fixed point of $f(y_b, I)$ to the right of w_0 .

C. The Evolution of the Income Distribution

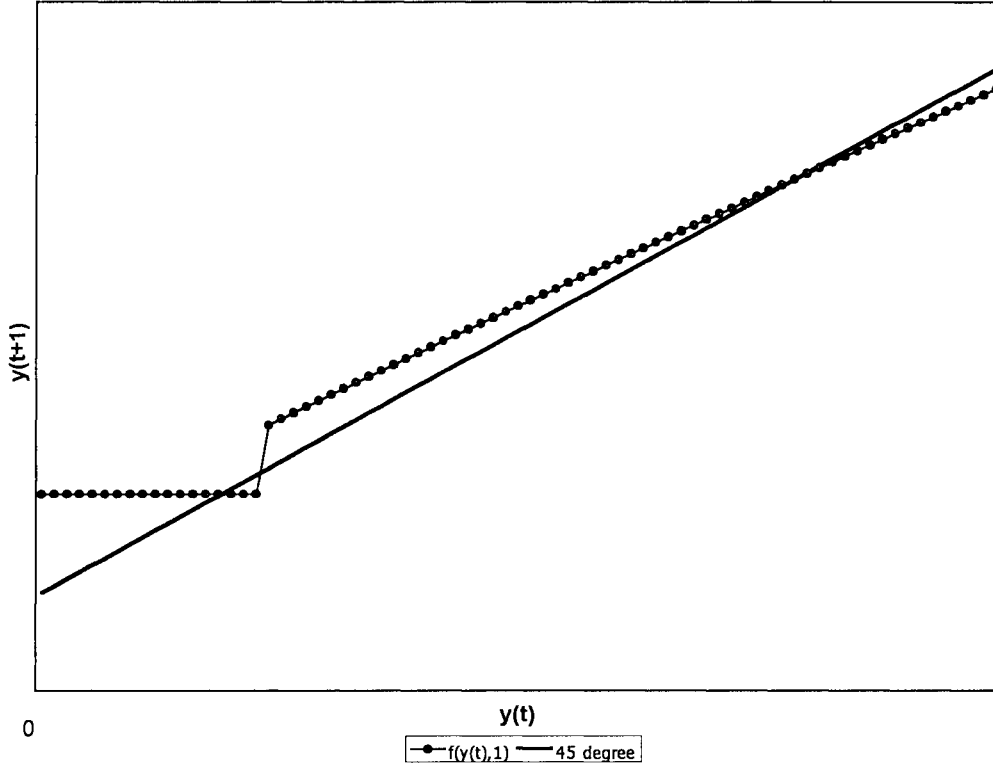
Main specification

Under assumptions A1-A3, and according to expression (13), the incomes of the different cohorts of a family form a Markov chain. In a chronically poor household, a person born in period T-1 has an adult income proportional to w_0 in period T, and the adult income of the person born at T in that household will be proportional to w_0 in period T+1 for sure.⁷ In other households, if the person born at T-1 has an adult income proportional to $y_T > y_{min}$, the income of the person born one period later is proportional to $y_{T+1} = f(y_T, I)$ with probability p , and to w_l with probability $1-p$.

⁶ The parameter values chosen to draw Figure 3 are $w_0=0.285$, $w_l=0.33$, $w_h=1$, $p=0.85$, $\beta=0.33$, $\alpha=1.475$, $\gamma=0.33$, $\mu=1-\beta$, and $\lambda=0.5$. For these parameters, $y_{min}=0.316$, and the maximum sustainable level of income is 0.741, the unique fixed point of $f(w_b, I)$ to the right of w_0 .

⁷ The constants of proportionality are the levels of socially available knowledge in each one of those periods.

Figure 4. The Difference Equation for Income in the Presence of Positive Shocks



For any initial distribution of the personal component of income among adults, $G_0(y)$, this stochastic process will eventually lead to a unique, invariant distribution of income. We can discuss this process as it unfolds one period at a time. After one period has passed, the first two brackets of the long-term income distribution will attain their definite values. The bottom bracket is characterized by $G_1(w_0) = G_1(y_{min}) = G_0(y_{min})$. The second lowest brackets is that of the temporarily poor, consisting of all adults who, after studying during $t=0$, experience an adverse shock $\varepsilon_1=0$, so that the personal component of their adult income is w_l ; this group attains its definite size in $t=1$: $G_1(w_l) - G_1(w_0) = (1-p)(1 - G_0(y_{min}))$. The rest of the households are headed by individuals who studied in $t=0$ and experienced a positive shock in $t=1$; their earnings are w_h $h(s_0) > w_l$. One period later, a new group attains its long-term size: that of the adults who grew up in temporarily poor households in $t=1$, but have good luck in $t=2$. This group is p times as large as the group of the temporarily poor. With each passing generation, a new income bracket is defined and attains its long-term number of occupants, so that, at any time $t > 1$, the distribution of income among adults is a mixed distribution function described by the following expression, where $f^{(m)}(y, 1)$ is the m^{th} iterate of the function $f(y, 1)$; that is, $f^{(2)}(y, 1) = f(f(y, 1), 1)$, and so on; we define $f^{(0)}(y, 1) = y$.

$$\begin{aligned}
 G_t(y) &= 0 && \text{for } y < w_0; \\
 &= G_0(y_{\min}) && \text{for } y \in [w_0, w_1); \\
 &= G_t(w_0) + (1 - G_0(y_{\min}))(1 - p) && \text{for } y \in [w_1, f(w_1, I)); \\
 &= G_t(f^{(m-2)}(w_1, I)) + (1 - G_0(y_{\min}))(1 - p) p^{m-1} && \text{for } y \in [f^{(m-1)}(w_1, I), f^{(m)}(w_1, I)), \\
 &&& \text{for all integers } m \text{ such that } 2 \leq m \leq t; \\
 &= G_t(f^{(t-1)}(w_1, I)) + p^t (G_0(x) - G_0(y_{\min})), && \text{where } x \text{ solves } y = f^{(t)}(x, I), \\
 &&& \text{for } y \geq f^{(t)}(y_{\min}, I).
 \end{aligned} \tag{17}$$

