

ECONOMIC RESEARCH REPORTS

*Tax and Education Policy in a
Heterogeneous Agent Economy:
What Levels of Redistribution
Maximize Growth and Efficiency?*

By

Roland Benabou

RR# 99-12

May 1999

**C.V. STARR CENTER
FOR APPLIED ECONOMICS**



NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, NY 10003-6687

Tax and Education Policy in a Heterogenous Agent Economy:
What Levels of Redistribution Maximize Growth and Efficiency?

Roland Bénabou ¹

New York University, CERAS–IDEI and GREMAQ–CNRS.

First Draft: March 1997

This Version: January 1999

¹I wish to thank Martin Lettau, Guillaume Rabault, Patrick Rey, Jean Tirole and seminar participants at IDEI and the University of Chicago Graduate School of Business for helpful comments. Federico Ravenna provided excellent research assistance. Financial support from the National Science Foundation (SBR–9601319), the MacArthur Foundation and the C.V. Starr Center is gratefully acknowledged.

Abstract

This paper studies the effects of progressive income taxes and education finance in a dynamic heterogeneous agent economy. Such redistributive policies entail distortions to labor supply and savings, but also serve as partial substitutes for missing credit and insurance markets. The resulting tradeoffs for growth and efficiency are explored, both theoretically and quantitatively, in a model which yields complete analytical solutions. Progressive education finance always leads to higher income growth than taxes and transfers, but at the cost of lower insurance. Overall efficiency is assessed using a new measure which properly reflects aggregate resources and idiosyncratic risks but, unlike a standard social welfare function, does not reward equality per se. Simulations using empirical parameter estimates show that the efficiency costs and benefits of redistribution are generally of the same order of magnitude, resulting in reasonable values for the optimal rates. Aggregate income and aggregate welfare provide only very crude lower and upper bounds around the true efficiency tradeoff.

Keywords: Income Distribution, Heterogenous Agents, Inequality, Growth, Redistribution, Education Finance.

JEL classification: D31, E62, I22, D63.

Introduction

Absent the representative agent assumption, the evolution of macroeconomic aggregates is determined jointly with that of the entire distribution of wealth. In this I paper present a framework where these usually complex dynamics remain analytically tractable, and use it to study fiscal and educational policy. I thus analyze the effects of progressive income taxes and redistributive education finance on aggregate income, inequality, social mobility, individual risk, and intertemporal welfare. For each policy I ask what degree of progressivity is efficient, or simply growth-maximizing; I also compare the relative merits of two forms of redistribution. For such questions to be posed realistically, two ingredients must be present. Redistributive policies must have costs, due to distortions in agents' effort or savings decisions. They must also have benefits, due to imperfections in asset markets; redistribution then provides both insurance and a means to relax the credit constraints which impede certain investments.

To analyze this tradeoff theoretically and quantitatively, I develop a stochastic model of human capital accumulation with endogenous effort and missing credit and insurance markets. The model also incorporates recursive preferences a la Kreps-Porteus where agents' risk aversion, which determines the insurance value of redistributive policies, is independent of their intertemporal elasticity of substitution. Explicit analytical solutions are obtained for all individual and aggregate variables, under constant or time-varying policies. Given a reasonably broad menu of fiscal instruments, intertemporal distortions are shown to be preventable: to each income tax or education finance policy can be associated a simple combination of consumption taxes and investment subsidies which restores savings to its (constrained) optimal level. The analysis also demonstrates that progressive education finance always leads to higher income growth than taxes and transfers. Both are equally effective at substituting for the missing credit market, but the second policy entails smaller distortions to labor supply and (in the absence of corrective measures) to savings, because it redistributes only a fraction of family income. This comes, however, at the cost of lower consumption insurance.

To evaluate more generally the extent to which market distortions and imperfections are worsened or improved by alternative policies, the paper proposes a new measure of overall economic efficiency. This criterion properly reflects (dynamic) variations in the aggregate consumption of goods and leisure and in the idiosyncratic risks that agents face; but, unlike a standard so-

cial welfare function, it does not reward interpersonal equality per se. The underlying idea is straightforward. Instead of aggregating individual incomes or consumptions (which washes out all idiosyncratic uncertainty), or individual utilities (which introduces a bias towards egalitarian allocations), one sums up consumption certainty-equivalents, so as to obtain a kind of risk-adjusted GDP. Aggregate efficiency is shown to be maximized at some strictly positive rate of redistribution which depends intuitively on parameters like the elasticity of labor supply, the variability of idiosyncratic shocks, and the growth losses from liquidity-constrained investments. Equity concerns can be incorporated as well, but through a separate degree of inequality-aversion which is a priori independent of individuals' attitudes towards risk and intertemporal fluctuations.

Complementing the theoretical analysis, the model is simulated with parameter estimates from the empirical literature. In the baseline specification long run income (or growth) is maximized by an average marginal tax-and-transfer rate of 21%, which corresponds to a share of redistributive transfers in GDP of 6%. Taking into account the value of insurance and leisure, efficiency maximization raises these numbers to 48% and 14% respectively. Under the alternative policy of progressive education finance, the income-maximizing equalization rate for school expenditures is 62%, the efficient one 68%. In both cases, the efficient policy results in the top 30% of families subsidizing the bottom 70%, whether through the fiscal or the education system. Maximizing average welfare would always imply much higher rates of redistribution.

Naturally, these numbers should only be taken as providing a rough assessment of the main policy tradeoffs. More important than the results corresponding to specific parameters are the general lessons emerging from an extensive sensitivity analysis. First, the efficiency costs and benefits of redistribution are typically of the same order of magnitude, so that neither side can be neglected. Second, aggregate income and aggregate welfare provide only very crude lower and upper bounds around the true efficiency tradeoff.

This paper relates to three strands of literature. First, and most directly, to the work on growth and distribution with imperfect credit markets (e.g., Lorry (1981), Galor and Zeira (1993), Banerjee and Newman (1993), Perotti (1993), Saint-Paul and Verdier (1994), Bénabou (1996a), Durlauf (1996), Aghion and Bolton (1997), Piketty (1997)).¹ This literature has remained es-

¹This list is far from exhaustive: see also Banerjee and Newman (1991), Glomm and Ravikumar (1992), Fernandez and Rogerson (1996), Bénabou (1996b), Gradstein and Justman (1997), Cooper (1998), and many others.

entially theoretical, with the notable exception of Fernandez and Rogerson (1998) who study the quantitative impact of redistributing educational expenditures in a model calibrated to US data. The framework proposed here allows new developments on both the theoretical and the quantitative fronts. The second strand of literature deals with the implications of imperfect insurance markets for savings behavior and wealth inequality on the one hand, (Laitner (1992), Aiyagari (1994), Heaton and Lucas (1996), Krusell and Smith (1998)), for public insurance on the other (Varian (1981), Persson (1983), Hansen and İmrohođlu (1992), Atkeson and Lucas (1995), Hopenhayn and Nicolini (1997), Acemoglu and Shimer (1998)).² The present paper abstracts from precautionary savings, which figure prominently in several of these models. On the other hand, it lets credit constraints bear not just on consumption smoothing, but also on investment. Finally, the paper relates methodologically to some of the asset pricing literature, both in its use of non-expected utility (Epstein and Zin (1989), Weil (1990)) and in obtaining exact aggregation in an economy with heterogeneous agents (Constantinides and Duffie (1996)). But whereas this literature focuses on explaining asset price movements with real quantities being exogenous (endowment economies), this paper explores the polar case of endogenous macroeconomic dynamics in an economy with missing asset markets.

Section 1 describes the model, then derives agents' optimal labor supply and savings behavior under progressive taxes and transfers. Section 2 does the same for progressive education finance. Section 3 solves for the dynamics and steady-state values of total income and its cross-sectional distribution. It also shows how to allow for endogenous growth in the model, without any of the results being affected. Sections 4 and 5 develop an efficiency criterion for such dynamic, stochastic economies with wealth heterogeneity, then relate it to aggregate income and social welfare functions. Section 6 explores the quantitative predictions of the model through a wide range of simulations. Section 7 concludes. All proofs are gathered in the appendix.

²Here again, the list only a partial one. See for instance Aiyagari and Peled (1995) and Castañeda, Díaz-Jimenez and Ríos-Rull (1998). A related line of work reassesses, under incomplete markets, some classical issues of optimal taxation such as the long-run tax on capital (Aiyagari (1997), Chamley (1997)) or tax smoothing (Bassetto (1998)).

1 The Model

1.1 Preferences and Technology

There is a continuum of infinitely-lived agents or dynasties, $i \in [0, 1]$. In period t , agent i chooses consumption c_t^i and labor supply l_t^i to maximize his intertemporal utility U_t^i , subject to technological and budget constraints described below. The simplest description of preferences I will consider is:

$$\ln U_t^i = \mathbf{E}_t \left[\sum_{k=0}^{\infty} \rho^k (\ln c_{t+k}^i - \delta (l_{t+k}^i)^\eta) \right]. \quad (1)$$

More generally, I use a specification of preferences which allows attitudes towards intertemporal substitution and towards risk to be distinguished (Kreps and Porteus (1979), Epstein and Zin (1989), Weil (1990)). Because the latter determines the value placed by agents on the insurance component of redistributive policies, it is important not to constrain it to any particular value. On the other hand, the model's analytical tractability requires a unitary intertemporal elasticity of substitution.³ Agent i 's intertemporal utility at time t is thus defined recursively by:

$$\ln U_t^i = \max_{l_t^i, c_t^i} \left\{ (1 - \rho) [\ln c_t^i - \delta (l_t^i)^\eta] + \rho \ln (\mathbf{E}_t [(U_{t+1}^i)^r])^{1/r} \right\}. \quad (2)$$

The degree of relative risk aversion to lotteries over U_{t+1}^i is $1 - r \geq 0$. When $r = 0$ the second term in (2) becomes $\rho \mathbf{E}_t [\ln U_{t+1}^i]$, and utility reduce to (1). When $r \neq 0$ preferences are not time-separable, and agents care about both the magnitude of uncertainty and the timing of its resolution.^{4,5} In period or generation t , agent i maximizes his intertemporal utility (2) subject to:

$$y_t^i = (h_t^i)^\lambda (l_t^i)^\mu \quad (3)$$

³This last assumption is ubiquitous in the literature on income distribution dynamics –generally in conjunction with a myopic bequest motive rather than the dynastic one assumed here. See Glomm and Ravikumar (1992), Galor and Zeira (1993), Banerjee and Newman (1993), Saint-Paul and Verdier (1993), Bénabou (1996a), (1996b), Aghion and Bolton (1997) or Gradstein and Justman (1997). In this model, however, it *will not* result in a constant (policy-invariant) savings rate, due to the progressivity of the tax schemes that will be considered.

⁴As explained in Weil (1990), agents prefer early resolution when their aversion to risk is larger than their aversion to intertemporal fluctuations, i.e. when $1 - r > 1$. Conversely, they prefer late resolution when $r > 0$.

⁵With respect to labor supply, the model easily extends to the more general specification of felicity $u = \ln c - v(l)$, $v', v'' > 0$. Glomm and Ravikumar (1992) and Gradstein and Justman (1997) use $v(l) = -\ln(1 - l)$. The specification (1)–(2) is more flexible, as the elasticity of labor supply is parametrized by $1/(\eta - 1)$. Specifically, the intertemporal elasticity of substitution for labor is $1/(\eta - 1)$, and in an intratemporal (or steady-state) context the compensated elasticity of labor supply given a wage w and non-wage income R is $1/[\eta - 1 + wl/(R + wl)]$.

$$\hat{y}_t^i = c_t^i (1 + \theta_t) + e_t^i \quad (4)$$

$$h_{t+1}^i = \kappa \cdot \xi_{t+1}^i \cdot (h_t^i)^\alpha ((1 + a_t) e_t^i)^\beta. \quad (5)$$

Human capital h_t^i is combined with labor l_t^i to produce output.⁶ The resulting *pre-tax* income y_t^i is then subject to taxes and transfers, resulting in a *disposable* income denoted \hat{y}_t^i . The government also taxes consumption c_t^i at the rate θ_t and subsidizes investment in education e_t^i at the rate a_t . These three dimensions of fiscal policy and the constraints linking them will be examined in more detail later on. Equation (4) reflects the absence of a credit market to finance human capital investment, requiring both consumption and education expenditures to come out of disposable income.⁷ Equation (5) describes a child's human capital h_{t+1}^i as the product of three inputs: innate ability ξ_{t+1}^i , the quality of the home or neighborhood environment as measured by parental human capital h_t^i , and purchased educational inputs such as teacher time, classrooms, books or computers, $(1 + a_t) e_t^i$. The uninsurable ability or productivity shocks ξ_t^i are i.i.d. and log-normally distributed: $\ln \xi_t^i \sim \mathcal{N}(-s^2/2, s^2)$, hence $E[\xi] = 1$.⁸ I shall also assume that $\ln h_0^i \sim \mathcal{N}(m_0, \Delta_0^2)$ and –without prejudging the distribution of $\ln h_t^i$, which is endogenous– denote its mean as m_t and its variance as Δ_t^2 . Finally, let $\eta \geq 1 > \rho$ and $\alpha + \beta\lambda < 1$.

The absence of any intertemporal trade is clearly an oversimplified (but quite common) representation of asset market incompleteness; it represents the main price of analytical tractability in the model. Intratemporal linkages between agents, on the other hand, could easily be incorporated –for instance a labor market with different skill levels being complements in production. I shall nonetheless not pursue this extension, in order to better on focus the interactions –both intra- and intertemporal– which arise through public policy.

⁶More generally, h_t^i could be any non-traded asset. Note that (3) should be interpreted as income net of the remuneration of physical capital (which can be collateralized), as in Barro, Mankiw and Sala-i-Martin (1995). Thus if the production function is $Y_t^i = (h_t^i)^\omega (k_t^i)^{\omega'} (l_t^i)^{\omega''}$, equating the marginal product of capital to a fixed world interest rate r makes $y_t^i \equiv Y_t^i - rk_t^i$ proportional to $(h_t^i)^\lambda (l_t^i)^\mu$, with $\lambda \equiv \omega/(1 - \omega')$ and $\mu \equiv \omega''/(1 - \omega')$.

⁷The simplest source of such incompleteness is the fact that children cannot be held responsible for the debts incurred by their parents. The human capital accumulation on which the model focuses thus corresponds best to early childhood, elementary and secondary schooling. It is much less relevant for college and beyond, where loans (both public and private) are more readily available, at least in developed countries. Note finally that the education expenditures e_t^i may be incurred directly, as with private school tuitions, or indirectly, in the form of property taxes and land rents conditioning access to a community's public schools.

⁸The model can be extended to serially correlated shocks ξ_t^i , say $ar(1)$. But it is much simpler, and qualitatively similar, to replace the resulting $ar(2)$ process by an $ar(1)$ with higher persistence, by increasing α .

1.2 Progressive Taxation

The government redistributes income using marginally *progressive* taxes and transfers, as is the case in most countries. I use the same log-linear scheme as in Bénabou (1996b) (1996c), so that disposable income is:

$$\hat{y}_t^i = (y_t^i)^{1-\tau_t} (\tilde{y}_t)^{\tau_t}, \quad (6)$$

with \tilde{y}_t defined by the balanced-budget constraint:

$$\int_0^1 (y_t^i)^{1-\tau_t} (\tilde{y}_t)^{\tau_t} di = y_t, \quad (7)$$

where y_t is per-capita income. The elasticity τ_t of disposable to market income measures the rate of (residual) progressivity in fiscal policy. Denoting $T(y_t^i) \equiv \hat{y}_t^i - y_t^i$ the net tax paid (after transfers) at income level y_t^i , note that both average and marginal rates are rising when $\tau_t > 0$.⁹ Furthermore, τ_t is equal to the income-weighted *average marginal tax (and transfer) rate*:

$$\int_0^1 T'(y_t^i) \cdot (y_t^i/y_t) di = \tau_t. \quad (8)$$

Incentive-compatibility precludes $\tau_t > 1$, but nothing in principle prevents a regressive tax, $\tau_t < 0$.

