A Pareto Efficiency Rationale for the Welfare State

Dilip Mookherjee and Stefan Napel[1]

Preliminary Draft: October 16, 2013

Abstract

Is there a Pareto improving rationale for a welfare state in dynamic economies with unobservable agent heterogeneity, when missing credit and insurance markets affect incentives to invest in human capital? If so, should the state provide transfers to the poor in the form of cash or in-kind transfers? In an occupational choice model, we show (a) every competitive equilibrium is interim-Pareto dominated by a policy providing education subsidies financed by income taxes, and (b) transfers conditional on educational investments similarly dominate unconditional transfers. The policies also result in macroeconomic improvements (higher per capita income and upward mobility, lower wage dispersion).

[1]Boston University, and University of Bayreuth respectively. We acknowledge helpful comments by participants of the Oslo THRED conference in June 2013, to Jianjin Miao for directions to the macro literature, and to Brant Abbott for providing us with details of the nature of inter vivos transfers in Abbott *et al.* (2013). We also thank Marcello d'Amato for numerous discussions and ideas concerning occupational choice and role of the welfare state.

1 Introduction

A central issue in discussions of a welfare state concerns its normative rationale, given incompleteness of financial markets which restrict the ability of households to borrow in order to finance education of their children, or insure against idiosyncratic risks. Standard theorems of welfare economics which relate Pareto optimal allocations and competitive equilibria are based on the assumption of complete markets, and therefore do not apply in such settings. Can anything be said at some level of generality regarding constrained Pareto efficiency (or lack thereof) of competitive equilibria when credit and insurance markets are missing, and budgetary and informational constraints on government policy are incorporated?

Dynamic models of investment in physical and/or human capital which incorporate missing credit and insurance markets and agent heterogeneity have been studied in the literature on macroeconomics based on specific functional forms for technology and preferences (Aiyagari (1994), Aiyagari, Greenwood and Sheshadri (2002), Bénabou (2002)). Versions of these models have been calibrated to fit data of real economies and numerically simulated to evaluate the welfare and macroeconomic effects of various fiscal policies (Heathcote (2003), Cespedes (2011), Abbott, Gallipoli, Meghir and Violante (2013)). Yet there is no clear answer available concerning the question whether there generally exist fiscal policies that result in Pareto improvements over laissez faire competitive equilibria.

Most of the theoretical occupational choice literature (e.g., Loury (1981), Banerjee and Newman (1993), Galor and Zeira (1993), Ljungqvist (1993), Ghatak and Jiang (2002), Matsuyama (2000, 2003), Mookherjee and Napel (2007)) does not provide any results concerning Pareto efficiency properties of competitive equilibria. An exception is Mookherjee and Ray (2003), who characterize efficiency properties of steady states in a model with homogenous ability, and show existence of steady states that are constrained Pareto-efficient. Whether this result extends to the case of heterogeneous ability has not been addressed.[2]

[2] See, however, D'Amato and Mookherjee (2013) who show in a model with ability heterogeneity, missing financial markets and job market signaling that laissez faire outcomes are Pareto dominated by a policy of educational loans to the poor, funded by bonds subscribed to by wealthy households.

A subsequent question is optimal design of fiscal policy, e.g., whether and how transfers to households should be designed. A particular question is whether transfers to poor households should be uniform/cash/unconditional rather than in-kind/conditional on wealth status and investments in human capital. While there are general arguments based on the Pareto criterion in favor of the former in static contexts — as in the Mirrlees (1971) or Atkinson-Stiglitz (1976) models – matters are more complicated in dynamic settings when effects on investments need to be incorporated. Cespedes (2011) uses a calibrated macro model fitted to data from the Mexican economy to evaluate the macroeconomic and welfare effects of a specific conditional cash transfer (CCT) program restricted to the poor that send their children to school, compared with unconditional lump sum transfers. His results show that the CCT resulted in macroeconomic improvements (higher education, per capita income, lower poverty and inequality), and a small average welfare improvement where a large majority of households (but not all) were better off. Apart from the need to understand the source of these welfare effects (e.g., evaluating attendant insurance effects), it leaves open the question whether there may exist other CCTs which could have resulted in a Pareto improvement, or what the effects might be in economies with different preferences and technology.

We consider a model with overlapping generations where parents invest in education of their children, and learning abilities of children are heterogeneous. Education cannot be financed by borrowing, and ability risk cannot be insured. Parents are altruistic towards their children, either nonpaternalistically (*à la* Barro and Becker (1989)) or paternalistically (as in Becker and Tomes (1979) or Mookherjee and Napel (2007)). The simple baseline model has only two occupations: unskilled and skilled, with the latter requiring costly investments that depend on the child's ability (which are observed by parents before deciding on how much to invest). Moreover, parents cannot supplement human capital investments with financial bequests. We abstract from endogenous labor supply and income risk (conditional on education). On the other hand, we do not restrict preferences or technology apart from standard assumptions of smoothness, monotonicity, concavity and constant returns to scale. We examine competitive equilibria where parents have perfect foresight concerning future wage rates, and evaluate the effects of fiscal policy where taxes and transfers can be conditioned on parental education/income status and on their spending on children's education.

Fiscal policy in this setting has complex effects on consumption insurance and investment incentives. Ability risk results in non-monotonic consumption fluctuations for instance. Parents spend less on children when they are extraordinarily gifted (whence the cost of education is low) and also when they have extraordinarily low ability (whence the cost of education is high enough that parents do not invest). They spend more when abilities of their children are intermediate. Investment effects are also complicated. Income transfers (funded by taxes on the rich) to the poor relax borrowing constraints that latter households are particularly prone to, reducing their opportunity costs of investing in education. However, by compressing consumption differentials, such fiscal policies lower future (post-tax/transfer) returns to education at the same time. As many authors (e.g., Aiyagari *et al.* (2002)) have shown, lowering uncertainty of future income can have a powerful depressing effect on investment incentives, owing to the reduced urgency of precautionary motives. A related concern with progressive fiscal policies is they may encourage 'welfare-dependence' among the poor, reducing their incentive to invest.

Moreover, it is difficult to ensure Pareto improvements owing to uneven incidence of benefits and taxes across different household groups classified by parental education and child ability realizations. Those with low ability children that do not invest in education obtain no benefits from education subsidies, while they end up paying the taxes that fund such subsidies. Further complications arise from possible general equilibrium feedback effects on wages in different occupations resulting from changes in relative supply of skilled and unskilled households, for which many empirical papers reported above have found strong evidence.

In our baseline model, we establish two analytical results. The first shows that any laissez faire competitive equilibrium is interim-Pareto dominated (using utilities evaluated after parent's education levels are given but before the realization of children's ability is known) by a balanced budget fiscal policy providing educational subsidies that are funded by income taxes. The second shows that any policy involving taxes or transfers not conditioned on educational investments is interim-Pareto dominated by a balanced budget policy involving transfers that are conditioned on these investments. In both cases, the welfare improvement is accompanied by a macroeconomic improvement (higher average education, per capita income, lower inter-occupation wage inequality, and greater upward mobility among the poor).

The main idea underlying these arguments is to focus on fiscal policies where educational subsidies or CCTs to parents in any given occupation are funded by income taxes *on the same occupation.* In other words, we do not focus on redistribution *across* occupations. Education subsidies accrue to those parents with children gifted with sufficient ability that those parents do find it worthwhile to invest. They do not benefit parents whose children are not able enough to justify any investment. Such policies redistribute from the latter 'non-investor' group in favor of the former 'investor' group. The former must consume more than the latter, since they have the same incomes. Hence the policy provides insurance against the uncertainty of the ability draw. At the interim stage, such insurance raises welfare. It is therefore possible for the government to run a budget surplus by offering an 'actuarially fair' policy, which leaves unchanged the interim expected utilities of parents in that occupation.

Such policies will raise educational investments, as the subsidies lower the costs of education, while future benefits of education are preserved by construction (as similar policies are introduced for every future generation). This generates macroeconomic improvements, but the general equilibrium feedback effects in future generations could worsen the welfare of parents in some occupations. These feedback effects can be neutralized by an accompanying set of taxes and transfers so as to leave after-tax wages unchanged in each occupation. Specifically, this will require the original policy to be bundled with a regressive policy in which taxes on unskilled wages fund subsidies on skilled wages. Such bundling allows the policy to generate a budget surplus while leaving interim welfare of all parents in all generations unaffected.

The last stage of the argument requires these surpluses to be rebated by lowering taxes (or raising transfers) to all subgroups of the population classified by parental occupation and investment decision, in a way that raises everyone's welfare by an equal amount, thereby leaving investment incentives unchanged. The final bundle must include this rebate package as well.