Expression (17) combines information on the initial position and subsequent fortunes of a household. The bottom income bracket includes those born in households whose members have been unskilled from the start; the second bracket from the bottom, those skilled adults currently experiencing an adverse shock; the third bracket, those whose parents suffered an adverse shock the previous period, but are currently experiencing a positive shock, and more generally, the m^{th} bracket contains adults living in households that have enjoyed runs of favorable shocks of length $m-2$ after their most recent negative shock. The rest of the adults live in households that have never experienced an adverse shock; the size of this group declines over time as a fraction $1-p$ of its members run into bad luck in any period. Thus, the long-run income distribution corresponds to that described in the first four items of expression (17), since the limit as t approaches infinity of the second term of the expression in the last line is zero. It will be useful to introduce a simpler notation for this limit distribution by letting ϕ_i denote the size of the i^{th} income group:

$$\begin{aligned}
 \phi_0 &= G_0(y_{\min}) && \text{for } y_t = w_0 \\
 \phi_n &= (1 - G_0(y_{\min}))(1 - p) p^{n-1} && \text{for } y_t = f^{(n-1)}(w_1, I), n \geq 1.
 \end{aligned} \tag{18}$$

Social mobility in the steady state is governed by our simple Markov chain. Let us say that, up to the constant representing socially available knowledge, state 0 is defined by household-head earnings of w_0 , state 1 by earnings of w_1 , and state j by earnings of $f^{(j-1)}(w_1, I)$. Then, we can write the long-run transition matrix for this chain in the following way:

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1-p & p & 0 & 0 & \dots \\ 0 & 1-p & 0 & p & 0 & \dots \\ 0 & 1-p & 0 & 0 & p & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (19)$$

The isolation of the chronically poor from the rest of society is reflected in that M can be partitioned into two submatrices representing two sets of states that do not communicate with each other: the submatrix consisting of the element in the first row and column of M , and the submatrix obtained by deleting the first row and column of M . This last submatrix corresponds to a Markov process where each state represents the length of a “success run” in a sequence of independent Bernoulli trials with parameter p . (This is a consequence of our assumption that any household headed by a skilled person can fall into temporary poverty). This chain is irreducible and it can be shown that all its states are aperiodic, persistent, and have finite recurrence times (i.e., they are ergodic). Moreover, its asymptotic behavior is described by the invariant distribution given by the expression $(1-p)p^n$ for runs of length $n \geq 0$, which we find embedded in the second line of expression (18); the mean recurrence time of each state is just the reciprocal of this quantity.⁸ In terms of our model, the states characterized by low incomes are also those with the shortest recurrence times.

A more general model with a possibility of escaping extreme poverty

In our model, extreme poverty is synonymous with permanent poverty because of the existence of an hermetic poverty trap. Here we consider moving to a more general problem. Assume that the income of an unskilled adult person is random: it equals w_0 if $\xi_{t+1}=0$, an event with probability $1-q$, and w_l if $\xi_{t+1}=1$, an event with probability q . Also, make downward mobility possible for the children of skilled workers experiencing an adverse shock. As before, these children can enjoy the full skilled income, $w_h h(\cdot)$, with probability p , but they can end up being forced into unskilled labor if they are hit by a bad shock, an event with probability r ; with probability $1-p-r$, they earn w_l . Our model so far is a special case of this problem, with $q=r=0$. But how would the description of the economy change if we allowed for the additional possibilities arising in this larger problem? In what follows, a superscript “g” will denote values associated with this more general problem

Expected lifetime utility from studying for all nonpoor individuals is still described by expression (6), and their optimal study effort is still given by expression (7). Remaining

⁸ See Feller (1968), Chapter XV; and Karlin and Taylor (1975), Chapter 2. The ergodicity of all states in the “success runs” chain can be verified by first establishing that state 1 is aperiodic, persistent and has a mean recurrence time of $1/(1-p)$, and then using the fact that all states in an irreducible chain are of the same type.

unskilled now yields utility $\ln(w_0 + \gamma y_l) + \lambda [q \ln(w_l) + (1-q) \ln(w_0)]$, higher than before because of the possibility of obtaining income w_l . Thus, the new excess utility from schooling, $\Delta V^g(y)$, is lower than $\Delta V(y)$ from expression (8) by the amount $\lambda q [\ln(w_l) - \ln(w_0)]$ for all non-poor households. Households where current adult earnings are w_l or less are assumed to be in danger of falling to w_0 with probability r despite investing in schooling. There, $V^g(y)$ is lower than $V(y)$ by $\lambda r [\ln(w_l) - \ln(w_0)]$, so that the excess utility from schooling, $\Delta V^g(y)$, is lower than $\Delta V(y)$ by $\lambda(r+q)[\ln(w_l) - \ln(w_0)]$. If a child living in such a household decides to study, however, optimal schooling is still given by expression (7) because the actual value of s_t is relevant to determine earnings only when the person enjoys a positive shock.

In summary, the new $\Delta V^g(y)$ is everywhere below the one in expression (8) and has a discontinuity at $y=w_l$, with the curve jumping down from the right at that point. We can then impose conditions analogous to A1–A3 to ensure that $\Delta V^g(w_0) < 0 = \Delta V^g(y_{min}^g) < \Delta V^g(w_l)$ and that $w_l < w_h h(s(y_{min}^g), y_{min}^g)$, so that our new excess utility function crosses the horizontal axis much like the old one did in Figure 3. As a result of the lower placement of $\Delta V^g(y)$ relative to $\Delta V(y)$, in this version of the model $y_{min}^g > y_{min}$. That is, the range of paternal incomes where studying is optimal is narrower in the more general model of this subsection. This reflects the fact that escape from deep poverty is possible without studying, increasing the expected utility from remaining unskilled. The meaning of the “states” is different too: earnings of w_0 signal deep rather than permanent poverty, while w_l can be considered moderate rather than temporary poverty. Social mobility is now described by a new transition matrix M^g and has an invariant income distribution with weights ϕ_i^g :

$$M^g = \begin{bmatrix} 1-q & q & 0 & 0 & 0 & \dots \\ r & 1-p-r & p & 0 & 0 & \dots \\ 0 & 1-p & 0 & p & 0 & \dots \\ 0 & 1-p & 0 & 0 & p & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \quad (20)$$

$$\phi_0^g = \left(1 + \frac{q}{r(1-p)} \right)^{-1}$$

$$\phi_1^g = \frac{q}{r} \phi_0^g = \frac{q(1-p)}{q + r(1-p)} \quad (21)$$

$$\phi_i^g = p^{i-1} \phi_1^g = \frac{q(1-p) p^{i-1}}{q + r(1-p)} \quad \text{for } i = 2, 3, \dots$$