1.3 The Shadow Value of Human Capital

Taking as given the policy sequence $\{\tau_t, \theta_t, a_t\}_{t=0}^\infty$, an agent with human wealth h solves the dynamic programming problem:

$$\begin{aligned} \ln U_t(h) &= \max_{l, \nu} \left\{ (1-\rho) [\ln((1-\nu)/(1+\theta_t)) + (1-\tau_t)(\lambda \ln h + \mu \ln l) + \tau_t \ln \tilde{y}_t - \delta(l)^\eta] \right. \\ &\quad \left. + \rho \ln(\mathbf{E}_t[(U_{t+1}(h')^r)]^{1/r}) \right\}, \end{aligned} \quad (9)$$

$$h' = \kappa \xi ((1+a_t)\nu)^\beta (h)^{\alpha+\beta\lambda(1-\tau_t)} (l)^{\beta\mu(1-\tau_t)}. \quad (10)$$

Clearly, optimal decisions will depend on the private marginal value of human capital, or equivalently on the elasticity $V_t \equiv \partial \ln U_t / \partial \ln h$. This shadow value will reflect future expected rates

⁹Note also that agents with average income are made better off ($\tilde{y}_t > y_t$) if and only if $\tau_t > 0$. A similar “constant residual progression” scheme turns out to have been used in a couple of earlier (and static) models, to study insurance or risk-taking (e.g., Feldstein (1969), Kanbur, (1979), Persson (1983)).

of redistribution, both directly and through their impact on the intergenerational persistence of human wealth, given by (10):

$$p(\tau_t) \equiv \alpha + \beta\lambda(1 - \tau_t). \quad (11)$$

The sequence of $p(\tau_t)$'s, which can also be thought of as inverse measures of social mobility, will play a fundamental role throughout the model. In particular, one shows:

Proposition 1 *The value function under fiscal redistributions is $\ln U_t^i = V_t \cdot (\ln h_t^i - m_t) + W_t$, where*

$$V_t = \lambda(1 - \rho) \sum_{s=0}^{\infty} \rho^s (1 - \tau_{t+s}) \prod_{k=0}^{s-1} p(\tau_{t+k}) \quad (12)$$

and W_t , which measures aggregate welfare in period t , is given in the appendix as a function of $\{\tau_{t+k}, \theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$.¹⁰

Note how the marginal value reflects current and future rates of redistribution, but is invariant to proportional consumption taxes and investment subsidies (a feature of logarithmic utility).

1.4 Labor Supply and Savings Decisions

The complete solution to the agent's problem is easily obtained from (12) and the first-order conditions associated to (9)–(10). I first consider labor supply.

Proposition 2 *Agents choose in every period a common level of effort, which decreases with current and expected future tax rates $\{\tau_{t+k}\}_{k=0}^{\infty}$:*

$$l_t = [(\mu/\delta\eta)(1 - \tau_t)(1 + \rho(1 - \rho)^{-1}\beta V_{t+1})]^{1/\eta}, \quad (13)$$

where V_{t+1} is defined by (12). Under a constant tax profile $\tau_t = \tau$, in particular:

$$l = \left(\frac{(\mu/\delta\eta)(1 - \tau)(1 - \rho\alpha)}{1 - \rho(\alpha + \beta\lambda(1 - \tau))} \right)^{1/\eta}. \quad (14)$$

Recall that $1/(\eta - 1)$ is the intertemporal elasticity of substitution in labor supply, with respect to variations in the real wage. The first result shows that the uncompensated elasticity

¹⁰Recall that, by definition, $m_t \equiv \int_0^1 \ln h_t^i di$; therefore $W_t = \int_0^1 \ln U_t^i di$. Throughout that paper I shall use the convention that $\prod_{k=s}^{s'} x_k \equiv 1$ and $\sum_{k=s}^{s'} x_k \equiv 0$ whenever $s' < s$, for any sequence of x_k 's.

with respect to $1 - \tau_t$, the net-of-tax (progressivity) rate, equals $1/\eta$. The second result makes transparent the role of the other parameters, and will be useful when focusing on steady-states later on. I now turn to agents' propensity to save out of disposable income.

Proposition 3 *Agents choose in every period a common savings rate, $\nu_t = e_t^i/\hat{y}_t^i$, which decreases with expected future tax rates $\{\tau_{t+k}\}_{k=0}^\infty$:*

$$\nu_t = \frac{\rho\beta V_{t+1}}{1 - \rho + \rho\beta V_{t+1}}, \quad (15)$$

where V_{t+1} is defined as before. Under a constant tax profile, $\tau_t = \tau$, in particular:

$$\nu = \frac{(1 - \tau)\rho\beta\lambda}{1 - \rho\alpha} \equiv (1 - \tau) \mathfrak{s}. \quad (16)$$

where \mathfrak{s} is the *laissez-faire* savings rate.

These results, and in particular the second one, show clearly how the progressive taxation of income (a constant share λ of which is derived from human capital) distorts private savings decisions. Income taxes, however, are not the only fiscal instrument available to policy-makers and the voters who elect them. In particular, consumption taxes and investment subsidies provide a means to correct intertemporal distortions.

1.5 Consumption Taxes and Investment Subsidies

Recall that the government taxes consumption at the rate θ_t and subsidizes education at the rate a_t , subject to the budget constraint:

$$\theta_t \int_0^1 c_t^i di = a_t \int_0^1 e_t^i di,$$

or, by Proposition 3,

$$\frac{\theta_t(1 - \nu_t)}{1 + \theta_t} = a_t \nu_t. \quad (17)$$

Given a savings rate ν_t and a relative price of education or similarly credit-constrained investment goods $1/(1 + a_t)$, each agent's actual investment rate is $(1 + a_t)\nu_t$. It can thus be restored to its (credit-constrained) optimal level $\mathfrak{s} \equiv \rho\beta\lambda/(1 - \rho\alpha)$ by a consumption tax of

$$\theta_t = \frac{\mathfrak{s} - \nu_t}{1 - \mathfrak{s}}, \quad (18)$$

whose proceeds are used to subsidize education. Moreover:

Proposition 4 *For any sequence of current and future rates of redistribution $\{\tau_{t+k}\}_{k=s}^{\infty}$, let $\{\theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$ be the unique corresponding sequence of consumption tax rates and investment subsidies such that, in every period $t+k$:*

(i) *the government budget is balanced, as described by (17);*

(ii) *agents' common investment rate is restored to its first-best level \mathfrak{s} , as described by (18).*

Every agent i of every generation t prefers the policy sequence $\{\tau_{t+k}, \theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$ to any alternative $\{\tau_{t+k}, \theta'_{t+k}, a'_{t+k}\}_{k=0}^{\infty}$.

This unanimity result means that while the policy space is two-dimensional (taking into account the budget constraint), the *Pareto set* is one-dimensional: distributional conflict concerns only the degree of progressivity $\{\tau_{t+k}\}_{k=0}^{\infty}$. Accordingly, I will from here on restrict attention to undominated policy mixes, setting $(1 + a_t) \nu_t = \mathfrak{s}$ for all t .

Given a reasonably broad menu of fiscal instruments, redistributive taxation thus causes only intratemporal distortions, namely those to labor supply.¹¹ This result is consistent with the empirical evidence from cross-country regressions, surveyed in Bénabou (1996c). There is no sign of a negative effect of redistribution (shares of various transfers in GDP, average and marginal tax rates) on national investment rates rates. In fact, the regression coefficient is more often positive than not. By contrast, there is a positive association between labor tax rates and national unemployment rates (Daveri and Tabellini (1997)).

If one wanted nonetheless to maintain intertemporal distortions in the model, one would simply constrain a_t and θ_t to zero. This might conceivably reflect the under-developed fiscal system of a poor country, or some informational constraints which make subsidies to human capital investment

¹¹The policy mix described in Proposition 4 is ultimately equivalent to using a progressive consumption tax to finance a program of progressive education subsidies (with the same rate τ_t) of the type studied in the next section. Because it expropriates only labor and existing human wealth, such a policy generates only effort distortions. Proposition 4 is thus related –but in an incomplete markets setting and with progressivity– to the classical public finance results about the superiority of consumption taxes over (capital plus labor) income taxes. See Atkinson and Stiglitz (1980), Chamley (1985) and Judd (1985).

impractical. Throughout the rest of the paper, all one would need to do is to replace \mathfrak{s} by ν_t . As shown by (16) this is particularly easy in steady-state, which will be our ultimate focus.

2 Education Finance

As an alternative to progressive income taxes and transfers, I consider the redistribution of education expenditures (mainly from early childhood to the secondary level). This may correspond to a policy of school funding equalization, such as those which have been mandated by constitutional courts in many U.S. states, or more generally of *subsidizing differentially* the education expenditures of rich and poor families or communities. Formally, let income itself remain untaxed, $\hat{y}_t^i = y_t$, while in (5) educational investment $(1 + a_t) e_t^i$ is replaced by

$$\hat{e}_t^i = (1 + a_t) (\tilde{y}_t / y_t^i)^{\tau_t} e_t^i. \quad (5')$$

This means that a family with income y_t^i faces the effective price $p_t^i = (1 + a_t)^{-1} (y_t^i / \tilde{y}_t)^{\tau_t}$ for education. The progressivity rate τ_t can also be thought as the degree of equalization of school inputs, while a_t still represents the average rate of education subsidization. Indeed, with all agents saving a fraction ν_t of their income (as shown below), $\hat{e}_t^i = (1 + a_t) \nu_t (y_t^i)^{1-\tau_t} (\tilde{y}_t)^{\tau_t}$. The polar cases of $\tau_t = 0$ and $\tau_t = 1$ correspond respectively to decentralized and egalitarian school funding, with most states in practice falling somewhere in-between.¹² Summing across agents, the net subsidy is $a_t \nu_t y_t$, to be funded as previously by a consumption tax.

The Bellman equation is now:

$$\begin{aligned} \ln U_t(h) = & \max_{l, \nu} \left\{ (1 - \rho) [\ln((1 - \nu)/(1 + \theta_t)) + \lambda \ln h + \mu \ln l_t - \delta (l)^\eta] \right. \\ & \left. + \rho \ln(\mathbf{E}_t[(U_{t+1}(h')^r)]^{1/r}) \right\}, \end{aligned} \quad (19)$$

with h' still given by (10). The solution has a similar structure to that obtained for fiscal redistributions.

¹²Some of the distributional and efficiency properties of these two polar regimes have been analyzed in Loury (1981), Glomm and Ravikumar (1992), Bénabou (1996a) (1996b), Gradstein and Justman (1997) and Fernandez and Rogerson (1998). Note that the redistributive scheme described in (5') is less detrimental to investment incentives than if educational budgets were being redistributed directly (e.g., $\hat{e}_t^i = (1 + a_t)(e_t^i)^{1-\tau_t}(\tilde{e}_t)^{\tau_t}$), as is the case in some US states such as California (see Hoxby (1998)).

Proposition 5 *The value function under progressive school finance is $\ln U_t^i = V_t \cdot (\ln h_t^i - m_t) + W_t$, where*

$$V_t = \lambda(1 - \rho) \sum_{s=0}^{\infty} \rho^s \prod_{k=0}^{s-1} p(\tau_{t+k}) \quad (20)$$

and aggregate welfare W_t is given in the appendix as a function of $\{\tau_{t+k}, \theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$.

The only difference with (12) is the absence of the factor $1 - \tau_{t+s}$ multiplying each discounted product term. As a result, human capital is more valuable (V_t is higher), for all non-negative sequences $\{\tau_{t+k}\}_{k=0}^{\infty}$. This reflects the fact that progressive redistribution now applies only to the fraction of income which is saved, and not to that which is consumed. As one would expect, this lessens both inter- and intratemporal distortions.

Proposition 6 *For any expected sequence of education finance equalization rates $\{\tau_{t+k}\}_{k=0}^{\infty}$, agents' common savings rate ν_t is still given by (15), but with V_t now defined by (20). In particular, under a constant education finance policy $\tau_t = \tau$,*

$$\nu = \frac{\rho\beta\lambda}{1 - \rho\alpha + \rho\beta\lambda\tau} \equiv \frac{\mathfrak{s}}{1 + \tau\mathfrak{s}}. \quad (21)$$

As to agents' labor supply, it is now

$$l_t = [(\mu/\delta\eta)(1 + \rho(1 - \rho)^{-1}(1 - \tau_t)\beta V_{t+1})]^{1/\eta}, \quad (22)$$

so that under a constant τ ,

$$l = \left(\frac{(\mu/\delta\eta)(1 - \rho\alpha)}{1 - \rho(\alpha + \beta\lambda(1 - \tau))} \right)^{1/\eta}. \quad (23)$$

Note that ν_t and l_t remain positive even for $\tau_{t+k} \equiv 1$, which corresponds in equilibrium to uniform funding of education, $\hat{e}_t^i = (1 + a_t)\nu_t y_t$. In particular, steady-state effort remains bounded below even as the intertemporal elasticity of substitution $\epsilon \equiv 1/(\eta - 1)$ becomes infinite.

As in the case of income taxes, the decline in savings can be fully offset by taxing consumption at the rate θ_t given by (18), and using the proceeds to finance the net (or average) education subsidy a_t which restores the investment rate to \mathfrak{s} . Since the distortion to ν_t is now smaller, the required rates of θ_t and a_t are lower. Moreover, conditional on any $\{\tau_{t+k}\}_{k=0}^{\infty}$ this policy will once again be supported unanimously, both within and across generations. As explained above, the

remainder of the paper will incorporate this Pareto-improving policy mix, but if it were for some reason infeasible one would simply replace \mathfrak{s} by ν_t in all derivations.

3 The Dynamics of Human Wealth and Income

Let us now take logarithms in (3) and (5) in the case of income taxes, or in (3) and (5') in that of progressive education finance. Under either redistributive scheme agent i 's net investment is $\mathfrak{s}(y_t^i)^{1-\tau_t}(\tilde{y}_t)^{\tau_t}$, so the law of motion of human capital takes the form:

$$\ln h_{t+1}^i = \ln \xi_{t+1}^i + \ln \kappa + \beta \ln \mathfrak{s} + \beta \mu (1 - \tau_t) \ln l_t + (\alpha + \beta \lambda (1 - \tau_t)) \ln h_t^i + \beta \tau_t \ln \tilde{y}_t. \quad (24)$$

This implies that both h_t^i and income $y_t^i = (h_t^i)^\lambda (l_t)^\mu$ remain log-normally distributed over time. If $\ln h_t^i \sim \mathcal{N}(m_t, \Delta_t^2)$, then

$$m_{t+1} = (\alpha + \beta \lambda) m_t + \beta \mu \ln l_t + \beta \tau_t (2 - \tau_t) \lambda^2 \Delta_t^2 / 2 + \beta \ln \mathfrak{s} + \ln \kappa - s^2 / 2, \quad (25)$$

$$\Delta_{t+1}^2 = (\alpha + \beta \lambda (1 - \tau_t))^2 \Delta_t^2 + s^2, \quad (26)$$

where the first equation is obtained by substituting into (24) the break-even level of income

$$\ln \tilde{y}_t = \ln y_t + (1 - \tau_t) \lambda^2 \Delta_t^2 / 2 = \lambda m_t + \mu \ln l_t + (2 - \tau_t) \lambda^2 \Delta_t^2 / 2, \quad (27)$$

defined by the budget constraint (7). Finally, (25)–(26) easily yield the following results.