In a subsequent section, we show that these arguments extend with alternative formulations of parental altruism, and when educational investments are perfectly divisible. However, they do not extend quite as straightforwardly when human capital investments can be supplemented by financial bequests. In particular, they do not apply to households wealthy (and altruistic) enough that they *always* make financial bequests, irrespective of how much they invest in education. Within such a wealth class, those who do not invest in education end up spending more on their children overall, and thus consume less than parents who do invest in education. This reverses the pattern of consumption variation with respect to realization of children's ability risk — the educational subsidy policy described above now imposes additional consumption risk, and thereby creates a welfare loss. Laissez faire competitive equilibria still continue to be constrained Pareto inefficient, however. The nature of a Pareto improving policy is now reversed: requiring education for the wealthy (as defined above) to be taxed, and these taxes to fund income transfers to the same class. At the other end of the wealth distribution, the nature of the Pareto improving policy is unchanged for poor households who invest if at all only in education and leave no financial bequests. The aggregate macroeconomic effects of such a Pareto improving policy are unclear, as the increased educational investments among the poor will be countered by falling investments among the wealthy.

The paper is organized as follows. Section 2 introduces the baseline model, followed by the main results for this model in Section 3. Extensions are discussed in Section 4. The relation to existing literature is described in Section 5, and Section 6 concludes.

2 Baseline Model

We first describe the dynamic economy in the absence of any government intervention.

There are two occupations: unskilled and skilled (denoted 0, 1 respectively). There is a continuum of households indexed by $i \in [0,1]$. Generations are denoted $t = 0, 1, 2, \ldots$ Each household has one adult and one child in each generation. The utility of the adult in household i in generation t is denoted $V_{it} = u(c_{it}) + \delta V_{i,t+1}$ where c_{it} denotes consumption in household i in generation t, $\delta \in (0,1)$ is a discount factor, and u is a strictly increasing, strictly concave and C^2 function defined on the real line. There is no lower bound to consumption, while u tends to $-\infty$ as c tends to $-\infty$.

Household i earns y_{it} in generation t, and divides this between consumption at t and investment in child education. Education investment I_{it} is indivisible, either 1 or 0. An educated adult has the option of working in either occupation, while an uneducated adult can only work in the unskilled occupation. The ability of the child in household i is represented by how little its parent has to spend to educate it. The cost of education x_{it} in household i in generation t is drawn randomly and independently according to a common distribution function F defined on the nonnegative reals. F is C^2 and strictly increasing; its density is denoted f. Hence the household budget constraint is $y_{it} = c_{it} + x_{it} I_{it}$. Every parent privately observes the realization of education cost of its child before deciding on whether to invest in education.

The key market incompleteness is that parents cannot borrow to finance their children's education. Neither can they insure against the risk that their child has low learning ability, the main source of (exogenous) heterogeneity in the model. The former arises owing to inability of parents to borrow against their children's future earnings. The latter could be due to privacy of information amongst parents regarding the realization of their children's ability.

Household earnings are defined by occupational wages: $y_{it} = w_{0t} + I_{it} \cdot (w_{1t} - w_{0t})$, where w_{ct} denotes the wage in occupation c in generation t determined in a competitive labor market.

Wages are determined as follows at any given date (so we suppress the t subscript for the time being). There is a CRS production function $G(\lambda, 1-\lambda)$ which determines the per

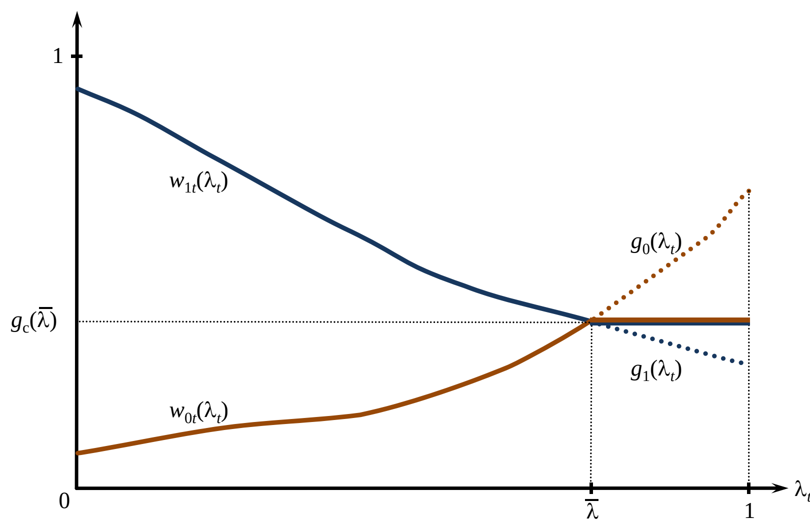

Figure 1: Labor Market

capita output in the economy in any generation t if the proportion of the economy that works in the skilled and unskilled occupations equal λ and $1 - \lambda$ respectively. We assume G is a C^2, strictly increasing, linearly homogenous and concave function. Let $g_c(\lambda)$ denote the marginal product of occupation $c = 0, 1$ workers when λ proportion of adults work in the skilled occupation. So g_1 is decreasing and g_0 is an increasing function. Moreover, $g_1(0) > g_0(0)$ while $g_1(1) < g_0(1)$. To avoid some technical complications we assume the functions g_i are bounded over $[0, 1]$. In other words, the marginal product of each occupation is bounded above even as its proportion in the economy becomes vanishingly small.[3]

Let $\bar{\lambda}$ denote the smallest value of λ at which $g_1(\lambda) = g_0(\lambda)$. Then in any given generation t, all educated workers will prefer to work in the skilled occupation, with $w_{1t} = g_1(\lambda_t)$

[3]When the production function satisfies Inada conditions, we obtain the same results if every household is able to resort to a subsistence self-employment earnings level $\underline{w}$ which is positive and exogenous. As the proportion of unskilled workers tends to one, the labor market will clear at an unskilled wage equal to $\underline{w}$, and the proportion of skilled households working for others will be fixed at a level where the marginal product of the unskilled equals this wage. The only difference is that wages in either occupation as a function of the skill ratio become kinked at the point where where the marginal product of the unskilled equals $\underline{w}$. Except at this single skill ratio, the wage functions are smooth, and our results continue to apply though with an 'almost everywhere' proviso.

and $w_{0t} = g_0(\lambda_t)$, if the proportion of educated adults is $\lambda_t < \bar{\lambda}$. And if $\lambda_t \geq \bar{\lambda}$, equilibrium in the labor market at t will imply that exactly $\bar{\lambda}$ fraction of adults will work in the skilled occupation, as educated workers will be indifferent between the two occupations, and $w_{1t} = w_{0t} = g_1(\bar{\lambda}) = g_0(\bar{\lambda})$. See Figure 1. When more than $\bar{\lambda}$ fraction of adults in the economy are educated, the returns to education are zero. Since education is costly, education incentives vanish if households anticipate more than $\bar{\lambda}$ proportion of adults in the next generation will be educated. Hence the proportion of educated adults will always be less than $\bar{\lambda}$ in any equilibrium with perfect foresight. We can identify the occupation of each household i in generation t with its education status $I_{i,t-1}$, and refer to λ_t as the skill ratio in the economy in generation t.

2.1 Dynamic Competitive Equilibrium under Laissez Faire

Definition 1 *Given a skill ratio fraction λ_0 in generation 0, a* dynamic competitive equilibrium under laissez faire (DCELF) *is a sequence $\{\lambda_t\}_{t=0,1,2,...}$ of skill ratios and investment strategies $\{I_{ct}(x)\}_{t=0,1,2,...}$ for every household in occupation c in generation t when its child's education cost happens to be x such that:*

(a) For each household and each t: $I_{ct}(x) \in \{0,1\}$ maximizes

$$u\big(w_{ct} - I_{ct}x\big) + \delta \mathsf{E}_{\tilde{x}} V_{t+1}(I_{ct}, \tilde{x}) \tag{1}$$

and the resulting value is $V_t(c,x)$.

(b)

$$\lambda_t = \lambda_{t-1}\mathsf{E}_x[I_{1t}(x)] + (1 - \lambda_{t-1})\mathsf{E}_x[I_{0t}(x)]. \tag{2}$$

(c) Every household correctly anticipates $w_{ct} = g_c(\lambda_t)$ for occupation $c = 0,1$ in generation t.

It is useful to note the following features of a DCELF.

Lemma 1 *In any DCELF and at any date t:*

(i) $V_t(1, x) > V_t(0, x)$ *for all* x *if and only if* $\lambda_t < \bar{\lambda}$.

(ii) $\lambda_t \in (0, \bar{\lambda})$, $w_{1t} > w_{0t}$.