As in Loury (1981), the long-term income distribution bears no connection whatsoever to the initial distribution. The reason is that the whole matrix M^B is irreducible. The nonconvexities in the human-capital accumulation function still give rise to a group of households where remaining unskilled is the strategy of choice; but this does no longer imply the existence of a permanent poverty trap. In the case where $r=q=0$, the first two income brackets attain their long-run size within one period, and in each period a new bracket does the same. In the more general case discussed in this subsection, convergence toward (21) does not occur stepwise, because all brackets attain their long-run sizes asymptotically. We will use this model in Section III to discuss income-stabilization schemes. For the moment, we will continue to work with the version of the model where $r=q=0$.

D. Aggregate Production and Economic Growth

We mentioned earlier that adult income is the product of a personal component, $y_{t,i}$, and a nonrival production factor, K_t . This factor is the stock of knowledge accumulated by the economy as a by-product of its activity. We will assume that the economy experiences learning by doing, with the change in the value of K_t between t and $t+1$ equal to aggregate production in t , which we will denote by $Y_t K_t$. By definition, then, the proportional rate of growth of K_t equals Y_t . In our model, $Y_t K_t$ is simply the weighted average of production in each household—with weights given by (18)—which includes adult production $y_{it} K_t$ and the labor income of the younger generation, $[1-s(y_{it})]w_0 K_t$.

$$\begin{aligned} \frac{K_{t+1} - K_t}{K_t} = Y_t = & 2\phi_0 w_0 + \sum_{i=1}^t \phi_i \left[w_0 \left(1 - s[f^{(i-1)}(w_1, 1)] \right) + f^{(i-1)}(w_1, 1) \right] \\ & + p^t \int_{y_{\min}}^{\infty} \left[w_0 \left(1 - s[f^{(t)}(x, 1)] \right) + f^{(t)}(x, 1) \right] dG_0(x) \end{aligned} \quad (22)$$

Expression (22) records the contribution to production (and growth) of chronically poor households in its first term, of households that have experienced at least one negative shock in the middle summation, and of households that have enjoyed an uninterrupted run of good luck in the last integral. Outside the steady state, their sum may change over time, altering the rate of growth of knowledge. The third term in expression (22) becomes less important with each passing generation, as fewer households continue to enjoy an uninterrupted run of good luck. Each period, an additional term is added to the summation in the middle of that expression, representing the production of households which have enjoyed the longest runs of good luck after suffering an adverse shock in $t=1$. The balance of these two effects determines whether the growth rate falls or rises toward its long-term value.⁹

⁹ This balance depends on the form of the initial distribution, G_0 , and on whether t is high enough that $s(f^{(t)}(w_b, 1))=1$. Imagine that all individuals whose adult income exceeded y_{\min} at time $t=0$ were earning $y_0 < w_b$, and that at the present moment, $t=T$, $s(f^{(T)}(w_b, 1))=1$. One period later, the right hand side of (20) should be larger, since a fraction $1-p$ of the households represented by the integral, with income $f^{(T)}(y_0, 1)$, are replaced by

(continued...)

In the steady state, the invariance of the distribution of income implies that Y_t is constant, with a value Y_∞ that equals the limit as t goes to infinity of the right-hand side of (22). Therefore, the steady-state rate of growth of aggregate output is equal to that of K_t , which is itself equal to Y_∞ . It is clear from (22) that Y_∞ rises with the probability that investment in human capital will result in upward mobility, p . It takes a bit more work to see the influence on steady-state growth of the initial distribution of income. That distribution continues to affect growth even in the long run through ϕ_0 . The following expression answers this question: if at time zero the number of the chronically poor, given by $G_0(y_{min}) = \phi_0$, were changed by altering the form of G_0 but without changing any parameter of the model, how would long run growth be affected? The answer is that chronic poverty may slow down growth in the steady state.

$$\begin{aligned} \frac{\partial Y_\infty}{\partial \phi_0} &= 2w_0 - (1-p) \sum_{i=1}^{\infty} p^{i-1} \left[w_0 \left(1 - s \left[f^{(i-1)}(w_p, 1) \right] \right) + p f^{(i-1)}(w_1, 1) \right] \\ &= - (1-p) \sum_{i=1}^{\infty} p^{i-1} \left[w_0 \left(1 - s \left[f^{(i-1)}(w_1, 1) \right] \right) + (1-p)w_1 + p f^{(i)}(w_1, 1) - 2w_0 \right]. \end{aligned} \quad (23)$$

The algebra behind (23) is straightforward. First, we transformed the weighted sum of the father-and-son incomes in households with skilled members into the weighted sum of the expected, undiscounted lifetime incomes of the young people living in those households. Then, we brought the income of the unskilled individuals under the summation sign. The expression now consists of the weighted sum of the net gain in expected lifetime productivity—over that of a permanently unskilled worker—accruing to those individuals who study. As we discussed above—see expression (15)—this net gain is increasing in the income of the household head, although it need not be positive for those individuals born in households where adults make incomes only a little higher than y_{min} . If p is relatively high, the weights given to the groups with the larger net gains in income will decline slowly, making it more likely that (23) will be negative. As we indicated above, lower values of γ and λ and higher values of β and μ strengthen the income-raising effect of education, also increasing the likelihood that (23) is negative.