Proposition 7 *The distribution of income at time t is $\ln y_t^i \sim \mathcal{N}(\lambda m_t + \mu \ln l_t, \lambda^2 \Delta_t^2)$, where m_t and Δ_t^2 evolve according to the linear difference equations (25)–(26) and $l_t = l(\tau_t)$ is given by Proposition 2 or 6. The growth rate of per capita income equals*

$$\ln y_{t+1} - \ln y_t = \ln \tilde{\kappa} - (1 - \alpha - \beta \lambda) \ln y_t + \mu (\ln l_{t+1} - \alpha \ln l_t) - \mathfrak{L}(\tau_t) \lambda^2 \Delta_t^2 / 2, \quad (28)$$

where $\ln \tilde{\kappa} \equiv \lambda (\ln \kappa + \beta \ln \mathfrak{s}) - \lambda (1 - \lambda) s^2 / 2$ is a constant and

$$\mathfrak{L}(\tau) \equiv \alpha + \beta \lambda (1 - \tau)^2 - (\alpha + \beta \lambda (1 - \tau))^2 > 0 \quad (29)$$

measures the extent to which income inequality slows down growth.

While Propositions 2 to 6 dealt with the costs of redistribution, the above results bring to light some of the benefits. From the point of view of equity, equations (24) and (26) show that a higher τ_t works to reduce both the persistence $p(\tau_t) = \alpha + \beta\lambda(1 - \tau_t)$ and the magnitude Δ_{t+1}^2 of disparities in human capital and income. From the point of view of efficiency, redistribution provides a partial substitute for the missing credit market. This *investment reallocation* effect is reflected in the last term of equation (28), which measures the shortfall in per capita income growth compared to that of a representative agent economy. Because decreasing returns and credit constraints imply that poorer families have a higher marginal return than wealthier ones, redistributing education resources directly or indirectly (through income taxes) reduces this loss, but only up to a point: $\mathfrak{L}(\tau)$ is minimized for $\tau = 1 - \alpha/(1 - \beta\lambda)$. The more important the *complementary inputs* provided by families and communities, i.e. the greater is α , the less redistribution is called for –at least in the short run, where Δ_t^2 is given.

Having solved for the economy’s aggregate and distributional dynamics under any policy profile $\{\tau_t\}_{t=0}^\infty$ I shall now explore –first analytically, then quantitatively–how fiscal and educational redistributions affect total income, inequality, risk–sharing and welfare. The exposition of these policy tradeoffs will often focus on steady–states, but the results are derived more generally.

3.1 Steady–State Income, Inequality and Redistribution

- *Asymptotic Values.* Given a constant rate of fiscal progressivity or school finance equalization τ , income inequality converges to $\lambda^2\Delta^2(\tau)$, with

$$\Delta^2(\tau) \equiv \frac{s^2}{1 - (\alpha + \beta\lambda(1 - \tau))^2}, \quad (30)$$

and long–run income per capita is:

$$\ln y(\tau) \equiv \frac{\ln \tilde{\kappa} + \mu(1 - \alpha) \ln l(\tau) - \mathfrak{L}(\tau)\lambda^2\Delta^2(\tau)/2}{1 - \alpha - \beta\lambda}. \quad (31)$$

Redistribution but has two opposing effects on long–run income. On one hand, it reduces labor supply $l(\tau)$; it would also depress savings, were it not for the offsetting effect of consumption taxes and investment subsidies. If those are for some reason infeasible, one simply replaces \mathfrak{s} by $\mathfrak{s}(1 - \tau)$ or $\mathfrak{s}/(1 + \tau\mathfrak{s})$ in the term $\ln \tilde{\kappa}$. The other effect is to alleviate the misallocation of education

resources due to credit constraints: by reducing $\mathfrak{L}(\tau)$, up to some point, as well as $\Delta^2(\tau)$, a positive τ tends to raise $y(\tau)$. The degree of progressivity τ_Y^* which maximizes long-run output is easily seen to: (a) decrease with the labor supply elasticity $1/\eta$, the share of labor in production μ , and the discount factor ρ ; (b) increase with the variability of shocks to ability or human wealth, s^2 . The effects of α and $\beta\lambda$ are generally ambiguous, and will be explored quantitatively.

- *Endogenous Growth.* The above results are easily transposed from long-run levels to *long-run growth rates*, by allowing for knowledge spillovers in the accumulation of human capital. This can be done in a “heterogeneity-neutral” manner (that is, without introducing additional costs or benefits of redistribution), by replacing the constant κ in (5) with the human capital index $\kappa_t \equiv \left(\int_0^1 (h_t^i)^\lambda di \right)^{\gamma/\lambda}$. All previous results remain unchanged, with κ_t simply substituted for κ everywhere. In steady-state, the only difference is that the denominator in (31) is now $1 - \alpha - \beta\lambda - \gamma$. For $\alpha + \beta\lambda + \gamma = 1$ the numerator gives the economy’s asymptotic growth rate, which behaves with respect to τ exactly as $\ln y(\tau)$ did in the original specification.¹³

- *Redistribution in Steady-State.* A tax rate τ yields inequality $\lambda\Delta(\tau)$ in earnings, but only $(1-\tau)\lambda\Delta(\tau)$ in disposable incomes. When τ is a rate of school finance progressivity, as in Section 2, this narrowing operates only on education spending. To assess the extent of redistribution implied by either policy, recall from (27) that the threshold \tilde{y} separating losers and gainers is always given by $\ln(\tilde{y}/y) = (1 - \tau)\lambda^2\Delta^2(\tau)/2$.¹⁴ This corresponds to a rank in the income distribution of $r(\tau) \equiv \Phi[(2 - \tau)\lambda\Delta(\tau)/2]$, where Φ denotes the c.d.f. of a standard normal. Moreover:

Proposition 8 *The Lorenz curve for a log-normal distribution with variance ω^2 is*

$$\Gamma(r; \omega) = \Phi(\Phi^{-1}(r) - \omega), \text{ for all } r \in [0, 1].$$

In the steady-state resulting from a rate of fiscal progressivity τ , the $1-r(\tau)$ richest households constituting the net tax base thus earn

¹³The reason why the above definition of κ_t is heterogeneity-neutral is that it aggregates individual human capital contributions with the same elasticity of substitution as total output. Alternative spillovers, embodying social costs or benefits of heterogeneity in human capital interactions, can easily be dealt with by letting the elasticity of substitution in κ_t be, respectively, smaller or greater than $1/(1 - \gamma)$. See Bénabou (1996a).

¹⁴By losers and gainers I mean: (i) families paying a net tax and those receiving a net transfer, when τ describes fiscal policy; (ii) families whose education expenditures are subsidized beyond the average rate $a(\tau)$ and those whose expenditures are taxed relative to $a(\tau)$, when τ describes school finance policy.

$$1 - \Gamma(r(\tau); \lambda\Delta(\tau)) = 1 - \Phi((2 - \tau)\lambda\Delta(\tau)/2 - \lambda\Delta(\tau)) = \Phi(\tau\lambda\Delta(\tau)/2) \quad (32)$$

of total pre-tax income. After redistribution their share falls to

$$1 - \Gamma(r(\tau); (1 - \tau)\lambda\Delta(\tau)) = 1 - \Phi((2 - \tau)\lambda\Delta(\tau)/2 - (1 - \tau)\lambda\Delta(\tau)) = \Phi(-\tau\lambda\Delta(\tau)/2). \quad (33)$$

The share of net transfers in national income is therefore:

$$b(\tau) \equiv 2\Phi(\tau\lambda\Delta(\tau)/2) - 1. \quad (34)$$

When redistribution occurs only in school expenditures, the top $1 - r(\tau)$ families' shares of disposable income, consumption and personal education spending are given by (32). The proportion of school inputs actually allocated to their children, however, is reduced to (33). Thus (34) now describes the share of the total educational budget which is transferred from those above \tilde{y} to those below it. Multiplying this number by \mathfrak{s} translates it into a percentage of total income.

4 A Criterion of Aggregate Economic Efficiency

Aggregate income growth is important from a macroeconomic perspective, but provides only an incomplete picture of the efficiency implications of redistributive policies. First, it omits the value of leisure and other non-market activities; this could be remedied by looking at the aggregate consumption-leisure bundle. More fundamentally, it fails to reflect redistribution's role as social insurance: by the law of large numbers, individual shocks cancel out when computing aggregate quantities.

Policies are most often evaluated according to some social welfare criterion such as $W_0 = \int_0^1 \ln U_0^i di$ or $T_0 \equiv \int_0^1 U_0^i di$. Risk and effort concerns are now properly embodied, and I shall indeed compute such utility aggregates. But the problem is that whereas *aggregate income underestimates* the efficiency value of redistribution, *aggregate welfare overestimates* it. Because of the concavity of individual utility, any such utilitarian criterion rises (keeping labor supply and savings fixed) with all current and future redistributions, even when there are no shocks to be insured against and no credit-constrained investments in need of reallocation.¹⁵

¹⁵For instance, in our model, even when $s = 0$ and either $(\alpha, \beta\lambda) = (0, 0)$ (no accumulation) or $\alpha + \beta\lambda = 1$ (accumulation with constant returns to investment).

I shall therefore propose an alternative measure of pure *aggregate economic efficiency*, which puts *zero value on equity* of consumption or income per se, and is affected by redistribution only to the (full) extent that such policies: (i) distort effort and savings decisions; (ii) relax the liquidity constraints which impede growth; (iii) reduce the idiosyncratic risk faced by individuals. The basic idea is very simple:

- First, replace agents' stochastic consumptions sequences with appropriate *certainty-equivalents*.

In this aggregation over states, the relevant parameter is the degree of risk-aversion.

- Second, *aggregate linearly* individuals' certainty-equivalent consumptions, which are thus treated as perfect substitutes. More generally, when aggregating over individuals the relevant parameter should be society's degree of inequality-aversion, which here is set to zero.

- Finally, aggregate *over time* using agents' common discount rate and intertemporal elasticity of substitution.

The point of this construction is not that this is “the right” social welfare function, nor that society should not care about equity. In fact, the next section will explicitly incorporate inequality-aversion into the analysis. The point is instead that one should be able to *separate* efficiency concerns –namely, the extent to which market distortions and imperfections are worsened or improved by policy– from pure equity concerns. For instance, inequality of initial endowments should not affect a measure of pure efficiency, unless wealth affects an agent's ability to invest or bear risk. Conversely, redistributions should affect an index of total efficiency only to the extent that they change the “size of the pie” (the path of aggregate consumption) or the riskiness of individual “slices”. We shall see that the index proposed above has these properties, whereas none of the usual (utilitarian) social welfare functions, such as W_0 or T_0 , do. (The latter also applies to standard total compensating variation). These points will be established in the context of the present model, but the underlying idea is clearly more general.

I shall now proceed to compute the efficiency criterion, focusing the exposition on redistribution through fiscal policy. The case of redistributive education finance is very similar; it is treated in the appendix, and the main results are presented at the end of the next section.

By Proposition 1, the intertemporal utility of agent i in period zero can be written as:

$$\ln U_0^i = (1 - \rho) \left(\sum_{t=0}^{\infty} \rho^t (1 - \tau_t) \prod_{k=0}^{t-1} p(\tau_k) \right) \lambda (\ln h_0^i - m_0) + \sum_{t=0}^{\infty} \rho^t (W_t - \rho W_{t+1}).$$

The first term captures the lasting effects of differences in initial endowments. The second represents the part that is common to all agents, including labor supply and risk concerns. The same level of intertemporal utility would clearly result from the deterministic, or *certainty-equivalent* consumption sequence $\{\bar{c}_t^i\}_{t=0}^\infty$ defined by:

$$\ln \bar{c}_t^i - \delta(l_t)^\eta \equiv (1 - \tau_t) \left(\prod_{k=0}^{t-1} p(\tau_k) \right) \lambda (\ln h_0^i - m_0) + (W_t - \rho W_{t+1}) / (1 - \rho), \quad (35)$$

with unchanged efforts $\{l_t\}_{t=0}^\infty$. Moreover, it is shown in the appendix that \bar{c}_t^i simplifies to:

$$\ln \bar{c}_t^i = \mathbf{E}_0 [\ln c_t^i | h_0^i] + r \left(\frac{\rho}{1 - \rho} \right) \left(\frac{V_{t+1}^2 s^2}{2} \right). \quad (36)$$

In the absence of shocks, $\bar{c}_t^i = c_t^i$. When $r = 0$, \bar{c}_t^i is the standard certainty-equivalent consumption given time-separable, logarithmic preferences. In the more general case there is an extra term which might be called (minus) the “resolution premium” for the shock ξ_{t+1}^i .¹⁶ It is negative when tastes favor early resolution of uncertainty ($r < 0$), positive in the reverse case.

The next step is to compute *total certainty-equivalent consumption*, then E_0 itself.

Definition 1 Let $\bar{C}_t \equiv \int_0^1 \bar{c}_t^i di$. The aggregate efficiency of a tax sequence $\{\tau_t\}_{t=0}^\infty$ is defined as:

$$E_0 \equiv (1 - \rho) \sum_{t=0}^{\infty} \rho^t [\ln \bar{C}_t - \delta(l_t)^\eta].$$

Given two policies $\{\tau_t\}_{t=0}^\infty$ and $\{\tau'_t\}_{t=0}^\infty$, $E'_0 - E_0$ can thus be expressed as a percentage difference in total consumption. The lognormality of the \bar{c}_t^i 's makes it easy to compute \bar{C}_t , and obtain:

$$E_0 = W_0 + (1 - \rho) \left(\sum_{t=0}^{\infty} \rho^t (1 - \tau_t)^2 \prod_{k=0}^{t-1} p(\tau_k)^2 \right) \left(\frac{\lambda^2 \Delta_0^2}{2} \right), \quad (37)$$

where $W_0 \equiv \int_0^1 \ln U_0^i di$.¹⁷ Thus E_0 differs from aggregate welfare by a term which increases with Δ_0 and declines with all present and future rates of redistribution. As shown below, this adjustment eliminates all effects of inequality except those relating to efficiency via market incompleteness, so that E_0 indeed satisfies properties (i) to (iii) postulated earlier. These results appear

¹⁶That shock has variance s^2 and affects the agent through h_{t+1}^i , which enters the intertemporal utility U_{t+1}^i with an elasticity equal to V_{t+1} ; see Proposition 1.

¹⁷By (35), $\ln \bar{C}_t = \int_0^1 \ln \bar{c}_t^i di + (1 - \tau_t)^2 \prod_{k=0}^{t-1} p(\tau_k)^2 \lambda^2 \Delta_0^2 / 2$, and $W_0 = \sum_{t=0}^{\infty} \rho^t \left(\int_0^1 \ln \bar{c}_t^i di - \delta(l_t)^\eta \right)$.

most clearly with time-invariant policies, but are established more generally in the appendix.

Proposition 9 (a) *The aggregate efficiency of a constant rate of progressive taxation τ equals:*

$$E_0 = (1 - \rho) \sum_{t=0}^{\infty} \rho^t [\ln y_t - \delta(l_t)^\eta] + \ln(1 - \mathfrak{s}) - \rho(1 - \tau)^2 \left(\frac{1}{1 - \rho p(\tau)^2} - \frac{r(1 - \rho)}{(1 - \rho p(\tau))^2} \right) \left(\frac{\lambda^2 s^2}{2} \right).$$

(b) *For any initial conditions (m_0, Δ_0^2) , E_0 is maximized at some strictly positive $\tau_{E,0}^*$.*

The interpretation is simple. The first term, which reduces to $\ln y(\tau) - \delta(l(\tau))^\eta$ in steady-state, captures the effects of redistribution on the path of total output, net of effort. These operate through its influence on the allocation of investment and on labor supply (plus possibly on savings, if \mathfrak{s} is replaced by $\nu(\tau)$), and were discussed earlier. Adding the second term yields the utility derived by a fictitious representative agent from aggregate consumption and leisure. More novel is the last term, which measures the disutility which agents suffer from uninsured idiosyncratic shocks –or conversely, the *insurance and uncertainty-resolution value* of redistribution. This risk premium is always positive, hence minimized for $\tau = 1$; it is larger, the greater is agents' risk-aversion $1 - r$. Part (b) of the proposition is also quite intuitive: starting from $\tau = 0$, a small tax increase generates only second-order losses from labor (and possibly savings) distortions; but due to market frictions it yields a first-order gain in insurance and in the allocation of investment resources across differentially credit-constrained families.