(iii) $I_{ct}(x) = 1$ *iff* $x < x_{ct}$, *where* x_{ct} *is defined by*

$$u(g_c(\lambda_t)) - u(g_c(\lambda_t) - x_{ct}) = \delta[W_{1,t+1} - W_{0,t+1}] \tag{3}$$

and $W_{ct} \equiv \mathsf{E}_x V_t(c, x)$

(iv) The investment thresholds satisfy $x_{0t} < x_{1t}$, *are uniformly bounded away from 0, and uniformly bounded above, while* λ_t *is uniformly bounded away from* 0 *and* $\bar{\lambda}$ *respectively. Consumptions of all agents are uniformly bounded.*

This Lemma shows that skilled wages always exceed unskilled wages, and those in skilled occupations always have higher utility. There is inequality of educational opportunity: children born to skilled parents are more likely to be educated. There is also upward and downward mobility: some talented children born to unskilled parents do receive an education, while some untalented children born to skilled parents fail to receive an education. Property (iv) shows that consumptions and utility differences are bounded, which will be useful in our subsequent analysis.

2.2 Competitive Equilibrium with Taxes

We now extend the model to incorporate fiscal policies. The government observes the occupation/income of parents as well as the education decisions they make for their children. Transfers can accordingly be conditioned on these. Fiscal policy is represented by four variables: $\tau_{1t}, \tau_{0t}, e_{0t}, e_{1t}$ in any generation t which are income transfers based respectively on parental occupation and education investment decision. In particular, the government does not observe the ability realization of any given child, the key informational constraint that prevents attainment of a first-best utilitarian optimum. We are also focusing on transfers that depend only on the current status of the household, thus ruling out educational loans. Similar to private agents, the government will also not be able to lend or borrow across generations, and will have to balance its budget within each generation.

Definition 2 *Given a skill ratio fraction λ_0 in generation* 0*, a* dynamic competitive equilibrium (DCE) given fiscal policy $\{\tau_{1t}, \tau_{0t}, e_{0t}, e_{1t}\}_{t=0,1,2,...}$ *is a sequence* $\{\lambda_t\}_{t=0,1,2,...}$ *of skill ratios and investment strategies* $\{I_{ct}(x)\}_{t=0,1,2,...}$ *for every household in occupation c in generation t when its child's education cost happens to be x such that for each c, t:*

(a) $I_{ct}(x) \in \{0, 1\}$ *maximizes*

$$u\big(w_{ct} + \tau_{ct} - I_{ct}\{x - e_{ct}\}\big) + \delta \mathsf{E}_{\tilde{x}} V_{t+1}(I_{ct}, \tilde{x}) \tag{4}$$

and the resulting value is $V_t(c, x)$.

(b)

$$\lambda_t = \lambda_{t-1}\mathsf{E}_x I_{1t}(x) + (1 - \lambda_{t-1})\mathsf{E}_x I_{0t}(x). \tag{5}$$

(c) Every household correctly anticipates $w_{ct} = g_c(\lambda_t)$ for occupation $c = 0, 1$ in generation t.

The government balances its budget if at every t it is the case that

$$\lambda_t\{\tau_{1t} + e_{1t}\mathsf{E}[I_t(1, x)]\} + (1 - \lambda_t)\{\tau_{0t} + e_{0t}\mathsf{E}[I_t(0, x)]\} \leq 0. \tag{6}$$

A DCELF with a (trivially) balanced budget obtains as a special case of a DCE when the government selects zero income transfers and educational subsidies.

It is easy to check that a DCE can also be described by investment thresholds x_{ct} satisfying the following conditions. Define the interim expected utility of consumption of a parent in occupation c in generation t as follows:

$$\mathcal{U}_{ct} \equiv u(w_{ct} + \tau_{ct})[1 - F(x_{ct})] + \int_0^{x_{ct}} u(w_{ct} + \tau_{ct} + e_{ct} - x)dF(x) \tag{7}$$

The thresholds must then satisfy

$$u(w_{ct}+\tau_{ct})-u(w_{ct}+\tau_{ct}+e_{ct}-x_{ct})=\delta\cdot\Delta W_{t+1} \tag{8}$$

where

$$\Delta W_t \equiv W_{1,t}-W_{0,t}=\sum_{k=0}^{\infty}\nu_k[\mathcal{U}_{1,t+k}-\mathcal{U}_{0,t+k}] \tag{9}$$

with $\nu_0=1$, $\nu_k=\delta^k\Pi_{l=0}^{k-1}[F(x_{1,t+l})-F(x_{0,t+l})]$ for $k\geq 1$. A DCE can be equivalently described by $\{\lambda_t,w_{1t},w_{0t},x_{1t},x_{0t},\mathcal{U}_{1t},\mathcal{U}_{0t}\}_{t=0,1,2,\dots}$ which satisfy (5) and (7)–(9), besides condition (c) above.

3 Results for the Baseline Model

Our first result is an efficiency as well as a macroeconomic role for fiscal policy. The efficiency criterion is *interim Pareto dominance*, which requires expected utility of every parent $W_{ct}\equiv \mathsf{E}_x V_t(c,x)$ to be higher for every c,t. The criterion of *macroeconomic dominance* is that the skill ratio λ_t must be higher at every t, and the investment threshold x_{ct} must be higher for every c,t. This ensures higher per capita skill and output at every date, as well as greater educational opportunity (in the sense of a higher probability for every child to become educated, both conditional on parent's occupation, and unconditionally).

Theorem 1 *Consider any DCELF starting from an arbitrary skill ratio λ_0 at $t=0$. There exists a balanced budget fiscal policy with educational subsidies for each occupation funded by income taxes, and an associated DCE which interim-Pareto (as well as macroeconomically) dominates the original DCELF.*

Our second result is that any fiscal policy involving income transfers alone is dominated in a similar way by a policy with educational subsidies.

Theorem 2 *Consider any DCE given an initial skill ratio λ_0 and a balanced budget fiscal policy consisting of income transfers alone ($e_{ct} = 0$ for all c, t), satisfying the following conditions:*

(a) $\tau_{0t} \geq \tau_{1t}$ for all t;

(b) there exists $\kappa > 0$ such that $-[\tau_{1t} - \tau_{0t}] < [g_1(0) - g_0(0)] - \kappa$ for all t;

(c) τ_{ct} is uniformly bounded.

Then there exists another balanced budget fiscal policy consisting of income transfers combined with educational subsidies ($e_{ct} > 0$ for all c, t) and an associated DCE which interim Pareto (as well as macroeconomically) dominates the original DCE.

Condition (a) requires the income transfers to be progressive in the weak sense that unskilled parents receive a higher transfer (or pay a lower tax), while (b) restricts the marginal tax rate to be less than (and bounded away from) 100%. Condition (c) is a technical restriction needed to ensure that competitive equilibria always involve bounded consumptions and investment thresholds. The role of (b) is to ensure that skilled households earn more both before and after government transfers, so agents always have investment incentives (that are bounded away from zero). Condition (a) ensures that the direct effect of any reduction in the proportion of unskilled households is to weaken the government budget balance constraint. These conditions imply that equilibria with fiscal policy continue to satisfy the same properties as equilibria under laissez faire that were shown in Lemma 1.

It is evident that Theorem 1 is a special case of Theorem 2. The latter result can be generalized further when the status quo policy includes educational taxes. In this case we need to impose the additional condition that these educational taxes are not large enough to destroy investment incentives. Specifically, the same result can be established when the status quo policy and DCE satisfy the following conditions:

(a) the fiscal policy is *progressive* in the sense that $\tau_{0t} + F(x_{0t})e_{0t} \geq \tau_{1t} + F(x_{1t})e_{1t}$ for all t;

(b) there exists $\kappa > 0$ such that $-[\tau_{1t} - \tau_{0t}] < [g_1(0) - g_0(0)] - \kappa$ for all t;

(c) τ_{ct}, e_{ct} are uniformly bounded;

(d) for some c: $e_{ct} \leq 0$ for all t,[4] while $\sup_{c,t}\{-e_{ct}\}$ is not too large so as to ensure x_{ct} is uniformly bounded away from zero.

In all these versions, the new policy provides educational subsidies to a given occupation which are funded by income taxes levied on the same occupation. The proof constructs a small increase $\epsilon(1-\mu_t)$ in education subsidy for occupation c which is financed by an increase $\epsilon F(x^*_{ct})$ in income taxes on this occupation.

The key observation is that parents' consumption varies with the realization of their child's ability. The nature of this variation is shown in Figure 2. It is nonmonotone with respect to x, the cost necessary to educate the child. If the child is a genius and can be costlessly educated, the parents consumption equals his earning. The same is true when the child has low enough ability that it is not educated. For intermediate abilities where the child is educated, the parent invests a positive amount, lowering consumption. Hence parental consumption varies non-monotonically with respect to x.

The educational subsidy introduced raises the consumption of the investors, while financing it by income taxes on the same occupation, which lowers the consumption of the non-investors. See Figure 3. If μ_t were zero, this would reduce risk faced in consumption associated with realization of the child's education cost (assuming that education is not subsidized to start with and hence non-investing households consume more than every household in the same occupation that does invest). This would result in a mean-preserving reduction in riskiness of parental consumption, thus raising the interim expected utility of current consumption in occupation c. The parameter μ_t is set so as to reduce the mean

[4] We can also relax the requirement 'for all t' to 'for some t'.