The result that chronic poverty may hinder steady-state growth is strengthened under alternative forms of the externality driving the growth of K_t . We may assume that knowledge is created only in the “modern sector,” so that $K_{t+1} - K_t$ equals the combined output of all skilled workers. Then, the rate of growth of knowledge would be obtained by dropping all terms involving w_0 from the right-hand side of (22). In this case, it is easy to see, with the help of expression (18), that ϕ_0 hurts growth. A similar conclusion follows if we assume, somewhat less plausibly, that new knowledge arises only during the transmission of existing

individuals making $f^{(T)}(w_t, 1)$, which is higher because $w_t > y_0$, while the remaining fraction p of those households see their income rise to $f^{(T+1)}(y_0, 1) > f^{(T)}(y_0, 1)$.

knowledge, so that $K_{t+1} - K_t$ is equal to the total amount of schooling. It is also possible to see that, in the more complex setting of Section II.C, summarized by expressions (20) and (21), steady-state growth may still be hampered by the extent of deep poverty, φ_o^g . However, in this case the initial income distribution G_o does not affect the extent of deep poverty in the steady state. Given the mechanics of the transition toward the ergodic income distribution noted at the end of Section II.C, however, the influence of the initial level of poverty, $G_o(y_{min}^g)$, on the extent of poverty at any future time diminishes only gradually.

III. POLICY ISSUES

A. Student Loans and Other Improvements in Financial Markets

We have assumed that there are no markets where individuals may borrow in order to study, and that the only vehicle for “saving” is investment in human capital. Introducing student loans and financial savings, while increasing the efficiency of the economy, can increase the dispersion of the distribution of human capital within a cohort. To see this, we write the expected marginal return to schooling (“MRS”) at the locally optimal schooling level given by expression (7):

$$MRS = (1 + Y_t) p \frac{w_h \alpha \beta y_t^\mu}{w_o} \left(\frac{\lambda \beta p}{1 + \lambda \beta p} \frac{w_o + \gamma y_t}{w_o} \right)^{\beta-1} \quad (24)$$

The derivative of MRS with respect to y_t is proportional to $\mu w_o + \gamma (\beta + \mu - 1) y_t$, showing that students born in households with different paternal income do not just have different average productivity in school, but also have different equilibrium marginal products of schooling. Assuming that $\beta + \mu \geq 1$, if credit markets operated perfectly, offering loans and taking deposits at the same interest rate, all takers of student loans would be children of better-off parents, whose marginal productivity at school is highest. The funding for those loans would come from poorer students who would have spent modest amounts of time engaged in schooling in the credit-constrained equilibrium, and who may now reduce their study efforts further, and even work full time, and use the financial markets as a more efficient vehicle to transfer income to their adult age. These savers are individuals whose motivation to study would have included a strong desire to smooth consumption, as was noted in the discussion of expression (15). If assumption A4 holds so that the children of the richest individuals are full-time students, then also these individuals, unable to study more because of a binding time constraint, would build up financial savings (a similar reversal of saving behavior could occur if $\beta + \mu < 1$). Even with different lending and borrowing rates, this pattern would persist.

In the equilibrium with operative credit markets, the usefulness of young people’s time endowments would increase: those who are more productive in school would get to study longer, and the efficiency gains would be shared by savers in the form of higher adult incomes inclusive of interest on savings. This translates automatically into higher growth rates in our model. The middle of the distribution of human capital would be thinned out, but this would no longer suffice to determine the distribution of wealth and income in a wider

sense—the long-term behavior and growth of a model with credit markets is harder to characterize, as the state space includes financial assets and not just labor income.

Relieving credit market imperfections by introducing (possibly subsidized) student loans is not a remedy for poverty. Student loans provide the financing for the more efficient use of a person's endowment, and chronic poverty is defined by the inability of parents to endow their child with enough resources to move up in life. Addressing credit market imperfections will increase the efficiency of the economy and its long-term growth rate, but it need not make it more egalitarian. It is, however, a policy that most individuals would welcome, but whose high implementation costs make very difficult to pursue.

B. Direct Poverty Reduction

Two common measures of poverty are the head-count ratio and the poverty gap. The first of them indicates the prevalence of poverty, and it is given by ϕ_0 in our model. The second indicates the amount of resources needed to lift the poor to the poverty line: $\phi_0(y_{\min} - w_0)K_t$. If we could give a chronically poor household head a transfer $z = K_t(y_{\min} - w_0)$, his child would study, escaping long-term poverty. In this section, we analyze the effects of a policy that redistributes income toward the chronically poor during the course of one generation. If chronic poverty is eradicated or reduced, long-term growth is likely to increase—see the discussion concerning expression (23). We will model this policy as assistance to the head of the household in order to take advantage of familiar quantities.¹⁰

We will require a revenue-raising instrument to finance transfers to the poor. Consider a constant tax of size τK_t to be paid by all skilled persons currently enjoying good luck (that is, persons with $\varepsilon_{t,i} = I$). The tax should be chosen so total revenues equal the poverty gap:

$$\begin{aligned} \phi_0(y_{\min} - w_0)K_t &= \sum_{i=2}^{\infty} (\tau \phi_i K_t) = (1 - \phi_0) p \tau K_t \\ \rightarrow \tau &= \frac{\phi_0(y_{\min} - w_0)}{(1 - \phi_0) p}. \end{aligned} \tag{25}$$

¹⁰ Targeting assistance directly to the young poor rather than to their parents is more complicated than just giving them γz , because the intrahousehold externality would remain limited at w_0^{μ} . In this case, the best approach would be to look for the scholarship that minimizes $x_0 + x_1 s(x_0, x_1; w_0)$, where x_0 is a “signing-up” bonus, x_1 is a per-period payment, and $s(x_0, x_1; w_0)$ is the best study effort by the child of an unskilled adult in response to this grant, subject to a participation constraint $E\{U[s(x_0, x_1; w_0), w_0]\} \geq U[0, w_0]$, and to a constraint ensuring that this child's descendants will be no worse off than those who are temporarily poor, $w_h h(s(x_0, x_1; w_0), w_0) \geq w_l$.