Finally, note that E_0 is independent of the distribution of initial endowments Δ_0^2 , except to the extent that it affects the present value of total output, through accumulation. Equation (37) makes clear that such is not the case of W_0 . It also shows that the *median voter*, whose intertemporal utility is W_0 , would always choose taxes exceeding the efficient level.

5 Efficiency, Equality, and Social Welfare

I now extend the preceding results to social welfare indices which incorporate both efficiency and equity concerns. The procedure is the same as for constructing E_0 , except that individual certainty-equivalents are aggregated with an *interpersonal* elasticity of substitution $\sigma \geq 0$, whose inverse $1/\sigma$ measures society's degree of inequality-aversion (as in Atkinson (1970)):

$$\bar{C}_t(\sigma) \equiv \left(\int_0^1 (\bar{c}_t^i)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}. \quad (38)$$

I will also compute more standard social welfare functions, which are aggregates of (intertemporal) *utilities* rather than risk-adjusted consumptions. These have the clearly desirable property that maximizing such a criterion ensures Pareto efficiency. On the other hand, it will be seen that they can not distinguish between the effects of policy which operate through its role as a substitute for missing markets, and those which reflect an implicit equity concern.

Definition 2 For any $\sigma \in \mathbb{R}$, define the two social welfare indices:

$$\begin{aligned} S_0(\sigma) &\equiv (1 - \rho) \sum_{t=0}^{\infty} \rho^t [\ln \bar{C}_t(\sigma) - \delta(l_t)^\eta], \\ T_0(\sigma) &\equiv \ln \left(\int_0^1 (U_0^i)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}. \end{aligned}$$

The log-normality of the \bar{c}_t^i 's makes it again straightforward to compute each $\bar{C}_t(\sigma)$, hence $S_0(\sigma)$.¹⁸ As to $T_0(\sigma)$, it is easily obtained by recalling that $\ln U_t^i = V_t(\ln h_0^i - m_0) + W_0$. Hence:

Proposition 10 The social welfare levels resulting from a tax sequence $\{\tau_t\}_{t=0}^{\infty}$ are equal to:

$$\begin{aligned} S_0(\sigma) &= W_0 + \left(\frac{\sigma-1}{\sigma} \right) \left((1 - \rho) \sum_{t=0}^{\infty} \rho^t (1 - \tau_t)^2 \prod_{k=0}^{t-1} p(\tau_k)^2 \right) \left(\frac{\lambda^2 \Delta_0^2}{2} \right); \\ T_0(\sigma) &= W_0 + \left(\frac{\sigma-1}{\sigma} \right) \left((1 - \rho) \sum_{t=0}^{\infty} \rho^t (1 - \tau_t) \prod_{k=0}^{t-1} p(\tau_k) \right)^2 \left(\frac{\lambda^2 \Delta_0^2}{2} \right). \end{aligned}$$

The first result, together with (37), implies that $S_0(\sigma)$ can be naturally *decomposed* into *efficiency and equity* concerns, with the latter's intensity being parametrized by $1/\sigma$:

$$S_0(\sigma) = E_0 - \left((1 - \rho) \sum_{t=0}^{\infty} \rho^t (1 - \tau_t)^2 \prod_{k=0}^{t-1} p(\tau_k)^2 \right) \left(\frac{\lambda^2 \Delta_0^2}{2\sigma} \right), \quad (39)$$

or, in the simpler case of a time-invariant policy τ :

¹⁸By (35), $\ln \bar{C}_t(\sigma) = \int_0^1 \ln \bar{c}_t^i di + \left(\frac{\sigma-1}{\sigma} \right) (1 - \tau_t)^2 \left(\prod_{k=0}^{t-1} p(\tau_k)^2 \right) \lambda^2 \Delta_0^2 / 2$.

$$S_0(\sigma) \equiv E_0 - \left(\frac{(1-\tau)^2}{1-\rho p(\tau)^2} \right) \left(\frac{\lambda^2 \Delta_0^2}{2\sigma} \right) \quad (40)$$

By contrast, there is *no value of σ* for which the utility-based criterion $T_0(\sigma)$ does not reward equity per se. Although $T_0(\sigma)$ can also be decomposed into aggregate efficiency and a pure social cost of inequality, the latter now has *two* components: $T_0(\sigma) = E_0 - (T_0(\infty) - T_0(\sigma)) - (S_0(\infty) - T_0(\infty))$. The first cost of inequality (and benefit of redistribution) arises naturally from society's aversion to disparities in welfare; given a constant τ , it is proportional to $(1-\tau)^2 \lambda^2 \Delta_0^2 / 2\sigma$. The second one, proportional to $(1-\tau)^2 \lambda^2 \Delta_0^2 / 2$, arises mechanically and inevitably from the concavity of individual preferences.¹⁹

The final observation is that the utilitarian criterion W_0 belongs to both families of indices, and therefore combines their two defining properties: exact decomposability and Pareto-compatibility. But it arbitrarily equates society's degree of *inequality aversion* $1/\sigma$, not even with individuals' risk aversion $1-r$, which might perhaps make sense in an ex-ante "veil-of-ignorance" perspective, but with the inverse of their (unitary) *intertemporal elasticity* of substitution.

In studying efficiency and its relation to aggregate output and social welfare, I have so far concentrated on policies of income redistribution. The case of progressive education finance is treated similarly in Section C of the appendix, where I show:

Proposition 11 *The economic efficiency and social welfare indices E_0 , $S_0(\sigma)$ and $T_0(\sigma)$ resulting from a sequence of education finance progressivity rates $\{\tau_t\}_{t=0}^\infty$ are given by the same expressions as in equations (37)–(40) and Propositions 9–10, except that the terms $(1-\tau_t)^2$ and $(1-\tau)^2$ are replaced by 1.*

The absence of the terms in $(1-\tau_t)^2$ reflects, once again, the fact that parental consumption is not expropriated. This was shown to reduce both intra- and intertemporal distortions, compared to the case of income taxes. The counterpart, as made clear by Proposition 11, is that redistribu-

¹⁹Recall that $U_0^i = (h_0^i)^{V_0} e^{-V_0 \ln m_0 + W_0}$, with $V_0 \in [0, 1]$. The fact that $S_0(\infty) - T_0(\infty)$ is indeed a positive welfare cost (whether or not τ is constant) follows from:

$$\frac{S_0(\infty) - T_0(\infty)}{\lambda^2 \Delta_0^2 / 2} = \left((1-\rho) \sum_{t=0}^{\infty} \rho^t (1-\tau_t)^2 \prod_{k=0}^{t-1} p(\tau_k)^2 \right) - \left((1-\rho) \sum_{t=0}^{\infty} \rho^t (1-\tau_t) \prod_{k=0}^{t-1} p(\tau_k) \right)^2 \geq 0,$$

since $(1-\rho) \left(\sum_{t=0}^{\infty} \rho^{2t/2} x_t \right)^2 \leq (1-\rho) \left(\sum_{t=0}^{\infty} \rho^t \right) \left(\sum_{t=0}^{\infty} \rho^t x_t^2 \right) = \left(\sum_{t=0}^{\infty} \rho^t x_t^2 \right)$ for all $x_t \geq 0$, by Schwartz's inequality.

tional education finance offers *less risk-sharing*. For instance, with income taxation $\tau = 1$ yields full consumption insurance (hence $E_0 = (1 - \rho) \sum_{t=0}^{\infty} \rho^t [\ln y_t - \delta(l_t)^{\eta}] + \ln(1 - \mathfrak{s})$); but when only educational inputs are equalized, individuals families remain exposed to significant risk .

6 Quantitative Analysis

That redistribution may generate output and efficiency gains when insurance and credit markets are incomplete has been understood for quite some time. Yet there has been little attempt to evaluate the size of these gains and compare them to the losses from distortionary taxation. Two notable exceptions are Inman (1978) and Fernandez and Rogerson (1998). Inman examines how different education finance schemes affect a social welfare function in a static model. Calibrating an overlapping generations model with educational bequests to U.S. data, Fernandez and Rogerson (1998) find that moving from local to state finance would increase steady-state output by about 3%. In both models the distortions induced by redistributive policies affect utilities but not incomes, as there is no labor supply or personal savings decision. Also absent are family or social inputs into education, which were seen earlier to potentially reduce the efficiency of redistributive schemes (and also clearly matter for mobility). Neither is there any role for insurance, which has the reverse effect.

The present framework incorporates these elements into a more comprehensive analysis of the *efficiency costs and benefits of redistribution*, be it fiscal or educational. Given the model's simplicity and the fact that the empirical literature provides only imprecise estimates of certain key parameters, however, the results should still be read as a only broad assessment of the main tradeoffs. Two points in particular should be kept in mind. On one hand, credit and insurance markets are completely absent from the model, rather than simply imperfect; there is also no precautionary savings. These factors all tend to overstate the benefits of redistribution, so I will compensate by using conservative values for agents' risk-aversion and the effectiveness of educational expenditures. On the other hand, by focusing on the efficiency criterion E_0 one abstracts from many potential sources of losses from inequality. These include of course pure equity concerns such as those embodied in standard social welfare functions, but also production complementarities (as in Tamura (1992) or Bénabou (1996a)), political instability, crime and other

forms of rent-seeking. These omissions all tend to understate the benefits of redistribution.²⁰

6.1 Parameter Values

- *Production.* The shares λ and μ of human capital and labor are determined by “maximizing out” physical capital, which is not subject to borrowing constraints, from a three-factor production function.²¹ This follows Barro, Mankiw and Sala-i-Martin (1995) and, like them, I use the shares estimated by Jorgensen, Gollop and Fraumeni (1987): .5 for human capital, .3 for physical capital, and .2 for labor. This yields $\lambda = .5/.8 = .625$ and $\mu = .3/.8 = .375$.

- *Accumulation.* Most estimates of intergenerational persistence $p(\tau) \equiv \alpha + \beta\lambda(1 - \tau)$ range from about .3 to .55 (Solon (1992), Zimmerman (1992), Cooper, Durlauf and Johnson (1996)). Mulligan (1995) finds values of .5 to .7 for family income, and up to .8 for consumption. I set $\alpha = .35$ and $\beta = .4$, which allows $p(\tau)$ to range from .35 to .60. The chosen value of β implies an elasticity of children’s income to education expenditures of $\beta\lambda = .25$. This is slightly above the value of .19 used by Fernandez and Rogerson (1998), but well below those of .35 and .45 used by Hendricks (1998) and Jones, Manuelli and Rossi (1993), respectively.²² Fernandez and Rogerson base their choice on the studies of Card and Krueger (1992), Johnson and Stafford (1973) and Wachtel (1976), which all imply elasticities of about .2. Hendricks appeals to the estimates of Haley (1976) and Heckman (1976). In a recent study using historical data on US States, Tamura (1998) finds the elasticity of income to school expenditures to be at least .4. Given its critical role in the accumulation of human capital, $\beta\lambda$ will be allowed to vary from 0 to .4 in the sensitivity analysis.

- *Inequality.* Given α and $\beta\lambda$, the variability of idiosyncratic shocks determines the feasible range for steady-state inequality: $\lambda\Delta(\tau) \in [\lambda s/\sqrt{1 - \alpha^2}, \lambda s/\sqrt{1 - (\alpha + \beta\lambda)^2}]$. If one approximates the US distribution of family incomes as a log-normal, the mean-to-median ratio implies

²⁰So will the fact that the simulations focus (for simplicity) on steady-states. Proposition 9 showed that, for any initial conditions (m_0, Δ_0^2) , $\tau_{E,0}^* \equiv \arg \max_{\tau} [E_0(\tau; m_0, \Delta_0^2)] > 0$; this takes into account the full transition path from (m_0, Δ_0^2) to $(m(\tau), \Delta^2(\tau))$. On the other hand, $\arg \max_{\tau} [E_0(\tau; m(\tau), \Delta^2(\tau))]$ need not be positive a priori.

²¹See footnote 6. Note that since the optimal amount of physical capital used by agent i is proportional to the reduced-form net output $y_i^i = (h_i^i)^\lambda (l_i^i)^\mu$, so is gross output from the three factors: $y_i^i/Y_i^i = 1 - \omega' = y_i/Y_i$.

²²As is well-known, the effect of (marginal) variations in school expenditures on educational and labor market outcomes is the subject of significant empirical controversy. The traditional view that there is little demonstrable effect at the elementary and secondary levels is represented by Hanushek (1986). In a recent and broader meta-analysis, Dewey, Husted and Kenny (1998) provide evidence that the lack of significant positive estimates is largely the result of systematic econometric misspecification in many earlier studies.

a standard deviation of log-incomes of about .61 in the 1990 Census and .69 in that of 1995. Computing the variance directly from data on the decile income distribution leads to higher values, between .75 (1990 Census) and 1.1 (fiscal data of Bishop, Formby and Smith (1991)). I set $s = 1.0$, so that the feasible range for steady-state inequality is [.67, .78].

- *Labor Supply.* A small number of microeconomic studies, surveyed in Browning, Hansen and Heckman (1998), provide direct estimates of the intertemporal elasticity of substitution, $\epsilon \equiv 1/(\eta - 1)$. Values for males vary between 0 and .40, with a median of about .20; the sole study on female participation yields a value of 1.6. All such estimates, however, are predicated on the assumption of a frictionless credit market –which is precisely the one we depart from. Intratemporal elasticities are not subject to this problem, and also more consistent with the present focus on lifetime outcomes and steady-states. Cross-sectional estimates of the compensated labor supply elasticity for men are surveyed by Ashenfelter (1984) and Pencavel (1986); they typically vary around .10. For women, Killingsworth and Heckman’s (1986) survey includes both high and insignificant values, with a median of about .43. In view of both inter- and intra-temporal estimates I choose a reference value of $\epsilon = .20$, but in the sensitivity analysis I will explore the full range from $\epsilon = 0$ to $\epsilon = \infty$.²³ As to the coefficient on labor disutility δl^η , it can be normalized to $\delta = 1$ without loss of generality.

- *Discounting.* The standard discount factor used in macroeconomic models is approximately .96 per year, which compounds to $.96^{25} = .36$ per generation (25 years). However, this computation is quite sensitive to the choice of the annual value: $.97^{25} = .47$, while $.98^{25} = .60$. I set $\rho = .4$ in the reference case, then let it vary between .2 and .8.

- *Risk-Aversion.* For the coefficient of relative risk aversion $1 - r$, I focus on two values of interest. The first is 1, which corresponds to the intertemporally separable log-utility. The second is 2, which is a more standard estimate of risk-aversion, yet still on the conservative side. In the sensitivity analysis $1 - r$ will vary from zero (static risk-neutrality) to 4.