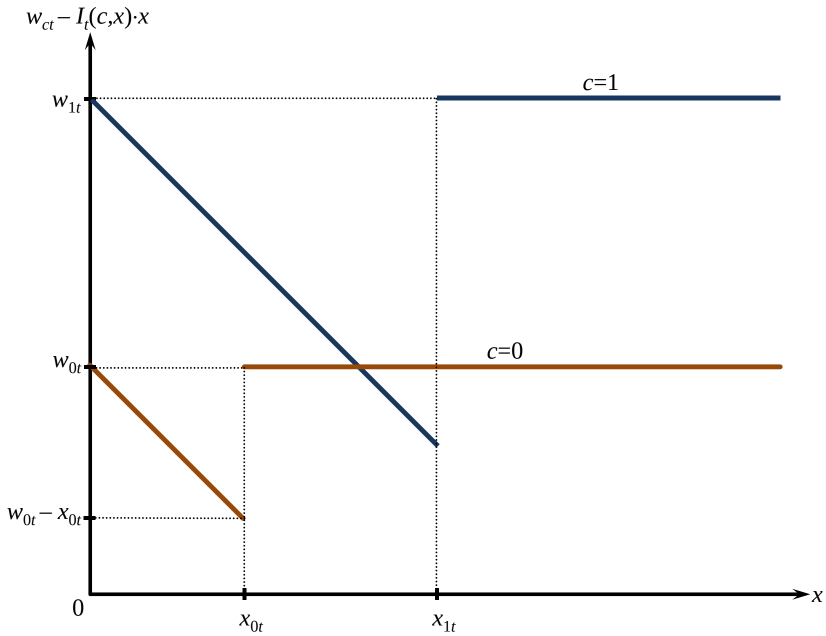

Figure 2: Variation of Parental Consumption with Education Cost

consumption enough that there is no change in the expected utility of current consumption at date t for each occupation.

Assuming wages are unchanged, this implies that dynastic utilities of both occupations are unchanged. Hence the future benefit of investment is unchanged. The subsidization of education in occupation c on the other hand lowers the sacrifice parents must endure to educate their children. Hence households invest more often. Aggregate investment in the economy will then rise, which will tend to lower skilled wages and raise unskilled wages. These general equilibrium changes will tend to reduce the benefits of investment.

Next fiscal policy is adjusted further to neutralize the wage changes, so as to preserve the rise in investment. This results in a new competitive equlibrium sequence with a higher skill ratio at every date, and a zero first-order effect on interim utilities. However, the government has a first order improvement in its surplus, owing to the rise in the skill ratio and the extraction of resources from households by setting $\mu_t > 0$. The progressivity of the original fiscal policy implies that the government budget surplus improves as a result of the

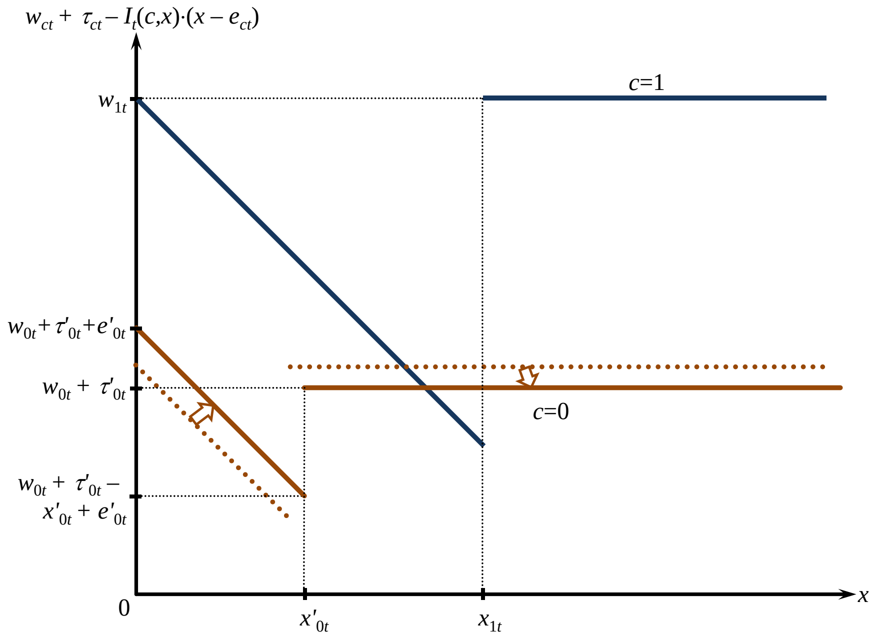

Figure 3: Effects of Steps 1 and 2 of Fiscal Policy Variation on Parental Consumption

decline in the proportion of unskilled households.

In the last step of the argument the government constructs another variation in its tax-subsidy policy to distribute the additional revenues to construct a strict Pareto improvement, while preserving investment incentives. Note that by construction after-tax wages have remained unchanged in each occupation. So wage dispersion between occupations is unchanged, while a fraction of agents move up from the unskilled to the skilled occupation in every generation.

4 Extensions

4.1 Paternalistic Altruism

Suppose parents do not have Barro-Becker dynastic preferences, and instead value (only) the earnings of their children according to a given increasing function $Y(w_{t+1})$, as in Becker and Tomes (1979) or Mookherjee and Napel (2007). A parent in occupation $c \in \{0, 1\}$ at date t with a child who costs x to educate then selects $I \in \{0, 1\}$ to maximize $u(w_{ct} - Ix) + IY(w_{1,t+1}) + (1 - I)Y(w_{0,t+1}))$. All preceding results continue to extend with this formulation of parental altruism. The wage neutralization policy preserves after-tax wages in each occupation, whence the altruistic benefit of investments remain unchanged. The costs of investing are lowered by providing educational subsidies, and at the same time riskiness of parental consumption is lowered. So investment incentives continue to rise, while enhancing interim expected utilities.

4.2 Continuous Investment Choices

What if educational investments can be varied continuously, rather than being indivisible? Our results can be shown to extend straightforwardly to this context, as we now explain.

Let the extent of education be described by a compact interval $E \equiv [0, \bar{e}]$ of the real line. Assume that the relation between wage earnings and education is given by a real-valued continuous function $w(e)$ defined on E. If the earnings function depends endogenously on the supply of workers with varying levels of education, the analysis can be extended using a similar strategy of following up on educational subsidy policies that increase the supply of more educated workers with a wage-neutralization policy that leaves the after-tax wage pattern unchanged. To illustrate how our results extend, it therefore suffices to take the earnings-education pattern in the status quo equilibrium as given.

Let $I(e'; x)$ denote the expenditure that must be incurred by a parent to procure education $e' \geq 0$ for its child whose learning ability gives rise to a learning cost parameter x which varies according to a distribution with full support on $[0, \infty)$, similar to the preceding section. The function I is strictly increasing and differentiable in both arguments. It satisfies $I(0; x) = 0$ for all x, while for any given $e' \geq 0$ the marginal cost $\frac{\partial I}{\partial e'}$ is increasing in x, approaching ∞ as $x \to \infty$.

The value function of a parent with education e and a child whose learning cost parameter is x is then

$$V(e|x) \equiv \max_{0 \leq e' \leq \bar{e}} \left[u(w(e) - I(e'; x)) + \delta W(e') \right] \tag{10}$$

where $W(e') \equiv \mathsf{E}_{\tilde{x}} V(e'|\tilde{x})$. Let the corresponding policy function be $e'(e; x)$. Given that wages are bounded above by $w(\bar{e})$, consumptions are also bounded above. Given this and the feature that u is unbounded below, consumptions can be bounded from below almost surely.[5] Hence the marginal utility of consumption is bounded almost surely, implying that $W'(0) \equiv \mathsf{E}_{\tilde{x}}[u'(w(0) - I(e'(0; \tilde{x}))] $ is bounded.

We can therefore define $x^*(e)$ as the solution for x in the equation $\frac{\partial I(0;x)}{\partial e'} = \frac{\delta W'(0)}{u'(w(e))}$. Then the optimal policy function takes the form $e'(e; x) = 0$ if $x \geq x^*(e)$ and positive otherwise.[6] In other words, parents decide to acquire no education for their children if and only if their learning cost parameter is larger than a threshold $x^*(e)$. These 'non-investors' consume their entire earnings $w(e)$ — just like those parents with the same education e whose children have a learning cost parameter of $x = 0$. For those whose children have intermediate learning ability, parents spend a positive amount on education.

We thus have a similar non-monotone pattern of variation of parental consumption with their children's learning costs, as in the two-occupation case. This ensures that a similar policy of educational subsidy funded by income taxes on all parents with the same education will reduce the riskiness of parental consumption, and thereby permit a Pareto improvement.

[5]Any policy where consumption approaches $-\infty$ with positive probability will be dominated by a policy where parents never invest.

[6]This follows since the value function is concave, owing to a direct argument.