Assuming society can afford it,¹¹ let us suppose that, starting from a steady-state position, the tax prescribed by expression (25) is levied, for one time only, in period T , and that proceeds from its collection are used to give each head of a chronically poor household a transfer $z = (y_{\min} - w_0)K_T$. Looking only at the adults living in period T , it is clear that the utility of the recipients of the transfer increases, while that of the taxpayers falls as a result of the changes in their disposable incomes. In turn, these changes affect the allowances adults give to their children. As intended, the children of the recipients of the transfer now receive a gift that allows them to study for a fraction $s(y_{\min})$ of their time. But the children of taxpayers will reduce their study effort from $s(y_T)$ to $s(y_T - \tau)$. The reverse side of the coin is that the poorest children reduce their labor supply, while the children of taxpayers increase theirs. The balance of these effects will determine whether output and growth change in the short run. The next expression measures the change in the growth rate at T :

$$\begin{aligned} \Delta Y_T &= -\phi_0 w_0 s(y_{\min}) + \sum_{i=1}^F \phi_{i+1} [s(f^{(i)}(w_1, 1)) - s(f^{(i)}(w_1, 1) - \tau)] \\ &= -\phi_0 w_0 \frac{\lambda \beta \rho}{1 + \lambda \beta \rho} \left[1 + \gamma(1 - p^F) + \frac{p^F \gamma y_{\min}}{w_0} \right] < 0. \end{aligned} \quad (26)$$

In expression (26), we have denoted by F the smallest integer such that $f^{(F+1)} - \tau > y_f$; that is, households so indexed have after-tax incomes sufficiently high that their children are full time students; if assumption A4 does not hold, there are no full-time students and we should then take the limit as F goes to infinity of (26). Expression (26) indicates that a temporary sacrifice of output and growth may be necessary in order to move the economy to a steady state with higher sustainable growth and without persistent poverty. If instead of depending on total output, the growth of knowledge depended on the output of skilled workers only, then poverty reduction would leave the growth rate of K_t unchanged, given that we have assumed that adults supply their labor inelastically (relaxing this assumption would allow distorting taxation to reduce the growth rate of K_t). In this case, there would be a fall in output, but not in the growth rate of K_t , which would, in fact, increase as early as in $T+1$, when the children of the recipients of the transfers become adults.

The result embodied in expression (26) suggests why effective poverty-reduction policies are so difficult to implement politically. Note that a taxpayer has lower utility as a result of this policy and that his child has been made unambiguously worse off too: he receives a lower gift at T and has lower human capital at $T+1$, when he enjoys an economy-wide externality K_{T+1} that may have shrunk because of the lower output in period T . In other words, one could not find support for this policy among the people living in taxpaying households at time T , whose horizon extends no further than period $T+1$. If each person had one vote, $\phi_0 > 1/2$ would be needed for this policy to be adopted by a majority vote. If collective decisions are

¹¹ In particular, we mean by this that the after-tax income of the poorest taxpayer is still above w_t .

especially influenced by those with economic power, a redistributive policy such as the one discussed here would be even less viable.

This result follows largely from the financing of social assistance with taxes levied contemporaneously, and could be strengthened if adult labor supply were elastic. But if the government could borrow funds from outside the system to finance assistance to the poor, the policy would be welcomed by most individuals. The low political appeal of the policy is also due to the assumption that parents obtain utility directly from their gifts to their children. If parents cared about their children's utility, the result may need to be qualified because, under this formulation, a person's horizon extends beyond his lifetime. In that case, an adult may give enough weight to the future—when the benefits of poverty reduction are reflected in the economy as a whole—to vote for this policy. In fact, if K_t grows with the output of skilled workers, the benefits of poverty reduction become noticeable only two generations down from that of the taxpayers, with K_{t+2} exceeding the level it would have had without poverty reduction, making a positive vote by these “visionaries” more likely.

C. Social Insurance

In Section II, we noted that study time is increasing in the probability of a positive shock to human capital, p . In our model, then, the fact that $p < 1$ results in a reduced level of investment in human capital relative to what would be observed if such investments were riskless. At the social level, this can lead to a steady state with relatively low-growth rates. However, a social insurance scheme that moderates the effects of an adverse shock on the individual could raise expected utility and lead to increased schooling, output and growth for the economy as a whole.

In our model, the natural form of a social insurance program is a fair insurance scheme for individuals investing in human capital, whereby adults receive a “benefit” zK_t if they experience a negative shock, and pay a “contribution” $zK_t(1-p)/p$ if they experience a positive shock. The parameter z is assumed to be small enough that $\varepsilon_t = 0$ continues to signal a worse state of the world than $\varepsilon_t = 1$. Young individuals planning to study would welcome such a scheme, because they are risk averse. This scheme has no redistributive bias *ex-ante*; however, it needs to be enforced externally because the transfers from the lucky to the unlucky are not self-enforcing. Both the shock $\varepsilon_{t,i}$ and the effort $s_{t-1,i}$ are observable.

The introduction of this scheme is announced at the start of period T with the proviso that the first payments made to and by the scheme will take place at $T+1$. At time T , news of the scheme lead young people to increase their schooling relative to what they would have done without the policy. To see this, we write the first order conditions for an interior solution to the problem with a social safety net like this:

$$s_t - \left(\frac{\lambda\beta p}{1 + \lambda\beta p} \frac{w_0 + \gamma y_t}{w_0} \right) = \frac{1}{1 + \lambda\beta p} \frac{z \frac{(1-p)}{p}}{w_h \alpha s_t^\beta y_t^\mu} . \quad (27)$$

Expression (27) implicitly defines a function $s_z(y_t; z)$; it is clear that $s_z(y_t; 0) = s(y_t)$ as we have defined it so far—see expression (7). The structure of (27) immediately indicates that $s_z(y_t; z) > s(y_t)$ if $z > 0$; using the implicit function theorem to analyze expression (27), we find that the partial derivatives of $s_z(y_t; z)$ with respect to its two arguments are positive, provided z is not too large.¹² Moreover, the second term in the left-hand side of (27) is equal to $s(y_t)$, so that the increase in schooling due to the introduction of the social insurance scheme is given by the right-hand side of that equation; this allows us to see that the increase in schooling brought about by the scheme is smaller among the children of richer individuals.