I report through a series of graphs the simulations for steady-state output $\ln y(\tau)$, utility from

²³Given the utility function (2), the uncompensated elasticity for a wage w and non-wage income R is $\zeta^u = R/((\eta - 1)R + \eta wl)$, and the marginal propensity to earn out of non-wage income is $mpe = -wl/((\eta - 1)R + \eta wl)$. Conversely, we can write $\epsilon = 1/(\eta - 1) = \zeta^c/[1 + mpe]$, where $\zeta^c = \zeta^u + mpe$ is the compensated elasticity. Thus typical values such as $\zeta^c \approx -mpe \in [.10, .20]$ yield ϵ in the range [.11, .25]. For instance, $\zeta^c = .11$ in Lucas’ (1990) parametrization, and $\zeta^c = .19$ in that of Hendricks (1998).

aggregate consumption and leisure $\ln z(\tau) \equiv \ln C(\tau) - \delta(l(\tau))^\eta$, efficiency $E(\tau)$, and welfare $W(\tau)$; all are measured as deviations from their values under laissez-faire ($\tau = 0$). The four functions are always single-peaked, with maxima at $\tau_Y^* < \tau_Z^* < \tau_E^* < \tau_W^*$. I also graph net transfers $b(\tau)$, intergenerational mobility $p(\tau)$, income inequality $\lambda\Delta(\tau)$, labor supply $l(\tau)$, private savings $\nu(\tau)$, and the combination of consumption taxes and investment subsidies $(\theta(\tau), a(\tau))$ required to restore the rate of accumulation to its undistorted level \mathfrak{s} . Finally, Tables I and II show how the optimal redistribution rates $(\tau_Y^*, \tau_R^*, \tau_E^*, \tau_W^*)$ vary with ϵ and $\beta\lambda$. This extensive sensitivity analysis also allows readers to replace our benchmark values for these two key parameters with their preferred ones.

6.2 Fiscal Redistribution: Benchmark Case

- *Income.* Figures 1a and 2a show that maximum per capita income occurs at $\tau_Y^* = 20.9\%$, which corresponds to a share of net transfers in GDP of $b(\tau_Y^*) = 6.2\%$. The output gain with respect to laissez-faire is 1.3%, representing the balance of a 2.7% shortfall due to the fact that agents reduce hours by 4.4% (Figure 2c), and a 5.0% increase from relaxing the liquidity constraints of poorer families. In a representative agent setting ($s^2 = 0$) only the first, negative, effect would be present. Figures 2d and 2e show that the savings distortion is fully offset by a consumption tax of 2.7%, used to finance a 26.4% subsidy for human capital investment.²⁴ Finally, note that in the endogenous growth version of the model the 1.3% net gain in long run income becomes a 0.5 percentage point rise in the long run growth rate.

- *Efficiency.* With $1 - r = 1$ total efficiency is maximized at $\tau_E^* = 48.5\%$, which corresponds to a transfer share $b(\tau_E^*) = 13.7\%$ (Figures 1a and 2a). The gain relative to $\tau = 0$ is the same as would result from an increase in every agent's consumption of 9.6%. The actual effect of τ_E^* on aggregate resources, however, is a decline of 0.9% (-7.4% from decreased labor supply versus $+6.5\%$ from the relaxation of credit constraints). But this is more than offset by the 12.1% reduction in effort, worth a 3.6% increase in consumption, and especially by the value of insurance, equivalent to another 6.8% of aggregate consumption. Neutralizing the savings distortion now requires a 6.4% consumption tax and a 94.3% subsidy for human capital investment.

²⁴The economy's investment rate then remains at $\mathfrak{s} = 11.6\%$ of aggregate income net of physical capital's remuneration, which corresponds to $\mathfrak{s}(1 - \omega') = 8.1\%$ of total factor income. Absent the corrective policy mix, the decline in savings would reduce aggregate output by $(\lambda\beta)(1 - \alpha - \beta\lambda)^{-1} \ln(\nu(\tau_Y^*)/\mathfrak{s}) = 14.6\%$.

Figures 1a–1b also illustrate one of the most important results, namely the relationship between *aggregate income*, *economic efficiency*, and a utilitarian *social welfare* function. In addition to the utility from leisure, the vertical distance between the $\ln y(\tau)$ and $E(\tau)$ curves measures the *insurance value* of taxation, which increases with risk aversion $1 - r$. The additional distance between $E(\tau)$ and $W(\tau)$ reflects the pure *equity value* of redistribution, for a degree of inequality aversion $1/\sigma = 1$. More generally, each value of $1/\sigma > 0$ defines a social welfare function $S(\sigma)$ above $E = S(\infty)$.

- *Inequality and Mobility.* Figure 2b plots intergenerational persistence and cross-sectional income inequality. Going from laissez-faire to $\tau_Y^* = 20.8\%$ reduces $p(\tau)$ from .60 to .55, and a further increase to $\tau_E^* = 48.5\%$ brings it down to .48. Equivalently, the annual convergence rate rises from 2.0% to 2.4% and then 2.9%. The effect on long-run inequality is similar: as τ rises from zero to τ_Y^* and then to τ_E^* , $\lambda\Delta(\tau)$ first falls from .78 to .75, then to .71.

- *Redistribution.* To assess the extent of redistribution implied by different values of τ , recall that it is also the (income-weighted) average marginal tax-and-transfer rate. For taxes alone, the weighted marginal rate in the US varies between 23% in the late 1970's and 17% in the late 1980's.²⁵ There is no readily available data on the incidence of transfers, but these are typically distributed much more progressively than taxes. A more relevant range for the true value of τ could thus be between .3 and .4, implying $\tau_Y^* < \tau < \tau_E^*$.²⁶ Another way to measure redistribution is to look at net transfers. In the United States the share of transfers in GDP is about 16% to 18%, but only about half are genuinely redistributive; the other half consists of social security and medicare payments. In the model, with $\tau_Y^* = 20.8\%$ the cutoff \tilde{y} between losers and gainers occurs at the 75th percentile, where family income is 1.25 times the average, or 1.65 times the median. The top 25% households earn 53.1% of total pre-tax income, but after redistribution their share falls to 46.1%.²⁷ Net transfers thus represent $b(\tau_Y^*) = 6.2\%$ of national income; see

²⁵These are typical estimates reported by Easterly and Rebello (1993) and Gouveia and Strauss (1994). Using the income tax data by decile reported by Bishop, Formby and Smith (1991), I computed that same number and again found values of 21% to 24% for 1979 and 17% to 19% for 1988.

²⁶For instance, Lambert (1993, chapter 10) computes an index of progressivity of the U.S. net fiscal system which is a properly weighted average of indices for taxes and benefits. He finds that the inclusion of the latter more than doubles overall progressivity. In the case of Chile, Engel, Galetovic and Raddatz (1998) find that the targeting of transfers and other public expenditures is by far the main source of income equalization in the fiscal system.

²⁷As a comparison, the richest 30% households in the United States earned 60% of total income in 1986 (Bishop, Formby and Smith (1991)).

Figure 2a. Under the efficient policy $\tau_E^* = 48.5\%$ the cutoff occurs at the 70th percentile (1.11 times the average, or 1.43 times the median). The pre- and post-redistribution shares of the top 30% families are 56.9% and 43.1%, hence a share of net transfers $b(\tau_E^*) = 13.7\%$. Finally, if policy is set by the median voter –or, equivalently, a social planner with inequality aversion $1/\sigma = 1$ – his purely distributional concerns will lead to $\tau_W^* = 61.2\%$, and transfers will rise to $b(\tau_W^*) = 17.0\%$ of national income. In summary, looking at either the average marginal tax rate or the transfer share suggests (quite tentatively, of course) that fiscal redistribution in the United States exceeds the income or growth-maximizing level, could be somewhat below the efficient level, and is markedly lower than the median family’s preferred outcome.²⁸

6.3 Fiscal Redistribution: Sensitivity Analysis

Figures 3a–3b show how the tradeoff between the efficiency costs and benefits of redistribution worsens as the *elasticity of labor supply* $\epsilon = 1/(\eta - 1)$ rises. Table I provides more detail on the optimal values $\tau_Y^*, \tau_Z^*, \tau_E^*$ and τ_W^* , which all decline with ϵ . Looking for instance at the row which contains the benchmark case, we see that in the absence of distortions τ_Y^* would be 66.8%, whereas beyond $\epsilon = 0.5$ redistribution only reduces total income. By contrast, the efficient tax rate and transfer share always remain positive. When $\epsilon = 1$, for instance, $\tau_E^* = 28.7\%$ and $b(\tau_E^*) = 8.4\%$, yielding a gain of 4.8% in aggregate certainty-equivalent consumption. Even as ϵ tends to infinity these numbers remain bounded below by 14.5% and 4.4% respectively (with a gain of 1.6%).

The effect of the *education return* parameter $\beta\lambda$ is more complex. Figures 4a–4b and Table I show a significant positive impact on τ_Y^* , but only a small negative one on τ_E^* ; $\beta\lambda$ mainly affects the size of the gains in $E(\tau)$. The intuition is as follows. When $\beta\lambda = 0$ there is no investment, so redistribution can only reduce output and deteriorate the consumption–leisure tradeoff; hence $\tau_Y^* = \tau_Z^* = 0$. Nonetheless, the insurance motive implies $\tau_E^* > 0$. Table I shows that τ_E^* is

²⁸There are several possible (complementary) explanations for this last result. First, it is well-documented that the propensity to engage in most forms of political participation (voting, contributing time or money to campaigns or parties, etc.), rises markedly with income and education; some of this evidence is presented in Bénabou (1996b). Second, the observed levels of progressivity and transfers may be consistent with a median voter outcome τ_W^* , if voters perceive sufficiently large tax distortions. As shown by Table I below, these would have to correspond to an intertemporal labor supply elasticity well above 2, and/or to an inability to offset savings distortions via consumption taxes and investment subsidies, as done in the model. Last, but not least, I am considering (here and in the next subsection) fiscal and educational redistribution separately. In reality both are present simultaneously, and their effects cumulative.

even fairly high since the labor disincentive is relatively small when human capital is not very persistent (see (14)). As $\beta\lambda$ rises the investment reallocation effect becomes increasingly important and gradually replaces insurance as the predominant benefit in the efficiency tradeoff with the (now stronger) labor supply effect. Consequently, τ_Y^* and τ_E^* move towards each other.^{29,30}

The role of the *family or neighborhood effect* is illustrated on Figures 5a–5b. We know from Propositions 6 and 7 that a higher α simultaneously exacerbates the effort distortion and reduces the gain from reallocating education resources toward poor families. On the other hand, it also makes human capital accumulation less sensitive to variations in investment, hence in effort, since more of it takes place “automatically” within families or neighborhoods (see the term $1 - \alpha$ multiplying $\mu \ln l(\tau)$ in (31)). This effect turns out to be dominant at relatively low values of τ , but dominated at higher ones. This is what explains the initially puzzling results that τ_Y^* increases with α while τ_E^* simultaneously decreases, in spite of the fact that more persistent shocks raise the value of insurance (as shown by the risk premium in (9), for $r = 0$).

Last but not least, Figure 6a shows the role of *risk-aversion*. The efficient rate of redistribution τ_E^* starts at 36.3% for $1 - r = 0$, then rises to 48.5%, 55.6% and 63.9% for $1 - r = 1, 2$ and 4 respectively. The corresponding shares of net transfers in national income are $b(\tau_E^*) = 10.5\%$, 13.7%, 15.5%, and 17.6%. These large variations make clear the value of working with the general preferences (2) rather than restricting attention to the separable case, $r = 0$. Finally, note that in all cases τ_E^* remains well below τ_W^* and well above τ_Y^* (or even τ_Z^*), as shown on Figure 1b.

Beyond the results obtained for specific parameters values, it is the *general message* from these simulations which is most important. They consistently show that the efficiency costs and benefits of redistribution are both quantitatively important, and that per capita income and aggregate welfare provide only very imperfect measures of the resulting tradeoff.

²⁹The only column of Table I along which τ_Y^* decreases with $\beta\lambda$ is the first one, where $\epsilon = 0$. Indeed, when there is no distortion to trade off against the credit-constraint effect $\mathfrak{L}(\tau)\Delta^2(\tau)$, the latter’s increasing magnitude with $\beta\lambda$ is irrelevant. All that matters is the fact that $\arg \min_{\tau} \{\mathfrak{L}(\tau)\Delta^2(\tau)\}$ declines with $\beta\lambda$, due the predominant influence of $\arg \min_{\tau} \{\mathfrak{L}(\tau)\} = 1 - \alpha/(1 - \beta\lambda)$.

³⁰The results remain essentially unchanged when α and $\beta\lambda$ are varied with $\alpha + \beta\lambda$ fixed at its reference value of .60, so as to leave total returns unchanged.

6.4 Education Finance: Benchmark Case

• *Income.* Because current consumption is not redistributed, policy distortions from equalizing educational investments are considerably smaller than from equalizing family incomes. Figure 7a shows that the output-maximizing rate is now $\tau_Y^* = 61.9\%$, leading to a sizeable 6.1% gain with respect to laissez-faire. The fall in labor supply is only 1.3%, as can be read off Figure 7c. On the savings side, a consumption tax of 0.9% and a 7.2.% average subsidy for education are now sufficient to maintain accumulation at its first-best level, still equal to $\mathfrak{s} = 11.6\%$.³¹ In the endogenous growth version of the model, the 6.1% increase in per capita income becomes a 2.4 percentage point rise in the long-run growth rate.

• *Efficiency.* With $1 - r = 1$, the efficient rate of education finance equalization is $\tau_E^* = 68.2\%$. In other words, about two-thirds of the variations in per-pupil expenditures reflecting differences in family incomes should be offset. The corresponding efficiency gain is 7.3%, of which 6.0% is due to increased aggregate income and consumption, 0.6% to lower effort, and only 0.7% to better risk-sharing. Figures 7a–7b clearly show the reduced value of redistributive education finance as a social insurance scheme, compared to that of taxes and transfers: $E(\tau)$ is closer to $\ln z(\tau)$ than on Figures 1a–1b, and much less sensitive to agents' degree of risk aversion.

• *Inequality and Mobility.* Since the optimal rate of progressivity is higher than for fiscal policy, intergenerational persistence and cross-sectional inequality are reduced further. Figure 8b shows that as τ rises from zero to τ_Y^* and then to τ_E^* , $p(\tau)$ falls from .68 to .45 and then to .43. Cooper (1998) finds evidence in the PSID data that redistribution of educational expenditures (measured by the percentage of school funding coming from outside each district) lowers intergenerational persistence between fathers' and sons' incomes. Consider next inequality in family incomes and per-student expenditures, $\lambda\Delta(\tau)$ and $(1 - \tau)\lambda\Delta(\tau)$. These are reduced from (.78, .78) under laissez-faire to (.70, .27) under τ_Y^* , (.69, .22) under τ_E^* , and (.68, .08) under $\tau_W^* = 87.4\%$. The expenditure numbers may be compared to Hoxby's (1998b) estimates of the coefficient of variation in per-pupil spending among local school districts in California, Illinois and Massachusetts: respectively .16, .25 and .28 for recent years. When normalized by state-wide income inequality,

³¹Absent this corrective scheme, the decline in savings would reduce from 6.6% to 2.3% the gain in long-run income resulting from $\tau_Y^* = 61.9\%$.

the corresponding ratios are all around .15. The values generated by the model thus tends to suggest that the extent of school finance redistribution in those three states exceeds the level that would maximize long-run income (or growth), but is below what the median agent would choose, or what would maximize average welfare. It may be somewhat lower or higher than the efficiency-maximizing level, depending on the value of risk aversion.³²

- *Redistribution.* Under the policy τ_Y^* , the break-even point \tilde{y} which separates families whose education is subsidized beyond the average rate $a(\tau)$ from those whose expenditures are taxed relative to $a(\tau)$ occurs at the 69th percentile (1.10 times per capita income, or 1.40 times the median). The 31% richest families earn 58.5% of total income, but their share of total educational expenditures is reduced to 41.4%. The fraction of total school spending reallocated from to relatively rich to relatively poor families or communities is thus $b(\tau_Y^*) = 17.1\%$. As illustrated on Figure 8a, it rises to 18.7% under the efficient policy τ_E^* (the cutoff remains almost unchanged, at the 68th percentile), and to 23.2% under the policy τ_W^* which maximizes average welfare and the utility of the median family.

6.5 Education Finance: Sensitivity Analysis

The most striking feature of the results reported in Figures 9a–9b and Table II is that *high rates of education finance equalization* remain optimal, no matter *how large* the intertemporal elasticity of labor supply ϵ . Even when preferences become linear in effort, $\epsilon = +\infty$, the income-maximizing and efficient rates of equalization are still $\tau_Y^* = 39.2\%$ and $\tau_E^* = 59.6\%$. The corresponding shares of educational resources being redistributed are $b(\tau_Y^*) = 11.3\%$ and $b(\tau_E^*) = 16.5\%$, and the resulting gains still amount to 2.8% for output and 5.9% for efficiency.