The essential argument is thus simple. Non-investing parents within any given occupation will by definition consume more than investing parents. The educational subsidy funded by the income tax on this occupation then redistributes consumption away from those consuming high amounts to those consuming less. Since these consumption variations arise from the 'ability lottery' of their children, the policy increases interim expected utilities of each occupation. The preceding details of the model were needed to ensure that there is a positive mass of investors and non-investors respectively, so as to allow a strict Pareto improvement.

4.3 Financial Bequests

There is however one important assumption underlying the above reasoning: that educational investments constitute the sole means by which parents transfer wealth to their children. In practice parents have other means as well, such as leaving them financial bequests or physical assets. The simple logic then breaks down: a parent that does not invest in his child's education owing to low learning ability of the latter could provide financial bequests instead. It no longer follows that education non-investors invest less when we aggregate across different forms of intergenerational transfers.

We now consider the consequences of allowing parents to leave financial bequests besides investing in their children's education. To simplify matters, suppose that the rate of return $(1 + r)$ on financial bequests is exogenously given, as in Becker and Tomes (1979) or Mookherjee and Ray (2010). This could correspond to a globalized capital market where the savings of any given country leaves the interest rate unaffected. Even if the interest rate depends on the supply of savings, a 'neutralization' policy allows policy-makers to ensure that the after-tax interest rate is unchanged. For the same reason we abstract from general equilibrium effects in the labor market and suppose that wages of different occupations are exogenously given.

Let us further simplify to the case of two occupations, skilled and unskilled, where the

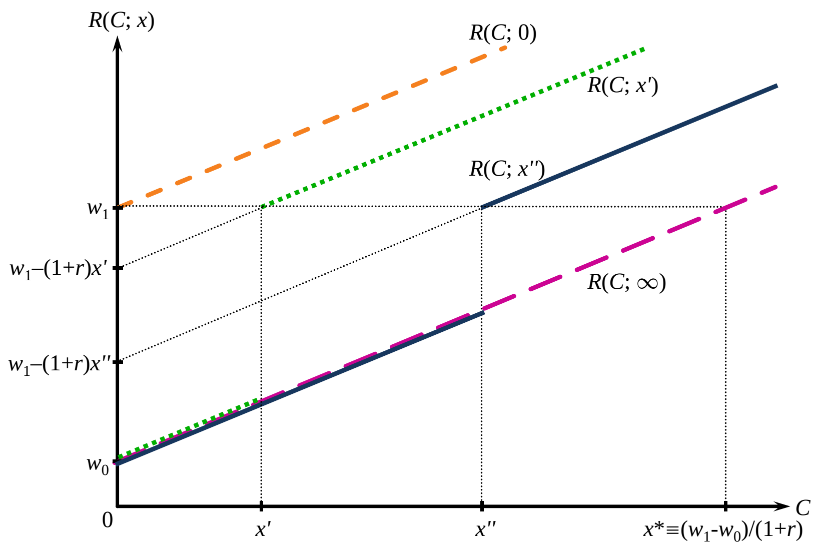

Figure 4: Child wealth as function of total investment expenditure C, given cost x

education cost of the former is denoted x and the latter equals zero. And suppose that parental altruism is paternalistic, where a parent with lifetime wealth W and education cost x chooses financial bequest $b \geq 0$ and education investment $I \in \{0, 1\}$ to maximize $u(W - b - Ix) + \delta Y(W')$ where Y is a strictly increasing and strictly concave function of the child's future wealth $W' = (1 + r)b + Iw_1 + (1 - I)w_0$.

This problem can be reformulated as follows. Let $C \equiv b + Ix$ denote the total parental investment expenditure on his child. An efficient way to allocate C across financial bequest and educational expenses is the following: $I = 0$ if either $C < x$, or $C \geq x$ and the rate of return on education is dominated by the return on financial assets: $\frac{w_1 - w_0}{x} < 1 + r$. Conversely, if the rate of return on education exceeds r and $C \geq x$, then $I = 1$, and $b = C - x$. Then the child ends up with wealth $W' \equiv R(C; x)$ given by

$$R(C; x) = \begin{cases} (1 + r)C + w_0 & \text{if } C < x \text{, or } C \geq x \text{ and } \frac{w_1 - w_0}{x} \leq 1 + r, \\ (1 + r)C + w_1 - (1 + r)x & \text{if } C > x \text{ and } \frac{w_1 - w_0}{x} > 1 + r; \end{cases} \tag{11}$$

and illustrated in Figure 4.

Define the *BT (Becker-Tomes) bequest* as the optimal bequest of a parent in the absence

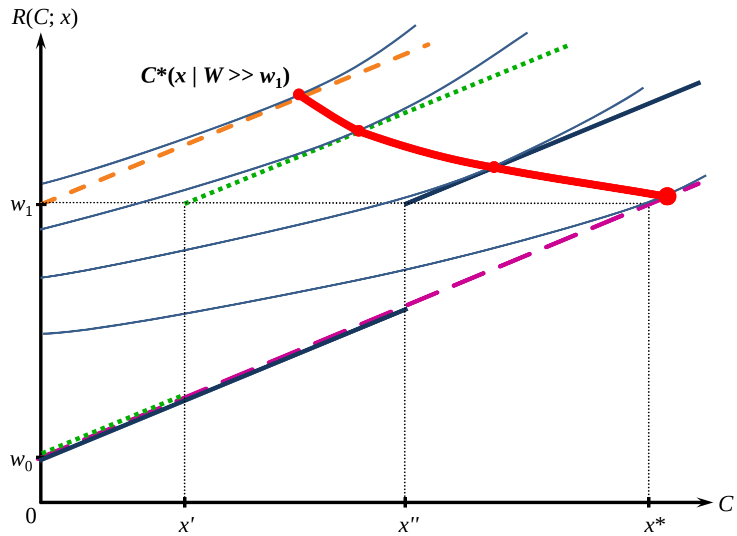

Figure 5: Investment expenditures of sufficiently wealthy parents (case A)

of any opportunity to invest in education, with a given flow earning w of the child when the parent leaves a zero bequest. This is the problem of choosing $C \geq 0$ to maximize $u(W - C) + \delta Y((1 + r)C + w)$. Denote the BT bequest by $C^{BT}(W; w)$. It is easily checked that this is increasing in parental wealth W and decreasing in w.

Recall that a parent will invest in education only if the child has enough ability to ensure that $x \leq x^* \equiv \frac{w_1 - w_0}{1+r}$. Whenever $x > x^*$, there will be no investment in education, and the optimal bequest equals the BT bequest $C^{BT}(W; w_0)$. When $x < x^*$, the optimization problem entails a nonconvexity and the solution is more complicated. The dotted and solid lines in Figure 4, for instance, respectively represent the nonconvex sets of feasible (C, W')-combinations for parents with children whose education costs x' and x'' lie below x^*.

Nevertheless we can illustrate the solution for some extreme cases, corresponding to different parental wealths.

Case A. W *sufficiently large:* Suppose W is large enough that $C^{BT}(W; w_1 - (1 + r)x) > x$

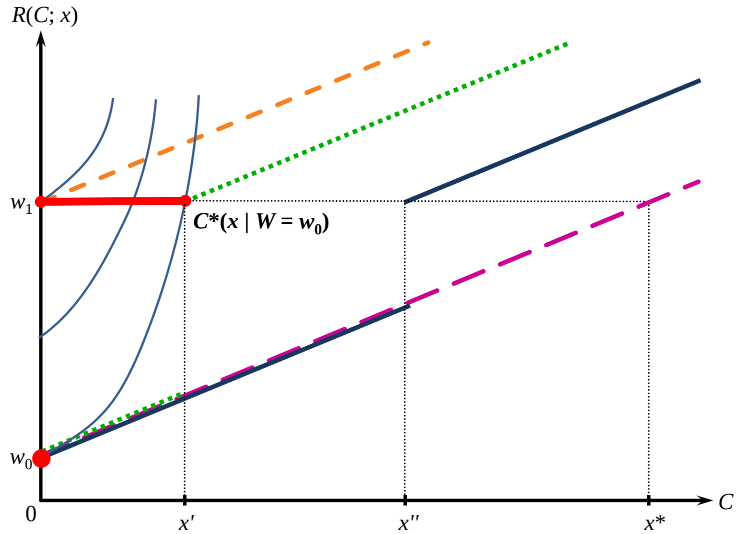

Figure 6: Investment expenditures of poor parents (case B)

for all $x \leq x^*$.[7] In words, irrespective of where x lies below x^*, the parent will always supplement education investments with a financial bequest. See Figure 5.

Case B. *W sufficiently small:* Suppose $W = w_0$, $\delta(1+r) \leq 1$ and $Y \equiv u$. Then the BT bequest $C^{BT}(w_0; w) = 0$ for all $w \geq w_0$, and the parent will never make a financial bequest. If however the child learning cost x is sufficiently small, the parent will invest in education. The optimal choice of expenditure C^* is illustrated in Figure 6, where the low parental wealth is reflected by very steep indifference curves.