The individuals who are now insured have a higher expected utility. The additional utility earned through participation in this scheme, evaluated at $z=0$, can be expressed as follows using the envelope theorem (the last inequality is due to assumption A2):

$$\left[\frac{\partial V(y_t, z)}{\partial z} \right]_{z=0} = \frac{1-p}{w_l} - \frac{1-p}{w_h h(y_t)} \geq (1-p) \left(\frac{1}{w_l} - \frac{1}{w_h h(y_{\min})} \right) > 0. \quad (28)$$

Expression (28) indicates that even a small social insurance program is to be welcomed. But it has further implications. Recall that, in the model of Section II, $V(w_0) < 0$ by assumption A1, and $V(y_{\min}) = 0$ by definition. Then, for sufficiently small z , even if $V(w_0, z)$ is still negative, the income that makes a person's child indifferent to schooling is lower than y_{\min} after the announcement of the scheme. In terms of Figure 3, the scheme shifts up the graph of ΔV , pushing y_{\min} to the left. That is, the insurance scheme makes it possible to eradicate poverty with a smaller fiscal effort, because it reduces the poverty gap, measured by $\varphi_0(y_{\min} - w_0)$. Moreover, it is possible that if w_0 and y_{\min} were not too far apart to begin with, this scheme could make $V(w_0, z)$ nonnegative: the introduction of this type of safety net could draw the poorest members of society to study, even though they are not the intended target of this policy. This situation is similar to that captured by migration models where the provision of services in one locality gives individuals residing elsewhere the incentive to move. In our example, however, “migrants” are moving from low productivity to high productivity activities, contributing to the economy-wide acceleration of learning by doing. In what follows, we will leave aside this potential effect of the social insurance program on chronic poverty in order to focus on its other effects.

Following the announcement of the insurance scheme, the younger members of households whose heads are not chronically poor study more and reduce their labor supply, causing a fall in output in T . This may cause a reduction in the rate of growth of knowledge, according to whether learning by doing responds to total output, to skilled output or to schooling, as we discussed at the end of Section II. Thus, with the possible exception of children living in

¹² It is easy to show that $\partial s_z(y, z) / \partial y$ is proportional to $\lambda \beta p (\gamma w_h h(\cdot) / s w_0) - (\mu / y) [z(1-p)/p]$. We have already required that z be small enough that $\varepsilon = 0$ continues to represent a bad outcome.

chronically poor households, who may remain unskilled while possibly seeing the growth in the nonrival factor K_t slow down, the introduction of the social insurance scheme would be welcome by most young people and not rejected by any member of the older generation. Thus, its political viability is high.¹³

But in order to understand the full implications of social insurance, we must analyze its impact on the long-term functioning of the economy. To this effect, we will define two new functions analogous to the function $f(y, I)$ we have used before, where the second argument, ε , is fixed at 1 to remind us that this function describes the transmission of social status during good times. The new functions are $f_z(y, I)$ and $g_z(y, I)$ and represent, respectively, the gross and net income of a skilled worker enjoying a positive shock to human capital in an economy with a social safety net. Just as we had $f(y, I) = w_h h(s(y), y)$, we must now have $f_z(y, I) = w_h h(s_z(y, z), y)$, and $g_z(y, I) = f_z(y, I) - z(1-p)/p$. We have seen that, for any given y , $s_z(y, z) \geq s(y)$, implying that $f_z(y, I) \geq f(y, I)$, with equality only if there are full-time students in the equilibrium without social safety nets. The differentiability of the human capital production function, $h(s, y)$, allows us to write the following Taylor expansion for f_z , where $R(y, z)$ is a remainder term of order less than $s_z(y, z) - s(y)$.

$$\begin{aligned} f_z(y, I) - f(y, I) &= \beta w_h \alpha s(y)^{\beta-1} y^\mu [s_z(y, z) - s(y)] + R(y, z) \\ &= \frac{f(y, I)}{f_z(y, I)} \left[\frac{w_0}{\lambda p (w_0 + \gamma y)} \frac{z(1-p)}{p} \right] + R(y, z) \\ \Leftrightarrow \frac{f_z(y, I) - f(y, I)}{f(y, I)} &= \beta \frac{s_z(y, z) - s(y)}{s(y)} + \frac{R(y, z)}{f(y, I)} \end{aligned} \quad (29)$$

The first line of (29) is the first order Taylor series for $w_h h(s_z(y, z), y)$, rewritten in the second line with the help of expressions (7) and (27) in a way that will be useful shortly; the third line divides through by $f(y, I)$ in order to give an idea of the proportions involved. As we noted above, $s_z(y, z) - s(y)$ is decreasing in y , and so we learn from the last line of expression (29) that the ratio of $f(y, I)$ to $f_z(y, I)$ must get closer to 1 as y increases. So, for high values of y , the difference between these two quantity approaches the quantity inside the square bracket in the middle line, and for sufficiently high values of y , the quantity inside the bracket falls short of the premium paid by individuals enjoying a positive shock. Since $g_z(y, I) = f_z(y, I) - z(1-p)/p$, this indicates that, for sufficiently high y , $g_z(y, I) < f(y, I)$.

Defining the iterates of f_z and g_z in the same way we did for f , the foregoing discussion leads to the conclusion that $g_z^{(t)}(w_l + z, I) = w_l + z > w_l = f^{(t)}(w_l, I)$, but possibly $g_z^{(t)}(w_l + z, I) < f^{(t)}(w_l, I)$

¹³ In reality, social insurance schemes often offer something to the older generations soon after their introduction, a fact that increases their popularity. Any imbalances are typically dragged for various generations before they must be corrected.

for sufficiently high t . If this is the case, the dispersion of after-transfer incomes among skilled workers is reduced as result of the operation of social insurance. This safety net increases the personal component of the disposable incomes of persons living in households with relatively short runs of good luck and reduces that of persons living in households with longer runs of good luck. But the actual level of output and growth depends on the gross income of these individuals, which, as we have seen, is higher under the insurance plan for all or most individuals. Thus, long-run economic growth is likely to be faster with a social insurance scheme, especially if knowledge grows with aggregate skilled output or schooling. This conclusion would be weakened if key individual information were not not observable. For example, if the scheme cannot separate a person earning w_0 from one earning w_l , the insurance scheme cannot be actuarially fair. Then, the reward to those adults who did not study will rise to $(w_0+z)K_t$, increasing y_{min} and reducing the incentive to study relative to the perfect information case.