The effect of the *education return* parameter is, once again, more complicated. A higher $\beta\lambda$ lowers both τ_Y^* and τ_E^* while at the same time significantly increasing the gains resulting from these policies, compared to $\tau = 0$. For instance, as $\beta\lambda$ rises from .05 (τ is irrelevant when $\beta\lambda = 0$) to .30, τ_Y^* declines from 68.5% to 59.6% and τ_E^* from 76.1% to 65.6%. At the same time, the corresponding gains rise from 0.9% to 8.2% for output, and from 1.1% to 9.7% for efficiency. The

³²With $1-r = 2$ the figures are $\tau_E^* = 74.9\%$, leading to $(\lambda\Delta(\tau_E^*), (1-\tau_E^*)\lambda\Delta(\tau_E^*)) = (.69, .17)$; and $\tau_W^* = 93.1\%$, leading to $(\lambda\Delta(\tau_W^*), (1-\tau_W^*)\lambda\Delta(\tau_W^*)) = (.67, .05)$. As to τ_Y^* , it is of course unchanged. Note that the data used by Hoxby pertains only to elementary and secondary school budgets. This leaves out spending disparities at the early childhood stage, which is often considered to play a critical role in shaping later educational outcomes.

intuition is as follows. With labor supply distortions and insurance both playing relatively modest roles, the key determinant of educational policy’s impact on growth and efficiency is the relaxation of credit constraints. This effect, measured by $\mathfrak{L}(\tau)\Delta^2(\tau)$ in (31), becomes more important as $\beta\lambda$ rises, i.e. as resource expenditure matters more for accumulation. Yet the optimal τ declines, because the accumulation technology gets closer to constant (private) returns; recall that $\mathfrak{L}(\tau)$ is minimized at $\tau = (1 - \alpha - \beta\lambda)/(1 - \beta\lambda)$.

The role of *family or peer effects* is depicted on Figures 11a–11b. The main impact of α is now to reduce the optimal degree of equalization of school inputs across students from different social backgrounds, as explained earlier. Thus, both τ_Y^* and τ_E^* decline with α . The importance of this result is best illustrated by starting from a case similar to that of Fernandez and Rogerson (1998): $\alpha = 0$ (no family or peer effect), $\epsilon = 0$ (inelastic labor), and $\beta\lambda = .19$. The optimal policy is then complete equalization ($\tau_Y^* = \tau_E^* = 1$), and brings a 3.8% gain in long run output over decentralized funding – a figure close to Fernandez and Rogerson’s 3.2%. Incorporating the benchmark labor response ($\epsilon = .20$) reduces τ_Y^* to 92.1%, and the maximum gain to 3.2%. But much more significant is the role of social background: with $\alpha = .35$ and $\epsilon = 0$, τ_Y^* falls very sharply, to 69.0%. On the other hand, long–run income gains are magnified by the higher total return, $\alpha + \beta\lambda$: 4.6% under τ_Y^* , but only 3.9% under $\tau = 1$. When social capital and labor supply effects are combined the results are very close to those in Table II, third column, third row: $\tau_Y^* = 63.8\%$ raises output by 4.4%, while complete equalization yields only a 3.3% gain.

As explained earlier, *risk–aversion* plays much less of a role than under fiscal redistribution. Figure 12a shows that the efficient progressivity rate τ_E^* starts at 61.6% for $1 - r = 0$ but rises relatively slowly, to 68.2%, 74.9% and 88.3% for $1 - r = 1, 2$ and 4 respectively. The corresponding fractions of school expenditures being redistributed are $b(\tau_E^*) = 17.0\%$, 18.7%, 20.3%, and 23.4%. Except for very low values of risk–aversion, aggregate income still underestimates the value of redistribution, although less so than with taxes and transfers; see Figures 7a–7b. The bias in aggregate welfare, on the contrary, is now much more severe. Of course, pure equity considerations (a positive value of $1/\sigma$) may well be more relevant in educational than in tax policy.

7 Conclusion

This paper has studied how progressive income taxation and education finance affect the level and distribution of income in a dynamic heterogeneous agent economy. The model was first solved analytically, then quantitative policy exercises were performed. A simple combination of consumption taxes and education subsidies can help restore investment to its undistorted level. Whether or not this additional policy instrument is used, education finance always dominates taxes and transfers from the point of view of growth, but is inferior from that of insurance. Simulating the model with empirical parameter estimates leads to generally reasonable results. For the benchmark specification efficiency is maximized with transfers equal to 14% of GDP, or with a 68% equalization rate for school expenditures. In both cases the richest 30% of families end up subsidizing the education (and possibly the consumption) of the remaining 70%, whether through school finance or through the fiscal system. More generally, the efficiency costs and benefits of redistribution remain of the same order of magnitude over a wide range of parameters values, so that omitting either side can seriously bias the policy analysis. Another robust conclusion is that per capita income and average welfare provide only crude lower and upper bounds around a more proper (risk-adjusted but distribution-free) measure of overall efficiency.

The model's analytical structure has a number of advantages, which hopefully justify the strong simplifying assumptions which lie behind it. One is the transparency of the insights obtained from complete closed-form solutions. Another is allowing anyone to easily generate alternative policy assessments, by replacing in the formulas our parameter choices with their preferred values. Finally, the model can be extended in several interesting directions. One could thus analyze fiscal and educational policy jointly rather than separately, and look for the optimal mix which alleviates the imperfections in the credit and insurance markets with minimal distortions. Another route is to endogenize the degree of redistribution through a political mechanism. This is pursued in Bénabou (1996b), which seeks to explain how countries with similar economic and political fundamentals can nonetheless choose very different fiscal and education systems.

Appendix

A Proofs for the Case of Income Taxation

Proofs of Propositions 1, 2 and 3

The first-order condition for the optimal savings rate in (9) is:

$$\begin{aligned} \frac{1-\rho}{1-\nu_t} &= \rho \frac{\mathbf{E}_t [(U_{t+1})^r (\partial \ln U_{t+1} / \partial \ln h') (\partial \ln h' / \partial \nu)]}{\mathbf{E}_t [(U_{t+1})^r]} \Leftrightarrow \\ \frac{\nu_t}{1-\nu_t} &= \frac{\rho\beta}{1-\rho} \times \frac{\mathbf{E}_t [(U_{t+1})^r (\partial \ln U_{t+1} / \partial \ln h')]}{\mathbf{E}_t [(U_{t+1})^r]} \end{aligned} \quad (\text{A1})$$

while for labor supply it is:

$$\begin{aligned} (1-\rho)[-(1-\tau_t)(\mu/l) + \delta\eta l^{\eta-1}] &= \rho \frac{\mathbf{E}_t [(U_{t+1})^r (\partial \ln U_{t+1} / \partial \ln h') (\partial \ln h' / \partial l)]}{\mathbf{E}_t [(U_{t+1})^r]} \Leftrightarrow \\ \delta\eta l_t^\eta &= \mu(1-\tau_t) \left[1 + \frac{\rho\beta}{1-\rho} \times \left(\frac{\mathbf{E}_t [(U_{t+1})^r (\partial \ln U_{t+1} / \partial \ln h')]}{\mathbf{E}_t [(U_{t+1})^r]} \right) \right]. \end{aligned} \quad (\text{A2})$$

We guess that the value function is of the form: $\ln U_t^i = V_t \ln h_t^i + B_t$. Substituting into the Bellman equation yields:

$$\begin{aligned} V_t \ln h_t + B_t &= \max_l \{ (1-\rho + \rho\beta V_{t+1})(1-\tau)\mu \ln l - (1-\rho)\delta(l)^\eta \} \\ &\quad + \max_\nu \{ (1-\rho) \ln((1-\nu)/(1+\theta_t)) + \rho\beta V_{t+1} \ln(\nu_t(1+a_t)) \} \\ &\quad + [(1-\rho)\lambda(1-\tau_t) + \rho V_{t+1}(\alpha + \beta\lambda(1-\tau_t))] \ln h_t^i + (1-\rho + \rho\beta V_{t+1})\tau_t \ln \tilde{y}_t \\ &\quad - (\rho/r) [rV_{t+1}(1-rV_{t+1})s^2/2] + \rho(B_{t+1} + V_{t+1} \ln \kappa). \end{aligned} \quad (\text{A3})$$

This problem is strictly concave in ν , as well as in l given $\eta \geq 1$ (strictly quasiconcavity in l is even ensured for all $\eta > 0$). Therefore, if (A3) does hold, the first-order conditions are sufficient to characterize the optimum. Moreover, in this case (A1)–(A2) immediately simplify to yield Propositions (2) and (3). Now, (A3) requires that

$$V_t = (1-\rho)(1-\tau_t)\lambda + \rho(\alpha + \beta\lambda(1-\tau_t))V_{t+1}, \quad (\text{A4})$$

which yields (12). As to B_t , it is given as the solution to the difference equation:

$$B_t - \rho B_{t+1} = (1 - \rho + \rho\beta V_{t+1})(1 - \tau_t)\mu \ln l_t - (1 - \rho)\delta(l_t)^\eta + \rho V_{t+1} \ln \kappa \quad (\text{A5})$$

$$+ (1 - \rho) \ln((1 - \nu_t)/(1 + \theta_t)) + \rho\beta V_{t+1} \ln(\nu_t(1 + a_t)) \\ + (1 - \rho + \rho\beta V_{t+1}) \ln \tilde{y}_t - (\rho/r)[rV_{t+1}(1 - rV_{t+1})]s^2/2. \quad (\text{A6})$$

with the transversality condition $\lim_{t \rightarrow \infty} (e^{-\rho t} B_t) = 0$, which implies that B_t equals the present value of the right-hand side terms. To simplify further, note that since all agents have the same $l_t^i = l_t$ and $\nu_t^i = \nu_t$ one can write:

$$\ln h_{t+1}^i = \ln \xi_{t+1}^i + \ln \kappa + \beta \ln \mathfrak{s}_t + \beta\mu(1 - \tau_t) \ln l_t + (\alpha + \beta\lambda(1 - \tau_t)) \ln h_t^i + \beta\tau_t \ln \tilde{y}_t, \quad (\text{A7})$$

where $\mathfrak{s}_t \equiv \nu_t(1 + a_t)$. Thus h_t^i remains log-normally distributed over time. If $\ln h_t^i \sim \mathcal{N}(m_t, \Delta_t^2)$, then (7) yields for \tilde{y}_t the value given by equation (27). Substituting into (A7) yields (25)–(26).

Let us now define $W_t \equiv V_t m_t + B_t$, so that $V_t(h) = V_t(\ln h - m) + W_t$. Substituting $B_t = W_t - V_t m_t$ into (A5), then using (25)–(26) to eliminate m_{t+1} and Δ_{t+1}^2 , the budget constraint (17) to eliminate θ_t , and (27) to eliminate $\ln \tilde{y}_t$, we obtain:

$$W_t - \rho W_{t+1} = (1 - \rho + \rho\beta V_{t+1})(1 - \tau_t)\mu \ln l_t - (1 - \rho)\delta(l_t)^\eta + \rho V_{t+1} \ln \kappa \\ + (1 - \rho) \ln(1 - \mathfrak{s}_t) + \rho\beta V_{t+1} \ln \mathfrak{s}_t - (\rho/r)[rV_{t+1}(1 - rV_{t+1})]s^2/2 \\ + (1 - \rho + \rho\beta V_{t+1})\tau_t [\lambda m_t + \mu \ln l_t + (2 - \tau_t)\lambda^2 \Delta_t^2/2] + V_t m_t \\ - \rho V_{t+1} [(\alpha + \beta\lambda)m_t + \beta\mu \ln l_t + \beta\tau_t(2 - \tau_t)\lambda^2 \Delta_t^2/2 + \beta \ln \mathfrak{s}_t + \ln \kappa - s^2/2].$$

Therefore, denoting $\mathcal{W}_t \equiv W_t/(1 - \rho)$:

$$\mathcal{W}_t - \rho \mathcal{W}_{t+1} = \mu \ln l_t - \delta(l_t)^\eta + \ln(1 - \mathfrak{s}_t) + \lambda m_t + \tau_t(2 - \tau_t)\lambda^2 \Delta_t^2/2 + r\rho(1 - \rho)^{-1} V_{t+1}^2 s^2/2, \quad (\text{A8})$$

where we used the recursion equation (A4) to simplify the coefficient on m_t . Thus:

$$\mathcal{W}_t \equiv W_t/(1 - \rho) = \sum_{k=0}^{\infty} \rho^k (\lambda m_{t+k} + \mu \ln l_{t+k} + \tau_{t+k}(2 - \tau_{t+k})\lambda^2 \Delta_{t+k}^2/2) \\ + \sum_{k=0}^{\infty} \rho^k (\ln(1 - \mathfrak{s}_{t+k}) - \delta(l_{t+k})^\eta + r\rho(1 - \rho)^{-1} V_{t+k+1}^2 s^2/2). \quad (\text{A9})$$

It just remains to compute $\sum_{t=0}^{\infty} \rho^t (\lambda m_t + \mu \ln l_t + \tau_t(2 - \tau_t)\lambda^2 \Delta_t^2/2)$ as a function of the initial distribution (m_0, Δ_0^2) and the policy sequence $\{\tau_t, a_t, \theta_t\}_{t=0}^{\infty}$. Now, (25) implies:

$$\begin{aligned}
& \sum_{t=0}^{\infty} \rho^t (\lambda m_t + \mu \ln l_t + \tau_t (2 - \tau_t) \lambda^2 \Delta_t^2 / 2) \\
&= (1 - \rho(\alpha + \beta\lambda))^{-1} \left(\lambda m_0 + \rho\lambda \sum_{t=0}^{\infty} \rho^t (\beta \ln \kappa - s^2/2 + \beta \ln \mathfrak{s}_t) \right) \\
& \quad + (1 + \rho\beta\lambda(1 - \rho(\alpha + \beta\lambda))^{-1}) \left(\sum_{t=0}^{\infty} \rho^t [\mu \ln l_t + \tau_t (2 - \tau_t) \lambda^2 \Delta_t^2 / 2] \right).
\end{aligned}$$

So, finally:

$$\begin{aligned}
\mathcal{W}_t &= \frac{\lambda m_t}{1 - \rho(\alpha + \beta\lambda)} + \sum_{k=0}^{\infty} \rho^k \left[\left(\frac{1 - \rho\alpha}{1 - \rho(\alpha + \beta\lambda)} \right) \mu \ln l_{t+k} - \delta l_{t+k}^\eta \right] \\
& \quad + \sum_{k=0}^{\infty} \rho^k \left[r(1 - \rho)^{-1} V_{t+k+1}^2 - \frac{\lambda}{1 - \rho(\alpha + \beta\lambda)} \right] \left(\frac{\rho s^2}{2} \right) + \sum_{k=0}^{\infty} \rho^k \left(\ln(1 - \mathfrak{s}_{t+k}) \right. \\
& \quad \left. + \frac{\rho\lambda(\ln \kappa + \beta \ln \mathfrak{s}_{t+k})}{1 - \rho(\alpha + \beta\lambda)} \right) + \left(\frac{1 - \rho\alpha}{1 - \rho(\alpha + \beta\lambda)} \right) \sum_{k=0}^{\infty} \rho^k \left[\tau_{t+k}(2 - \tau_{t+k}) \frac{\lambda^2 \Delta_{t+k}^2}{2} \right]. \quad (\text{A10})
\end{aligned}$$

In particular, under a constant policy $\{\tau_t = \tau\}_{t=0}^{\infty}$ aggregate welfare simplifies to:

$$W_0 = \bar{u}_0(\tau) + (1 - \rho)\Omega_\Delta(\tau) \left(\frac{\lambda^2 \Delta_0^2}{2} \right) + \rho\Omega_s(\tau) \left(\frac{\lambda^2 s^2}{2} \right), \quad (\text{A11})$$

where:

$$\begin{aligned}
\bar{u}_0(\tau) &\equiv \frac{\lambda[(1 - \rho)m_0 + \rho(\ln \kappa + \beta \ln \mathfrak{s})]}{1 - \rho(\alpha + \beta\lambda)} + \ln(1 - \mathfrak{s}) + \left(\frac{1 - \rho\alpha}{1 - \rho(\alpha + \beta\lambda)} \right) \mu \ln l(\tau) - \delta l(\tau)^\eta, \\
\Omega_\Delta(\tau) &\equiv \left(\frac{1 - \rho\alpha}{1 - \rho(\alpha + \beta\lambda)} \right) \left(\frac{\tau(2 - \tau)}{1 - \rho p(\tau)^2} \right), \\
\Omega_s(\tau) &\equiv \Omega_\Delta(\tau) - \frac{1/\lambda}{1 - \rho(\alpha + \beta\lambda)} + r(1 - \rho) \left(\frac{1 - \tau}{1 - \rho p(\tau)} \right)^2.
\end{aligned}$$

In steady-state it simplifies further, as Δ_0^2 becomes equal to $\Delta_\infty^2(\tau) = s^2/(1 - p(\tau)^2)$.