The implied consumption patterns of sufficiently wealthy and poor households are illustrated in Figure 7. For parents with very small wealth W, investment decisions are exactly as in our simple model without any financial bequests, and 'non-investors' consume more than the 'investors'. The situation is very different, however, for sufficiently wealthy parents. Their parental consumption (conditional on wealth W) is strictly decreasing in x over $x \in [0, x^*]$, and constant thereafter. The 'non-investors' (those with $x > x^*$) now consume *less* than the 'investors', opposite to the pattern in the model without any financial bequests.

[7] A sufficient condition for this is $C^{BT}(W; w_1) > x^*$.

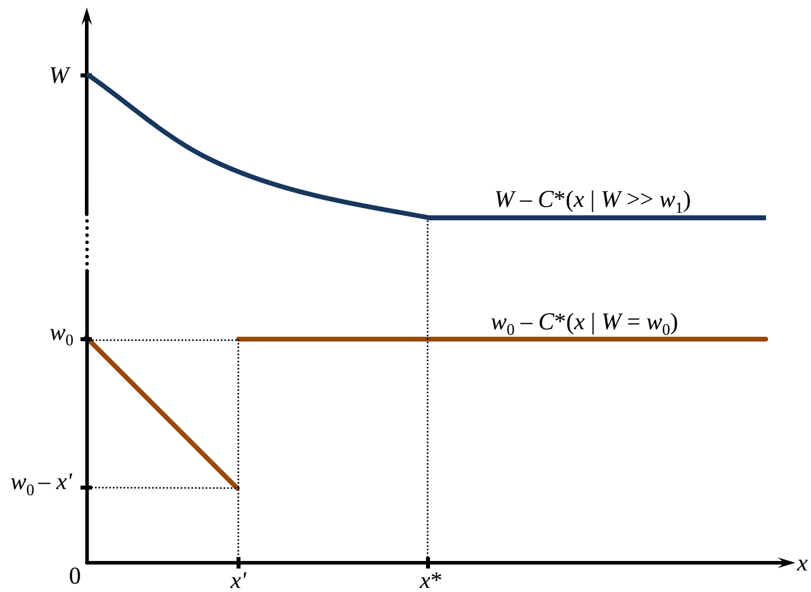

Figure 7: Consumption of sufficiently wealthy and poor parents

The argument that educational subsidies (financed by income or wealth taxes) lower consumption risk no longer applies to wealthy households in class A. They would instead raise risk. *So an opposite result holds here: an educational tax for parents with wealths falling in case A which funded a wealth subsidy (or income tax break) on the same set of households would reduce risk.* Starting with laissez faire, such a policy would be Pareto improving. It would, however, have opposite macroeconomic effects, as educational investments among such parents would fall. The resulting decline in skilled agents implies that the result about superiority of conditional transfers may not apply if the status quo policy is progressive, as this would worsen the government's fiscal balance.

On the other hand, our previous arguments would continue to apply for poor households in case B, who never make any financial bequests, and behave exactly as described in previous sections. For such poor housholds, therefore, our previous results remain unchanged: educational subsidies funded by income taxes would be Pareto improving as well as generate macro improvements.

For other classes of households, whether parents make financial bequests typically depends on the child's ability: they are made when the child is of sufficiently high ability, as well as when ability is low. For intermediate abilities, they make no financial bequests and make educational investments alone. The comparison of consumptions across 'investors' and 'non-investors' can go either way depending on the child's ability.

This suggests that arguments for educational subsidies should be limited to household wealth classes which make little or no financial bequests. The exact range of such households is an empirical matter. In the model of Abbott *et al.* (2013) calibrated to fit the NLSY 1997 data, all parents in the bottom quartile of the wealth distribution make inter-vivos transfers (inclusive of imputed value of rent when children lived with parents) to their children (when the latter were between ages of 16–22) which were smaller than what the latter spent on educational tuitions. The same was true for most of the second quartile as well. On the other hand, many parents in the top quartile transferred more than education tuition costs, and this happened to be true for *all* parents in the top 5%. This suggests case A applies to the top 5% of the US population, while case B applies to the bottom third of the population.

Indeed, our results suggest that it may be optimal for the government to use mixed policies of the following form: educational taxes for the population in case A, and subsidies for those in case B. The effects on educational investments in these two classes could then offset each other, leaving aggregate education investments unaltered. The composition of the educated would however change: since marginal children in case B are likely to be of higher ability than those in case A, there would be a rise in the average returns to education which would augment the efficiency benefits from the risk effects. We conjecture it is generally possible to construct such mixed policies which Pareto dominate a policy consisting of income-based transfers alone.

5 Relation to Literature

Loury (1981) provided a pioneering analysis of human capital investments by altruistic parents in an environment with ability shocks and no financial markets. Most of his analysis concerned the characterization of dynamic properties of competitive equilibria. He showed that redistributive policies could raise aggregate output and welfare, but did not explore the efficiency properties of laissez faire equilibria.

Bénabou (2002) also considered a model with human capital investment alone. An important difference from our model is that there are ability and productivity shocks, which are realized after investments are made. His model is more general by incorporating endogenous labor supply, while he restricts attention to specific functional forms for utility and production functions and specific distributions for shocks. He also restricts attention to specific forms of progressive taxes on income accompanied by linear consumption taxes and education subsidies.[8] His paper provides no analytical results concerning Pareto efficiency or inefficiency of laissez faire outcomes. Proposition 4 in his paper shows the Pareto criterion can be used to restrict attention to combinations of progressive income taxes, linear consumption taxes and linear education subsidies which support the laissez faire investment policy. This enables him to reduce the dimensionality of the fiscal policy to a single parameter of income tax progressivity. Thereafter his paper evaluates the comparative welfare (an aggregate efficiency measure) and macro effects of varying the level of progressivity. This corresponds to inter-occupation redistribution issues that we abstract from. Since varying progressivity does not seem to generate Pareto improvements, his results suggest that laissez faire outcomes are constrained efficient in his model. The underlying reason for this difference from our results is that ability shocks are not realized before parents make investment decisions in his model. This removes the essential heterogeneity in our model which generates fluctuations in parental consumption according to the realization of ability risk.

Aiyagari, Greenwood and Sheshadri (2002) study a model of human capital investment where education entails fixed and variable resource costs, besides child care. Education takes the form of increasing efficiency units of homogenous labor acquired by the child, as a function of the child's ability realization, parental resource and child care expenses. Apart from incorporating child care, the model is more general than ours by incorporating physical capital and financial bequests. The main focus of their paper is to characterize first-best

[8]He later examines progressive educational subsidies.

Pareto efficient allocations which can be decentralized with complete markets, and contrast these to allocations that result when there are no credit or insurance markets. They do not consider the effects of fiscal policy.

Our model is related to the occupational choice literature (Banerjee and Newman (1993), Galor and Zeira (1993), Ljungqvist (1993), Freeman (1996), Aghion and Bolton (1997), Bandopadhyay (1997), Lloyd-Ellis and Bernhardt (2000), Matsuyama (2000, 2006), Ghatak and Jiang (2002), Mookherjee and Ray (2002, 2003, 2008, 2010), Mookherjee and Napel (2007)), but this literature does not examine efficiency properties of competitive equilibria. The only exception is Mookherjee and Ray (2003) who abstract entirely from ability heterogeneity. Their model has steady state allocations that are constrained efficient. Hence ability heterogeneity has a dramatic effect on the efficiency properties of equilibria. Mookherjee and Napel (2007) study a closely related model with ability heterogeneity but do not analyze welfare properties of equilibria.

D'Amato and Mookherjee (2013) study efficiency properties of equilibria in a closely related model with ability heterogeneity, where the labor market is additionally characterized by signaling (i.e., productivity depends on ability in addition to education). They examine effects of educational loan programs provided by the government, funded by bonds released to the public. They obtain a result similar to our first result, viz. competitive equilibria are Pareto dominated by such a loan program. This intervention works differently by changing the composition of the educated in favor of children from low-income families who have higher abilities than children from high income families. Per capita education and output in the economy are unchanged. In our paper we abstract from policies (such as loans) which condition transfers on past education decisions, and more generally from efficiency effects operating through the composition rather than quantum of investments.

A number of papers (Abbott *et al.* (2013), Heathcote (2005), Cespedes (2011)) calibrate detailed macro models to US or Mexican data to study the effects of fiscal policy. These models involve investments in human and/or physical capital. There is ability heterogeneity and income risk, while credit and insurance markets are missing. Their models are more detailed by incorporating labor supply and savings decisions, different stages of life cycle of households, and various stages of education. Numerical simulations show welfare and macro effects of fiscal policies. Consistent with our results, Abbott *et al.* (2013) and Cespedes (2011) find positive aggregate welfare and macro effects of educational subsidy programs. They also find significant long run general equilibrium effects which attenuate the short-run effects, while rendering some groups of the population worse off (owing to changes in wage rates and interest rates).