D. Income Stabilization Schemes for the Poor

The generalized model of Section II.C, with its treatment of poor households' incomes as random, is best suited to illustrate the possible effects of a low-income stabilization scheme. Denote the mean income of an unskilled adult by $w_p = qw_l + (1-q)w_0$; by construction, its variance is $q(1-q)(w_l - w_0)^2$. An income stabilization scheme modeled as a fair insurance mechanism paying w_p for sure to the poorest household heads would increase the well being of unskilled individuals. But if $w_p < y_{min}^g$, the child of an insured unskilled worker will not study, even if his parent experiences a positive shock $\xi=1$: deep poverty would become chronic, if more tolerable. Moreover, chronic poverty would become a sink because a transition from the state where paternal income is w_l to the state where it equals w_0 is still possible. This policy would be ultimately unsustainable.

We can also consider a case where, even if the incomes of the poor are random, it is still true that $r=q=0$. In that case, by assumption, even during good times the income of an unskilled adult is less than y_{min} . An income stabilization scheme would again be welfare increasing for the poor, but it would entrench them more deeply in poverty even in this case, an effect that can be represented by a downward shift of $\Delta V(.)$ in Figure 3. By raising y_{min} , this scheme would make the task of poverty eradication harder.

E. Extensions

Country risk and growth

Imagine two small open economies differing only in the strength of their institutions. With free capital flows, w_h may differ between these two countries in response to the interest rates inclusive of country risk demanded by foreign investors. The less stable country, facing a higher exogenous interest rate, would have a lower level of physical capital, reducing the productivity of human capital in the more productive sectors of the economy—that is, reducing w_h . The main consequences of a lower w_h are less investment in human capital in

steady state, a larger size of the group of chronically poor people, and slower economic growth in the long run.

Economic crises

Assume that, due to an exogenous crisis of the modern sector, during a generation p is perceived as lower, or w_h or w_l fall while w_0 stays unchanged. Even if temporary, these events reduce the investment in human capital of the younger generation and increase y_{min} and could—if w_l falls and/or y_{min} rises sufficiently—throw families experiencing a negative household shock into permanent poverty. Even after the crisis has passed, the fact that it led children to abandon schooling prematurely could permanently increase the size of φ_0 , reducing long-run growth. This scenario suggests the need for a flexible social safety net that may be easily reinforced during a crisis.¹⁴

Mandatory basic education

Mandatory minimum education can be interpreted in two different ways in our model. For example, we may simply consider that the definition of an unskilled worker already incorporates an existing mandatory minimum education: what matters for the determination of income is what one is able to do beyond the socially enforced minimum standard. Under this interpretation, raising the mandated minimum education would only lead to a parallel upward shift of the whole schedule of educational attainment. An alternative interpretation would be that mandatory education programs require individuals to make at least a nonzero investment in human capital, s_r . This policy may only be binding on the poorest members of society and may be suboptimal for them; in addition, if s_r is insufficiently high, the policy will leave chronic poverty unchanged. These effects would have to be traded off against some of the benefits—which we have not discussed in this paper—that would justify a mandatory minimum education policy in the first place, such as the dissemination of the basic rules of social interaction.

IV. SUMMARY AND CONCLUSION

We developed a model of endogenous growth to examine some varieties of social policy. In the model, liquidity-constrained young people must sacrifice current consumption in order to undertake lumpy investments in human capital. In this setting, youngsters growing up in richer households are able to invest in their own human capital, while the children of poorer parents choose to study less. Under some conditions, there might be a group of very poor households where children grow up to be unskilled workers who themselves will head poor households. In the other families, social status is passed on from one generation to the next

¹⁴ See Chu and Gupta (1998) for an explanation of the concept of social safety nets, and Gupta et al. (1998) for a discussion of the strengthening of social safety nets during the recent macroeconomic turmoil in Asia. Gupta et al. (2000) discusses how social safety nets have been incorporated into IMF programs.

with a high probability, although bad luck can push them into temporary poverty. Over time, the distribution of income tends toward a stable form that may depend on the initial size of the group comprising the poorest households. The distribution of income affects output and growth. In fact, a reduction in poverty today is likely to yield permanent gains in economic growth.

We used the model to explore the effects of various policies to reduce poverty and to mitigate the negative effect of adverse productivity shocks. One of our main results is that a temporary (one generation) redistributive scheme that successfully eradicates poverty can put the economy on a trajectory leading to a steady state with higher sustainable growth. The scheme produces this result by enabling the poor to invest in their own human capital, thus allowing them to escape the low-productivity trap. But while any transfer program will increase the welfare of the poor in the short run, a successful poverty eradication program must be sufficiently large to halt the intergenerational transmission of poverty.

However, the political viability of a poverty-reduction effort cannot be taken for granted. The collection of tax revenue to finance this effort makes taxpayers and their children worse off, and can lead to a temporary fall in output and economic growth. Given this trade-off between present personal sacrifice and future social benefits, the nonpoor members of society would oppose the redistributive scheme, and in light of the fact that political influence is often correlated with economic clout, the odds are against the adoption of poverty-reduction policies.