Proof of Proposition 4

Let us now examine consumers' preferred time paths of consumption taxes and associated investment subsidies $\{\theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$, given any expected redistribution path $\{\tau_{t+k}\}_{k=0}^{\infty}$. Note that once the θ_{t+k} 's are eliminated from the budget constraint (17), choosing $\{a_{t+k}\}_{k=0}^{\infty}$ is equivalent to choosing $\{\mathfrak{s}_{t+k} \equiv \nu_{t+k}(1 + a_{t+k})\}_{k=0}^{\infty}$. Since an agent i 's utility is $\ln U_t^i = V_t (\ln h_t^i - m_t) + W_t$ and V_t is independent of $\{\theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$, all agree on the optimal path of $\{\mathfrak{s}_{t+k}\}_{k=0}^{\infty}$, namely the one that maximizes aggregate welfare W_t . By (A10), this sequence is given by:

$$\frac{1 - \rho(\alpha + \beta\lambda)}{1 - \mathfrak{s}_{t+k}} = \frac{\rho\beta\lambda}{\mathfrak{s}_{t+k}}, \quad \text{or} \quad \mathfrak{s}_{t+k} = \frac{\rho\beta\lambda}{1 - \rho\alpha} = \mathfrak{s}.$$

Proof of Proposition 8

Let x be a normal random variable with mean m , variance Δ^2 , and c.d.f. therefore equal to $F(x) = \Phi((x - m)/\Delta)$. By direct computation one easily establishes that, for any θ and $\lambda \in \mathbb{R}$,

$$\int_{-\infty}^{\theta} e^{\lambda x} dF(x) = e^{\lambda(m + \lambda\Delta^2/2)} F(\theta - \lambda\Delta^2). \quad (\text{A12})$$

Consider now the log-normal variable $y \equiv e^x$. Its c.d.f. is $G(y) = F(\ln y)$, and the corresponding Lorenz curve is $\Gamma(r) = \left(\int_0^{G^{-1}(r)} y dG(y) \right) / \left(\int_0^{\infty} y dG(y) \right) = \left(\int_0^{F^{-1}(r)} e^x dF(x) \right) / \left(\int_{-\infty}^{\infty} e^x dF(x) \right)$. Applying (A12) with $\lambda = 1$ and $\theta = F^{-1}(r) = m + \Delta\Phi^{-1}(r)$ yields the claimed result.

Proof of Proposition 9

First, observe that (36) follows directly from (A8) and (24), given the definition of \bar{c}_t^i in (35). Turning now to the proposition itself, we can rewrite (A8) at $t = 0$ as:

$$\begin{aligned} W_0/(1 - \rho) &= \sum_{t=0}^{\infty} \rho^t (\ln y_t - (1 - \tau_t)^2 \lambda^2 \Delta_t^2 / 2 - \delta(l_t)^\eta + \ln(1 - \mathfrak{s}_t) + r\rho(1 - \rho)^{-1} V_{t+1}^2 s^2 / 2) \Rightarrow \\ &E_0/(1 - \rho) \equiv \sum_{t=0}^{\infty} \rho^t (\ln y_t - \delta(l_t)^\eta + \ln(1 - \mathfrak{s}_t)) \\ &+ \sum_{t=0}^{\infty} \rho^t (1 - \tau_t)^2 \lambda^2 \left[r\rho(1 - \rho) \left(\frac{1}{1 - \rho p(\tau_t)} \right)^2 \left(\frac{s^2}{2} \right) - \frac{\Delta_t^2}{2} + \left(\prod_{k=0}^{t-1} p(\tau_k)^2 \right) \frac{\Delta_0^2}{2} \right]. \end{aligned}$$

Now, from (26) we know that

$$\Delta_t^2 = \left(\prod_{k=0}^{t-1} p(\tau_k)^2 \right) \Delta_0^2 + \left(\sum_{k=1}^t \prod_{q=k}^{t-1} p(\tau_q)^2 \right) s^2. \quad (\text{A13})$$

Therefore:

$$\frac{E_0(\tau)}{1 - \rho} = \sum_{t=0}^{\infty} \rho^t [\ln y_t - \delta(l_t)^\eta] + \ln(1 - \mathfrak{s}) + \sum_{t=0}^{\infty} \rho^t (1 - \tau_t)^2 \left(\frac{r\rho(1 - \rho)}{1 - \rho p(\tau_t)^2} - \sum_{k=1}^t \prod_{q=k}^{t-1} p(\tau_q)^2 \right) \left(\frac{\lambda^2 s^2}{2} \right). \quad (\text{A14})$$

With a constant τ the last present value, times $1 - \rho$, becomes $-\rho(1 - \tau)^2 \phi(\tau, r) \lambda^2 s^2 / 2$, where

$$\phi(\tau, r) \equiv \left(\frac{1}{1 - \rho p(\tau)^2} - \frac{r(1 - \rho)}{(1 - \rho p(\tau_t))^2} \right) \geq \phi(\tau, 1) = \frac{\rho(1 + p(\tau)^2 - 2p(\tau))}{(1 - \rho p(\tau)^2)(1 - \rho p(\tau_t))^2} \geq 0,$$

with strict inequality for all $p(\tau) < 1$. This establishes the proposition, together with the fact that the risk premium is always positive and minimized at $\tau = 1$.

Finally, we prove the second result in the proposition. From (A11) and (37), we have:

$$W_0 = \bar{u}_0(\tau) + (1 - \rho)\Omega_\Delta(\tau) \left(\frac{\lambda^2 \Delta_0^2}{2} \right) + \rho\Omega_s(\tau) \left(\frac{\lambda^2 s^2}{2} \right), \quad (\text{A15})$$

$$E_0 = W_0 + (1 - \rho) \left(\frac{(1 - \tau)^2}{1 - \rho p(\tau)^2} \right) \left(\frac{\lambda^2 \Delta_0^2}{2} \right) \quad (\text{A16})$$

First, it is easily verified from (14) and (16) that $\bar{u}'_0(0) = 0$; this holds whether the investment rate is s or $\nu(\tau)$. Next, straightforward but somewhat tedious derivations show that $\Omega'_\Delta(0) + (\partial/\partial\tau)_{\tau=0} [(1 - \tau)^2 / (1 - \rho p(\tau)^2)] \geq 0$, with strict inequality unless $\rho\beta\lambda = 1$; and that $\Omega'_s(0) \geq 0$, with strict inequality unless $\alpha + \beta\lambda = 1$. Therefore $E'_0(0) > 0$, hence $\tau_{E,0}^* > 0$.

B Proofs for the Case of Education Finance

Proofs of Propositions 5 and 6

The first-order condition for the optimal savings rate in (9) is unchanged:

$$\begin{aligned} \frac{1 - \rho}{1 - \nu_t} &= \rho \frac{\mathbb{E}_t [(U_{t+1})^r (\partial \ln U_{t+1} / \partial \ln h') (\partial \ln h' / \partial \nu)]}{\mathbb{E}_t [(U_{t+1})^r]} \Leftrightarrow \\ \frac{\nu_t}{1 - \nu_t} &= \rho\beta(1 - \tau_t) \frac{\mathbb{E}_t [(U_{t+1})^r (\partial \ln U_{t+1} / \partial \ln h')]}{\mathbb{E}_t [(U_{t+1})^r]}, \end{aligned} \quad (\text{B1})$$

while for labor supply it becomes:

$$\begin{aligned} (1 - \rho)[- \mu/l + \delta\eta l^{\eta-1}] &= \rho \frac{\mathbb{E}_t [(U_{t+1})^r (\partial \ln U_{t+1} / \partial \ln h') (\partial \ln h' / \partial l)]}{\mathbb{E}_t [(U_{t+1})^r]} \Leftrightarrow \\ \delta\eta l_t^\eta &= \mu \left[1 + \rho\beta(1 - \tau_t) \left(\frac{\mathbb{E}_t [(U_{t+1})^r (\partial \ln U_{t+1} / \partial \ln h')]}{\mathbb{E}_t [(U_{t+1})^r]} \right) \right]. \end{aligned} \quad (\text{B2})$$

We guess once again that the value function is of the form: $\ln U_t^i = V_t \ln h_t^i + B_t$. Substituting into the Bellman equation yields:

$$\begin{aligned} V_t \ln h_t + B_t &= \max_l \{ (1 - \rho + \rho\beta(1 - \tau)V_{t+1}) \mu \ln l - (1 - \rho)\delta(l)^\eta \} \\ &\quad + \max_\nu \{ (1 - \rho) \ln((1 - \nu)/(1 + \theta_t)) + \rho\beta V_{t+1} \ln(\nu_t(1 + a_t)) \} \end{aligned}$$

$$\begin{aligned}
& + [(1 - \rho)\lambda + \rho V_{t+1}(\alpha + \beta\lambda(1 - \tau_t))] \ln h_t^i + (\rho\beta V_{t+1})\tau_t \ln \tilde{y}_t \\
& - (\rho/r) [rV_{t+1}(1 - rV_{t+1})s^2/2] + \rho(B_{t+1} + V_{t+1} \ln \kappa). \tag{B3}
\end{aligned}$$

Thus:

$$V_t = (1 - \rho)\lambda + \rho(\alpha + \beta\lambda(1 - \tau_t))V_{t+1}, \tag{B4}$$

which yields (20), as claimed. Proposition 6 then follows immediately from the first-order conditions (B1)–(B2), which are again sufficient. The transition equation for $\ln h_t^i$, the formula for $\ln \tilde{y}_t$ and dynamics of the distribution (m_t, Δ_t^2) remain unchanged from the case of income taxes. Finally, B_t is given as the solution to the difference equation:

$$\begin{aligned}
B_t - \rho B_{t+1} &= (1 - \rho + \rho\beta(1 - \tau_t)V_{t+1})\mu \ln l_t - (1 - \rho)\delta(l_t)^\eta + \rho V_{t+1} \ln \kappa \\
& + (1 - \rho) \ln((1 - \nu_t)/(1 + \theta_t)) + \rho\beta V_{t+1} \ln(\nu_t(1 + a_t)) \\
& + (\rho\beta V_{t+1}) \ln \tilde{y}_t - (\rho/r)[rV_{t+1}(1 - rV_{t+1})]s^2/2 \tag{B5}
\end{aligned}$$

with the transversality condition $\lim_{t \rightarrow \infty} (e^{-\rho t} B_t) = 0$, which implies that B_t equals the present value of the right-hand side terms. Defining again $W_t \equiv V_t m_t + B_t$, so that $V_t(h) = V_t(\ln h - m) + W_t$, then substituting the budget constraint (17) as well as $\ln \tilde{y}_t$ from (27), we obtain:

$$\begin{aligned}
W_t - \rho W_{t+1} &= (1 - \rho + \rho\beta(1 - \tau_t)V_{t+1})\mu \ln l_t - (1 - \rho)\delta(l_t)^\eta + \rho V_{t+1} \ln \kappa \\
& + (1 - \rho) \ln(1 - \mathfrak{s}_t) + \rho\beta V_{t+1} \ln \mathfrak{s}_t - (\rho/r)[rV_{t+1}(1 - rV_{t+1})]s^2/2 \\
& + (\rho\beta V_{t+1})\tau_t [\lambda m_t + \mu \ln l_t + (2 - \tau_t)\lambda^2 \Delta_t^2/2] + V_t m_t \\
& - \rho V_{t+1} [(\alpha + \beta\lambda)m_t + \beta\mu \ln l_t + \beta\tau_t(2 - \tau_t)\lambda^2 \Delta_t^2/2 + \beta \ln \mathfrak{s}_t + \ln \kappa - s^2/2] \Leftrightarrow \\
W_t - \rho W_{t+1} &= (1 - \rho + \rho\beta(1 - \tau_t)V_{t+1})\mu \ln l_t - (1 - \tau_t)\rho\beta V_{t+1}(\mu \ln l_t) \\
& + \tau_t(2 - \tau_t)\lambda^2 \Delta_t^2/2[\rho\beta V_{t+1} - \rho\beta V_{t+1}] + m_t[V_t - \rho V_{t+1}(\alpha + \beta\lambda(1 - \tau_t))] \\
& + (1 - \rho) \ln(1 - \mathfrak{s}_t) - (1 - \rho)\delta(l_t)^\eta + (\rho r V_{t+1}^2)s^2/2,
\end{aligned}$$

where we used (B4) to simplify the coefficient on m_t . Thus, denoting again $\mathcal{W}_t \equiv W_t/(1 - \rho)$:

$$\mathcal{W}_t - \rho \mathcal{W}_{t+1} = \mu \ln l_t - \delta(l_t)^\eta + \ln(1 - \mathfrak{s}_t) + \lambda m_t + r\rho(1 - \rho)^{-1}V_{t+1}^2 s^2/2, \tag{B6}$$

and finally:

$$\mathcal{W}_t = \sum_{k=0}^{\infty} \rho^k (\lambda m_{t+k} + \mu \ln l_{t+k}) + \sum_{k=0}^{\infty} \rho^k (\ln(1 - \mathfrak{s}_{t+k}) - \delta (l_{t+k})^\eta + r\rho(1 - \rho)^{-1} V_{t+k+1}^2 s^2 / 2).$$

It just remains to compute $\sum_{k=0}^{\infty} \rho^k (\lambda m_{t+k} + \mu \ln l_{t+k})$ as a function of the initial conditions (m_0, Δ_0^2) and the policy sequence $\{\tau_t, a_t, \theta_t\}_{t=0}^{\infty}$. Now, (25) implies:

$$\begin{aligned} \sum_{t=0}^{\infty} \rho^t (\lambda m_t + \mu \ln l_t) &= (1 + \rho\beta\lambda(1 - \rho(\alpha + \beta\lambda))^{-1}) \left(\sum_{t=0}^{\infty} \rho^t \mu \ln l_t \right) + \rho(\alpha + \beta\lambda)^{-1} \\ &\quad \times \left(\lambda m_0 + \rho\lambda \sum_{t=0}^{\infty} \rho^t (\ln \kappa - s^2 / 2 + \beta \ln \mathfrak{s}_t + \beta \tau_t (2 - \tau_t) \lambda^2 \Delta_t^2 / 2) \right). \end{aligned}$$

So, ultimately:

$$\begin{aligned} \mathcal{W}_t &= \frac{\lambda m_t}{1 - \rho(\alpha + \beta\lambda)} + \sum_{k=0}^{\infty} \rho^k \left[\left(\frac{1 - \rho\alpha}{1 - \rho(\alpha + \beta\lambda)} \right) \mu \ln l_{t+k} - \delta l_{t+k}^\eta \right] \\ &\quad + \sum_{k=0}^{\infty} \rho^k \left[r(1 - \rho)^{-1} V_{t+k+1}^2 - \frac{\lambda}{1 - \rho(\alpha + \beta\lambda)} \right] \left(\frac{\rho s^2}{2} \right) + \sum_{k=0}^{\infty} \rho^k \left(\ln(1 - \mathfrak{s}_{t+k}) \right. \\ &\quad \left. + \frac{\rho\lambda(\ln \kappa + \beta \ln \mathfrak{s}_{t+k})}{1 - \rho(\alpha + \beta\lambda)} \right) + \left(\frac{\rho\beta\lambda}{1 - \rho(\alpha + \beta\lambda)} \right) \sum_{k=0}^{\infty} \rho^k \left[\tau_{t+k} (2 - \tau_{t+k}) \frac{\lambda^2 \Delta_{t+k}^2}{2} \right]. \quad (\text{B7}) \end{aligned}$$

Proof of the Education Finance Analogue of Proposition 4

Once again, choosing $\{\theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$ subject to the budget constraint (17) is equivalent to choosing $\{\mathfrak{s}_{t+k} \equiv \nu_{t+k}(1 + a_{t+k})\}_{k=0}^{\infty}$. Since an agent i 's utility is $\ln U_t^i = V_t (\ln h_t^i - m_t) + W_t$ and V_t is independent of $\{\theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$ all agree on the optimal path of $\{\mathfrak{s}_{t+k}\}_{k=0}^{\infty}$, namely the one that maximizes aggregate welfare W_t . By (B7), this yields $\mathfrak{s}_{t+k} = \mathfrak{s}$ once again.