6 Concluding Observations

We have provided theoretical arguments for Pareto-superiority of fiscal policies involving educational subsidies funded by income taxes imposed on the same income/occupational class. These dominate laissez faire outcomes, as well as policies where transfers are not conditioned on education decisions. The results apply quite generally, irrespective of specific assumptions on preferences or technology, provided parents do not supplement education investments with financial bequests. In the presence of financial bequests, laissez faire outcomes would continue to be Pareto dominated by similar policies applied only to poor households that do not leave financial bequests. For wealthy household classes that always leave financial bequests, Pareto optimality requires an opposite policy involving educational taxes which fund income transfers within the same class.

The main contribution of the paper is to provide qualitative results that depend relatively little on detailed assumptions concerning preferences or technology, or on the nature of social preferences for redistribution. They provide suggestions for policies based only on the Pareto criterion, that would generate no distributional conflict. And they help provide insights into the source of estimated welfare effects of educational subsidy policies in calibrated macro models.

Nevertheless, the model abstracted from endogenous labor supply, income risk, and intra-family correlations in ability draws. The design of Pareto optimal policies in the presence of financial bequests also needs to be explored further. These are left for future research.

Appendix

Proof of Lemma 1: (i) follows from the fact that $w_{1t} > w_{0t}$ if and only if $\lambda_t < \bar{\lambda}$, and $V_t(1,x) > V_t(0,x)$ for any x if and only if $w_{1t} > w_{0t}$. If (ii) is false and $\lambda_t \geq \bar{\lambda}$ at some date, we have $V_t(1,x) = V_t(0,x)$ for all x, implying that no parent with a child with $x > 0$ will want to invest in education at $t-1$, so $\lambda_t = 0 < \bar{\lambda}$, a contradiction. For (iii) note that (3) follows straightforwardly from the optimization problem faced by parents. And $x_{0t} < x_{1t}$ follows from (ii) above. To show (iv), suppose that it is not true and we can find a subsequence $t_n, n = 1, 2, ..$ along which x_{c,t_n} for some occupation c either tends to 0 or ∞. In the former case, (3) implies $[W_{1,t_n+1} - W_{0,t_n+1}]$ must converge to 0, which in turn requires λ_{t_n+1} to converge to $\bar{\lambda}$. Then x_{d,t_n} for both occupations $d = 0, 1$ must tend to 0, and (2) implies λ_{t_n+1} converges to 0, a contradiction. In the latter case $[W_{1,t_n+1} - W_{0,t_n+1}]$ must converge to ∞, implying x_{d,t_n} for both occupations $d = 0, 1$ must tend to ∞ by virtue of (3). (2) then implies λ_{t_n} approaches 1. This contradicts (ii) above. Since $\lambda_t \geq F(x_{0t})$ (owing to (2) and $x_{1t} > x_{0t}$), it follows that λ_t is uniformly bounded away from 0. Moreover, the argument used in (iv) (for the former case) also ensures λ_t is bounded away from $\bar{\lambda}$. The bounds on consumption follows from the bounds on wages and on investment thresholds. ■

Proof of Theorem 2

A useful preliminary result shows that any government budget surplus can be disposed of in an ex post Pareto improving manner while leaving investment incentives unchanged.

Lemma 2 *Given any sequence of positive budgetary surpluses $\{R_t\}_{t=0,1,...}$ resulting from a fiscal policy $\{\tau_{ct}, e_{ct}\}_{c;t}$ and an associated DCE $\{\lambda_t, w_{ct}, x_{ct}, \mathcal{U}_{ct}\}_{c;\,t}$, there exists another fiscal policy $\{\tau'_{ct}, e'_{ct}\}_{c;\,t}$ with $\tau'_{ct} > \tau_{ct}, e'_{ct} > e_{ct}$ for all $c = 0, 1$ and $t = 0, 1, \ldots$ with an associated DCE with the same skill ratios, wages and thresholds $\{\lambda_t, w_{ct}, x_{ct}\}_{c;\,t}$ which ex post Pareto dominates the original DCE (i.e., with $\mathcal{U}'_{ct} > \mathcal{U}_{ct}$ for all c, t).*

Proof of Lemma 2: Let the original DCE involve wages $\{w_{ct}\}_{t=0,1,2,\ldots}$ and investment thresholds $\{x_{ct}\}_{t=0,1,2,\ldots}$ in occupation c. For any period t and positive budgetary amount $R_{ct} \leq R_t$ to be disposed of to households in occupation c in t, select $\Delta\tau_{ct}(R_{ct}) \geq 0, \Delta e_{ct}(R_{ct}) \geq 0$ as defined by the unique solution to:

$$\begin{aligned} u(w_{ct}+\tau_{ct}) - u(w_{ct}+\tau_{ct}+e_{ct}-x_{ct}) &= u(w_{ct}+\tau_{ct}+\Delta\tau_{ct}) - u(w_{ct}+\tau_{ct}+\Delta\tau_{ct}+e_{ct}+\Delta e_{ct}-x_{ct}) \\ R_{ct} &= \alpha_{ct}[\Delta\tau_{ct}+F(x_{ct})\Delta e_{ct}] \end{aligned} \tag{12}$$

where α_{ct} equals λ_t if $c = 1$ and $1 - \lambda_t$ otherwise. This results in a change in interim consumption utility of a household in occupation c in period t by

$$\begin{aligned} \Delta\mathcal{U}_{ct}(R_{ct}) &= \big[u(w_{ct}+\tau_{ct}+\Delta\tau_{ct}) - u(w_{ct}+\tau_{ct})\big](1-F(x_{ct})) \\ &\quad + \int_0^{x_{ct}} \big\{u(w_{ct}+\tau_{ct}+\Delta\tau_{ct}+e_{ct}+\Delta e_{ct}-x) - u(w_{ct}+\tau_{ct}+e_{ct}-x)\big\} dF(x) \end{aligned}$$

provided the investment threshold remains x_{ct}.

$\Delta\tau_{ct}(R_{ct}), \Delta e_{ct}(R_{ct})$ and $\Delta\mathcal{U}_{ct}(R_{ct})$ are continuous, strictly increasing functions, taking the value 0 at $R_{ct} = 0$. By the Intermediate Value Theorem, for any $R_t > 0$ there exist R_{0t} and R_{1t} such that $R_{0t} + R_{1t} = R_t$ and $\Delta\mathcal{U}_{1t}(R_{1t}) = \Delta\mathcal{U}_{0t}(R_{0t})$. This ensures that $\mathcal{U}_{1t} - \mathcal{U}_{0t}$ is unchanged.

Because the definition of $\Delta\tau_{ct}$ and Δe_{ct} in (12) keeps investment sacrifices constant for threshold types x_{1t}, x_{0t}, the same investment strategies remain optimal for households in period t if they expect an unchanged welfare difference $W_{1,t+1} - W_{0,t+1}$. The sequence $\{W_{1t} - W_{0t}\}_{t=0,1,2,\ldots}$ remains unchanged given that there is no change to the sequence of consumption utility differences $\{\mathcal{U}_{1t} - \mathcal{U}_{0t}\}_{t=0,1,2,\ldots}$. The policy is constructed precisely to assure this, where preservation of the original investment thresholds also preserves skill ratios $\{\lambda_t\}_{t=1,2,\ldots}$ and associated pre-tax wages $\{w_{1t}, w_{0t}\}_{t=1,2,\ldots}$. The government budget is then balanced, while transfers to all households have increased. ■

The proof of Theorem 2 itself proceeds in five steps. Here we prove the more general version stated in the text following the statement of the result, where conditions (a)–(d) are satisfied by the status quo.

Step 1: Conditions (a)–(d) imply that the status quo fiscal policy and DCE satisfy the following properties:

(i) there exists $\bar{\lambda}_t \in (0,1)$ such that $\lambda_t \in (0, \bar{\lambda}_t)$ for all t, and λ_t is uniformly bounded away from 0;

(ii) x_{ct} is uniformly bounded above, and uniformly bounded away from zero;

(iii) consumptions of all agents are uniformly bounded.

To see this, define $\bar{\lambda}_t$ by the property that $g_1(\bar{\lambda}_t) - g_0(\bar{\lambda}_t) = \tau_{1t} - \tau_{0t}$. By virtue of condition (b), $\bar{\lambda}_t \in (0,1)$ for all t, and is uniformly bounded away from 0. Note that $\lambda_t \geq \bar{\lambda}_t$ implies that after-tax wages are equalized across the two occupations, since skilled agents can always work in the unskilled occupation. Then $W_{1,t} = W_{0,t}$, implying that parents at $t-1$ have no incentive to educate their children. This implies $\lambda_t = 0$, a contradiction. Hence equilibrium always involves $\lambda_t \in (0, \bar{\lambda}_t)$ for all t. Since $[g_1(\lambda_t) + \tau_{1t}] - [g_0(\lambda_t) + \tau_{0t}] > \kappa > 0$ for all t, $[W_{1t} - W_{0t}]$ is bounded away from 0 for all t. Since the distribution of x has full support over $(0, \infty)$, the proportion of parents at $t-1$ investing in their children's education is bounded away from zero in each occupation. Hence λ_t is uniformly bounded away from 0, which establishes (i). It also follows that x_{ct} is uniformly bounded above, which combined with condition (d) implies property (ii). Finally, (iii) follows from (ii) combined with condition (c).