A social safety net has to be permanent, given the recurrent nature of the shocks capable of reducing the productivity of skilled individuals. Even when these schemes have a modest size, they can be Pareto-improving in an expected-value sense for the group of individuals alive during the introduction of the schemes. This makes them easy to accept by society at large, a fact consistent with the popularity of social safety nets in both developed and developing countries. We also found that, by raising the well being of the high-productivity groups with access to them, social safety nets can induce poor families to make the sacrifices needed to invest in human capital. In a sense, a migration toward the more modern, policy-protected sectors of the economy may take place. Under certain circumstances, this effect alone can help reduce chronic poverty. In any case, the adoption of social safety nets lowers the cost of poverty reduction, making unambiguously redistributive policies more acceptable to any "visionaries" in the economy capable of caring about a future extending beyond their lifetimes. By contrast, income-stabilization schemes for the poor may entrench poverty by making it less vexing. This type of policy could then have the effect of making radical poverty-reduction efforts more difficult. We also noted that the provision of student loans, while enhancing the economy's efficiency, is most useful to those individuals who start out with high household-provided endowments. Thus, student loans alone cannot be relied upon to reduce poverty.

In sum, even a modest social safety net or a student-loan scheme can be politically popular and economically sensible; but a poverty-reduction initiative has to be of a certain size to succeed, and financing it is politically difficult. Mobilizing those resources requires

convincing taxpayers that short-term sacrifices may be needed to usher in a more egalitarian and productive economy.

References

- Aghion, P., and Peter Howitt, 1998, *Endogenous Growth Theory* (Cambridge, Mass.: MIT Press).
- Azariadis, C., and A. Drazen, 1990, "Threshold Externalities in Economic Development," *Quarterly Journal of Economics*, Vol. 105 (May) pp. 501–26.
- Bane, M., and D. Ellwood, 1986, "Slipping Into and Out of Poverty: The Dynamics of Spells," *Journal of Human Resources*, Vol. 21 (Winter), pp. 1–23.
- Barro, R., and X. Sala-I-Martin, 1995, *Economic Growth* (New York: McGraw-Hill).
- Becker, G., and N. Tomes, 1979, "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, Vol. 87 (December) pp. 1153–89.
- , 1986, "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, Vol. 4 (July), pp. S1–S39.
- Benabou, R., 1996, "Heterogeneity, Stratification and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review*, Vol. 86 (June), pp. 584–609.
- Benhabib, J., and M. Spiegel, 1994, "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data," *Journal of Monetary Economics*, Vol. 34 (October), pp. 143–73.
- Birdsall, N., and J. Londoño, 1997, "Asset Inequality Does Matter: Lessons from Latin America," Inter-American Development Bank OCE Working Paper.
- Caballe, J., 1995, "Endogenous Growth, Human Capital and Bequests in a Life Cycle Model," *Oxford Economic Papers*, Vol. 47 (January), pp. 156–81.
- Caldwell, J., 1990, "Cultural and Social Factors Influencing Mortality Levels in Developing Countries," *Annals of the American Academy of Political and Social Science*, Vol. 510 (July), pp. 44–59.
- Chu, K., and S. Gupta, 1998, *Social Safety Nets: Issues and Recent Experiences* (Washington: International Monetary Fund).
- Feller, W., 1968, *An Introduction to Probability Theory and its Applications, Volume 1* (New York: John Wiley).

- Galor, O., and D. Tsiddon, 1996, "Income and Growth: The Kusnetz Hypothesis Revisited," *Economica*, Vol. 63, pp. S103–S117.
- , 1997, "The Distribution of Human Capital and Economic Growth," *Journal of Economic Growth*, Vol. 2 (March), pp. 93–124.
- Galor, O., and J. Zeira, 1993, "Income Distribution and Macroeconomics," *Review of Economic Studies*, Vol. 60 (January), pp. 35–52.
- Glomm, G., and B. Ravikumar, 1992, "Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality," *Journal of Political Economy*, Vol. 100 (August), pp. 813–34.
- Gupta, S., C. McDonald, C. Schiller, M. Verhoeven, Z. Bogetic, and G. Schwartz, 1998, "Mitigating the Social Costs of the Economic Crisis and the Reform Programs in Asia," IMF Paper on Policy Analysis and Assessment 98/7 (Washington: International Monetary Fund).
- Gupta, S., L. Dicks-Mireaux, R. Khemany, C. McDonald, and M. Verhoeven, 2000, *Social Issues in IMF-Supported Programs*, Occasional Paper 191, IMF, Washington.
- Karlin, S., and H. Taylor, 1975, *A First Course in Stochastic Processes* (New York: Academic Press).
- Levy, F., and R. Murnane, 1992, "US Earnings Levels and Earnings Inequality: A Review of Recent Trends," *Journal of Economic Literature*, Vol. 30 (September), pp. 1333–81.
- Loury, G., 1981, "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, Vol. 49 (June), pp. 843–67.
- Lucas, R., 1988, "On the Mechanics of Economic Development," *Journal of Monetary Economics*, Vol. 22 (July), pp. 3–42.
- Owen, A., and D. Weil, 1997, Intergenerational Earnings, Mobility, Inequality and Growth, National Bureau of Economic Research Working Paper 6070.
- Park, K., 1998, "Distribution and Growth: Cross Country Evidence," *Applied Economics*, Vol. 30 (July), pp. 943–49.
- Psacharopoulos, G., 1994, "Returns to Investment in Education," *World Development*, Vol. 22 (September), pp. 1325–43.
- Quah, D., 1997, "Empirics for Growth and Distribution: Stratification, Polarization and Convergence Clubs," *Journal of Economic Growth*, Vol. 2 (January), pp. 27–59.

- Rosenzweig, M., 1995, "Why Are There Returns to Schooling?," *American Economic Review*, PPAEA, Vol. 85 (May), pp. 153–58.
- Stevens, A., 1994, "The Dynamics of Poverty Spells: Updating Bane and Ellwood," *American Economic Review*, PPAEA, Vol. 84 (May) pp. 34–37.
- Social Security Administration, 1995, *Social Security Systems throughout the World* (Washington: Social Security Administration).
- Stokey, N., and R. Lucas, 1989, *Recursive Methods in Economic Dynamics* (Cambridge: Harvard University Press).
- Tanzi, V., K. Chu, and S. Gupta, 1999, *Economic Policy and Equity* (Washington: International Monetary Fund).