C Aggregate Efficiency: the Case of Education Finance

By Proposition 5, the intertemporal utility of an agent i in period zero can be written as:

$$\ln U_0^i = (1 - \rho)\lambda \left(\sum_{t=0}^{\infty} \rho^t \prod_{k=0}^{t-1} p(\tau_k) \right) (\ln h_0^i - m_0) + \sum_{t=0}^{\infty} \rho^t (\ln W_t - \rho \ln W_{t+1}).$$

The *certainty-equivalent* sequence $\{\bar{c}_t^i\}_{t=0}^{\infty}$, to the random stream $\{c_t^i\}_{t=0}^{\infty}$ is now defined as:

$$\ln \bar{c}_t^i - \delta(l_t)^\eta \equiv \lambda \left(\prod_{k=0}^{t-1} p(\tau_k) \right) (\ln h_0^i - m_0) + (W_t - \rho W_{t+1})/(1 - \rho). \quad (C1)$$

Given (A8) and (24), this implies once again that $\ln \bar{c}_t^i$ and $E_0 [\ln c_t^i | h_0^i]$ are related by (36). As to the aggregate certainty-equivalent consumption, it is now equal to:

$$\ln \bar{C}_t \equiv \ln \left(\int_0^1 \bar{c}_t^i di \right) = \int_0^1 \ln \bar{c}_t^i di + \prod_{k=0}^{t-1} p(\tau_k)^2 \lambda^2 \Delta_0^2 / 2. \quad (C2)$$

Finally, economic efficiency is defined as before: $E_0 \equiv (1 - \rho) \sum_{t=0}^{\infty} \rho^t [\ln \bar{C}_t - \delta(l_t)^\eta]$. The following results underlie the claims of Proposition 11 in the text.

Proposition 12 *The aggregate efficiency of a sequence of education finance equalization rates $\{\tau_t\}_{t=0}^{\infty}$ equals:*

$$E_0 = W_0 + (1 - \rho) \left(\sum_{t=0}^{\infty} \rho^t \prod_{k=0}^{t-1} p(\tau_k)^2 \right) \left(\frac{\lambda^2 \Delta_0^2}{2} \right).$$

Under a constant educational policy τ , this becomes:

$$E_0(\tau) = (1 - \rho) \sum_{t=0}^{\infty} \rho^t [\ln y_t - \delta(l_t)^\eta] + \ln(1 - \mathfrak{s}) - \rho \left(\frac{1}{1 - \rho p(\tau)^2} - \frac{r(1 - \rho)}{(1 - \rho p(\tau))^2} \right) \left(\frac{\lambda^2 s^2}{2} \right).$$

Proof of Proposition 12

We can rewrite the equation for W_t at $t = 0$ as:

$$\begin{aligned} W_0/(1 - \rho) &= \sum_{t=0}^{\infty} \rho^t (\ln y_t - \lambda^2 \Delta_t^2 / 2 - \delta(l_t)^\eta + \ln(1 - \mathfrak{s}_t) + r\rho(1 - \rho)^{-1} V_{t+1}^2 s^2 / 2) \Rightarrow \\ E_0/(1 - \rho) &= \sum_{t=0}^{\infty} \rho^t (\ln y_t - \delta(l_t)^\eta + \ln(1 - \mathfrak{s}_t)) \\ &\quad - \sum_{t=0}^{\infty} \rho^t \lambda^2 \left[r\rho(1 - \rho) \left(\frac{1}{1 - \rho p(\tau_t)} \right)^2 \left(\frac{s^2}{2} \right) - \frac{\Delta_t^2}{2} + \left(\prod_{k=0}^{t-1} p(\tau_k)^2 \right) \frac{\Delta_0^2}{2} \right]. \end{aligned}$$

Substituting (A13) into the last term, we obtain the analogue to (A14) but with the $(1 - \tau_t)^2$'s replaced by 1. Under a constant policy, in particular, this yields the claimed result. Similar derivations can be carried out for any social welfare function associated to an interpersonal elasticity of substitution σ , hence the results which parallel those of Proposition 10.

References

- Acemoglu, D. and Shimer, R. (1998) "Efficient Unemployment Insurance," NBER Working Paper 6686, August.
- Aghion, P. and Bolton, P. (1997) "A Trickle-Down Theory of Growth and Development with Debt Overhang," *Review of Economics Studies*, 64(2), 151–172.
- Aiyagari, R. (1993) "Optimal Income Taxation With Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 32,
- Aiyagari, R. and Peled, D. (1995) "Social Insurance and Taxation Under Sequential Majority Voting and Utilitarian Regimes," Federal Reserve Bank of Minneapolis Research Department Staff Report, August.
- Ashenfelter, O. (1984) "Macroeconomic and Microeconomic Analyses of Labor Supply," *Carnegie–Rochester Conference Series on Public Policy*, Autumn, 21, 17-157.
- Atkeson, A. and Lucas, R.E., Jr. (1995) "Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance," *Journal of Economic Theory*, 66 (1), 64–88.
- Atkinson, A. (1970) "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–263.
- Atkinson, A. and Stiglitz, J. (1980) "Lectures on Public Economics," Mc–Graw Hill, New York.
- Banerjee, A. and Newman, A. (1991) "Risk-Bearing and the Theory of Income Distribution," *Review of Economic Studies*, 58, 211-236.
- Banerjee, A. and Newman, A. (1993) "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101, 274-298.
- Barro, R., Mankiw, G. and Sala-i-Martin, X. (1995) "Capital Mobility in Neoclassical Models of Growth," *American Economic Review*, 85, 1, 103-115.
- Bassetto, M. (1998) "Optimal Taxation with Heterogenous Agents," University of Chicago mimeo.
- Bénabou, R. (1996a) "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and Education Finance," *American Economic Review*, 86(3), 584–609.
- Bénabou, R. (1996b) "Unequal Societies: Income Distribution and the Social Contract," NBER Working Paper 5583, May.
- Bénabou, R. (1996c) "Inequality and Growth," *NBER Macroeconomics Annual*, Ben Bernanke

and Julio Rotemberg eds., MIT Press, 11–74.

Bishop, J., Formby, J., and Smith, J. (1991) “Lorenz Dominance and Welfare: Changes in the U.S. Distribution of Income, 1967–1986,” *Review of Economics and Statistics*, 73, 1, 134–139.

Browning M., Hansen L. and Heckman, J. (1999) “Micro Data and General Equilibrium Models,” forthcoming in the *Handbook of Macroeconomics*, John Taylor and Michael Woodford eds., North Holland.

Castañeda, A., Díaz–Jimenez J. and Ríos–Rull, V. (1998) “Earnings and Wealth Inequality and Income Taxation: Quantifying the Trade–Offs of Switching the U.S. to a Proportional Income Tax System,” Pennsylvania University mimeo, August.

Chamley, C. (1985) “Efficient Tax Reform in a Dynamic Model of General Equilibrium,” *Quarterly Journal of Economics*, May, 335–356.

Chamley, C. (1997) “Capital Income Taxation, Wealth Distribution and Borrowing Constraints,” DELTA mimeo, August.

Constantinides, G. and Duffie, D. (1996) “Asset Pricing With Heterogenous Consumers,” *Journal of Political Economy*, 104, 219–240.

Cooper, S. (1992) “A Positive Theory of Income Redistribution,” *Journal of Economic Growth*, 3, 171–195.

Cooper, S., Durlauf, S. and Johnson, P. (1994) “On the Evolution of Economic Status Across Generations,” *American Statistical Association, Business and Economics Section, Papers and Proceedings*, 50–58.

Daveri, F. and Tabellini, G. (1997) (1997) “Unemployment, Growth and Taxation in Industrial Countries,” CEPR Discussion Paper 1961, August.

Dewey, J., Husted, T. and Kenny, L. (1998) “The Ineffectiveness of School Inputs: A Product of Misspecification,” University of Florida mimeo, June.

Durlauf, S. (1996) “A Theory of Persistent Income Inequality,” *Journal of Economic Growth*, 1, 75–94.

Easterly, W. and Rebello, S. (1993) “Marginal Income Tax Rates and Economic Growth in Developing Countries,” *European Economic Review*, 37, 409–417.

Engel, E., Galetovic, E. and Raddatz, C. (1998) “Taxes and Income Distribution in Chile: Some Unpleasant Redistributive Arithmetic” NBER Working Paper 6828, December.

Epstein, L. and Zin, S. (1989) “Substitution, Risk Aversion, and the Temporal Behavior of Asset Returns: A Theoretical Framework,” *Econometrica* 57(4), 937–969.

Feldstein, M. (1969) “The Effects of Taxation on Risk Taking,” *Journal of Political Economy*, 77, 755–764.

Fernandez, R. and Rogerson, R. (1996) “Income Distribution, Communities and the Quality of Public Education,” *Quarterly Journal of Economics*, 101(1) 35–164.

Fernandez, R. and Rogerson, R. (1998) “Public Education and the Dynamics of Income Distribution: A Quantitative Evaluation of Education Finance Reform,” *American Economic Review*, 88(4), 813–833.

Galor, O. and Zeira, J. (1993) “Income Distribution and Macroeconomics,” *Review of Economic Studies*, 60, 35–52

Glomm, G. and Ravikumar, B. (1992) “Public vs. Private Investment in Human Capital: Endogenous Growth and Income Inequality,” *Journal of Political Economy*, 100, 818–834.

Gouveia, M. and Strauss, R. (1994) “Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis,” *National Tax Journal*, June, 337–339.

Gradstein M. and Justman, M. (1997) “Democratic Choice of an Education System: Implications for Growth and Income Distribution,” *Journal of Economic Growth*, 2, 169–183.

Haley W. (1976) “Estimation of the Earnings Profile from Optimal Human Capital Accumulation,” *Econometrica*, 44, 1223–38.

Hansen, G. and İmrohoğlu, A. (1992) “The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard,” *Journal of Political Economy*, 100 (1), 118–142.

Heaton J. and Lucas D. (1996) “Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing,” *Journal of Political Economy*, 104, 443–487.

Heckman J. (1976) “A Life–Cycle Model of Earnings, Learning, and Consumption,” *Journal of Political Economy*, (84), S11–S44.

Hendricks, L. (1998) “Taxation and Long–Run Growth,” forthcoming, *Journal of Monetary Economics*.

Hopenhayn, H. and Nicolini, J.P. (1997) “Optimal Unemployment Insurance,” *Journal of Political Economy*, 105 (2), 412–438.

Hoxby, C. (1998) “All School Finance Equalizations Are Not Created Equal,” NBER Working Paper 6792, November.

Hoxby, C. (1998b) “How Much Does School Spending Depend on Family Income? The Historical Origins of the Current School Finance Dilemma,” *American Economic Review Papers and Proceedings*, 88(2), 309–314.

Inman, R. (1978) "Optimal Fiscal Reform of Metropolitan Schools: Some Simulation Results," *American Economic Review*, 68, 107-122.

Jones, L., Manuelli, R. and Rossi, P. "Optimal Taxation in Models of Endogenous Growth," *Journal of Political Economy*, 101(3), 485-517.

Jorgenson, D., Gollop, F. and Fraumeni, B. (1987) *Productivity and US Economic Growth*, Cambridge, MA: Harvard University Press.

Judd, K. (1985) "Redistributive Taxation in a Perfect Foresight Model," *Journal of Public Economics*, 28, 59-83.

Kanbur, S.M. (1979) "Of Risk Taking and the Personal Distribution of Income," *Journal of Political Economy* 87(4), 769-797.

Killingsworth, M.K. and Heckman, J. (1986) "Female Labor Supply: A Survey," Chapter 2 in *Handbook of Labor Economics*, Ashenfelter O.C. and Layard R. eds., vol. 1., North-Holland, New York.

Kreps, D. and Porteus, E. (1979) "Dynamic Choice Theory and Dynamic Programming," *Econometrica*, 47, 91-100.

Krusell, P., and Smith, A. (1998) "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867-896.

Laitner, J. (1992) "Random Earnings Differences, Lifetime Liquidity Constraints and Altruistic Intergenerational Transfers," *Journal of Economic Theory*, 58, 113-170.

Lambert, P. (1993) "The Distribution and Redistribution of Income: A Mathematical Analysis," 2nd. ed. Manchester: Manchester University Press.

Loury, G. (1981) "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49, 843-867.

Lucas, R.E. Jr. (1990) "Supply Side Economics: an Analytical Review," *Oxford Economic Papers*, 42, 293-316

Mulligan, C. (1995) "Some Evidence on the Role of Imperfect Markets for the Transmission of Inequality," University of Chicago mimeo, July.

Pencavel, J. (1986) "Labor Supply of Men", Chapter 1 in *Handbook of Labor Economics*, North-Holland, New York.

Perotti, R. (1993) "Political Equilibrium, Income Distribution, and Growth," *Review of Economic Studies*, 60, 755-776.

Persson, M. (1983) "The Distribution of Abilities and the Progressive Income Tax," *Journal*

of *Public Economics*, 22, 73–88.

Piketty, T. (1997) “The Dynamics of the Wealth Distribution and Interest Rate with Credit-Rationing,” *Review of Economic Studies*, 64(2), 173–190.

Saint-Paul, G. and Verdier, T. (1993) “Education, Democracy and Growth,” *Journal of Development Economics*, 42, 2, 399-407.

Solon, Gary (1992) “Intergenerational Income Mobility in the United States,” *American Economic Review*, 82, 3, 393-408.

Stockey, N. and Rebello, S. (1995) “Growth Effects of Flat Taxes,” *Journal of Political Economy*, 103(3), 519–550.

Tamura, R. (1992) “Efficient Equilibrium Convergence: Heterogeneity and Growth,” *Journal of Economic Theory*, 58(2), 355–376.

Tamura, R. (1998) “Teachers, Growth and Convergence,” Clemson University mimeo.

Varian, H. (1989) “Redistributive Taxation as Social Insurance,” *Journal of Public Economics* 14, 49–62.

Weil, P. (1990) “Nonexpected Utility in Macroeconomics,” *Quarterly Journal of Economics*, 105 (1), 29-42.

Zimmerman, D. (1992) “Regression Toward Mediocrity in Economic Stature,” *American Economic Review*, 82, 3, 409-429.