Step 2: For arbitrary $\epsilon > 0$, construct the following policy change. Denote the status quo DCE by an * superscript. Fix an occupation c^*, and leave the fiscal policy $\{\tau_{dt}, e_{dt}\}_{t=0,1,2,\ldots}$ for occupation $d \neq c^*$ unchanged. For $c = c^*$, take $\tau'_{ct}(\epsilon) = \tau_{ct} - \epsilon F(x^*_{ct}), e'_{ct}(\epsilon) = e_{ct} + \epsilon(1 - \mu_t)$ where

$$\mu_t \equiv (1 - F(x^*_{ct}))\Big[1 - \frac{F(x^*_{ct})u'(w^*_c + \tau_{ct})}{\int_0^{x^*_{ct}} u'(w^*_c + \tau_{ct} + e_{ct} - x)dF(x)}\Big]. \tag{13}$$

It is evident that $\mu_t \in (0,1)$ for all t. By Step 1 and the concavity of u, it is uniformly bounded away from 0 and 1 respectively.

In what follows we use c to denote c^*, and d the other occupation. For arbitrary thresholds x_{ct}, x_{dt} define

$$\mathcal{U}_{ct}(x_{ct},\epsilon) \equiv u\big(w^*_{ct}+\tau_{ct}-\epsilon F(x^*_{ct})\big)[1-F(x_{ct})]+\int_0^{x_{ct}} u\big(w^*_{ct}+\tau_{ct}-\epsilon F(x^*_{ct})+e_{ct}+\epsilon(1-\mu_t)-x\big)dF(x) \tag{14}$$

$$\mathcal{U}_{dt}(x_{dt},\epsilon) \equiv u(w^*_{dt}+\tau_{dt})[1-F(x_{dt})]+\int_0^{x_{dt}} u(w^*_{dt}+\tau_{dt}+e_{dt}-x)dF(x) \tag{15}$$

By construction of μ_t we have

$$\frac{\partial \mathcal{U}_{ct}(x^*_{ct},0)}{\partial \epsilon} = 0 = \frac{\partial \mathcal{U}_{dt}(x^*_{dt},0)}{\partial \epsilon}. \tag{16}$$

Considering bounded but otherwise arbitrary sequences of investment thresholds $\{x_{ct}\}_{t=0,1,2,\ldots}$ and $\{x_{dt}\}_{t=0,1,2,\ldots}$, define

$$\Delta W_t(\epsilon) \equiv \sum_{k=0}^{\infty} \nu_k \cdot [\mathcal{U}_{1,t+k}(x_{1t},\epsilon) - \mathcal{U}_{0,t+k}(x_{0t},\epsilon)]. \tag{17}$$

The associated interim welfare of a parent in the unskilled occupation at t is

$$W_{0t}(\epsilon) = \sum_{k=0}^{\infty} \delta^k \mathcal{U}_{0t}(x_{0t},\epsilon) + \delta \sum_{k=0}^{\infty} \delta^k F(x_{0,t+k}) \Delta W_{t+1+k}(\epsilon) \tag{18}$$

and that of a parent in the skilled occupation at t is

$$W_{1t}(\epsilon) = W_{0t}(\epsilon) + \Delta W_t(\epsilon). \tag{19}$$

These are the correct expressions for the value functions corresponding to the specified investment thresholds, *assuming that the policy change leaves after-tax wages for each occupation unchanged at every* t. Throughout Step 2, we shall continue to assume this.

The series in (17)–(18) converge uniformly in a non-empty interval I around $\epsilon = 0$.[9] Moreover, for any given bounded sequences $\{x_{ct}\}_{t=0,1,2,\ldots}$ and $\{x_{dt}\}_{t=0,1,2,\ldots}$, the partial sums

$$\sum_{k=0}^{T} \delta^k \frac{\partial \mathcal{U}_{ot}(x_{ot},\epsilon)}{\partial \epsilon}$$

[9] Convergence is uniform because $\delta < 1$, $\nu_k < \delta^k$, and $|\mathcal{U}_{ot}(x_{ot},\epsilon)|$ is bounded uniformly on $I = [-\xi,\xi]$, $\xi > 0$, by, e.g., $\kappa_o = \max_{\epsilon \in I}\Big\{|u(g_1(\bar{\lambda})+\tau_{ot})|, |u(\tau_{ot}+e_{ot}+\epsilon-\sup_t\{x^*_{1t},x^*_{0t}\})|\Big\}$ for $o = 0, 1$.

converge uniformly on I because $\left|\frac{\partial \mathcal{U}_{ot}(x_{ot},\epsilon)}{\partial \epsilon}\right|$ is uniformly bounded on I given that u is C^2 on $\mathbb{R}$ and wages and thresholds are uniformly bounded. We can therefore exchange the order of summation and differentiation when evaluating the welfare effects of a small change of ϵ. Moreover, applying the Envelope Theorem at each t, we can neglect induced changes in the investment thresholds at $\epsilon = 0$. So one obtains

$$\frac{\partial \Delta W_t(0)}{\partial \epsilon} = \sum_{k=0}^{\infty} \nu_k \left[\frac{\partial \mathcal{U}_{1,t+k}(x_{1t}^*,0)}{\partial \epsilon} - \frac{\partial \mathcal{U}_{0,t+k}(x_{0t}^*,0)}{\partial \epsilon}\right] = 0 \tag{20}$$

from (16); and combining (18)–(20) allows us to conclude that

$$\frac{\partial W_{0t}(0)}{\partial \epsilon} = 0 = \frac{\partial W_{1t}(0)}{\partial \epsilon}. \tag{21}$$

Since u is C^2 on $\mathbb{R}$, the second derivatives of $\mathcal{U}_{ot}(x_{ot},\epsilon)$ are also uniformly bounded on I. Hence we can also exchange the order of summation and differentiation when considering the derivative of $\partial \Delta W_t(\epsilon)/\partial \epsilon$, and conclude that $\Delta W_t(\epsilon)$ is C^1 because $\partial^2 \Delta W_t(\epsilon)/\partial \epsilon^2$ exists.

Optimal investment thresholds $x_{ct}(\epsilon)$ and $x_{dt}(\epsilon)$ are determined by

$$u\big(w_{ct}^* + \tau_{ct} - \epsilon F(x_{ct}^*)\big) - u\big(w_{ct}^* + \tau_{ct} + e_{ct} + \epsilon(1-\mu_t) - \epsilon F(x_{ct}^*) - x_{ct}(\epsilon)\big) = \delta \cdot \Delta W_{t+1}(\epsilon) \tag{22}$$

and

$$u(w_{dt}^* + \tau_{dt}) - u\big(w_{dt}^* + \tau_{dt} + e_{dt} - x_{dt}(\epsilon)\big) = \delta \cdot \Delta W_{t+1}(\epsilon). \tag{23}$$

Since all involved terms are C^1 functions of ϵ, we can conclude from the Implicit Function Theorem that $x_{ct}(\epsilon)$ and $x_{dt}(\epsilon)$ are C^1 on an interval I around $\epsilon = 0$.

As we vary ϵ from 0, the threshold x_{ct} undergoes a first-order increase, while the first-order change in x_{dt} is zero. Namely, the derivatives of the right-hand sides of (22) and (23) w.r.t. ϵ at 0 are zero, given that (20) holds for every t. Differentiating (23) directly implies $\frac{\partial x_{dt}(0)}{\partial \epsilon} = 0$. Differentiating (22) yields

$$\frac{\partial x_{ct}(0)}{\partial \epsilon} = F(x_{ct}^*) \frac{u'(w_{ct}^* + \tau_{ct})}{u'(w_{ct}^* + \tau_{ct} + e_{ct} - x_{ct}^*)} + (1 - \mu_t - F(x_{ct}^*)) \tag{24}$$

and the concavity of u implies

$$\int_0^{x_{ct}^*} u'(w_{ct}^* + \tau_{ct} + e_{ct} - x)dF(x) < F(x_{ct}^*)u'(w_{ct}^* + \tau_{ct} + e_{ct} - x_{ct}^*). \tag{25}$$

Hence

$$\mu_t < (1 - F(x_{ct}^*))\Big[1 - \frac{u'(w_{ct}^* + \tau_{ct})}{u'(w_{ct}^* + \tau_{ct} + e_{ct} - x_{ct}^*)}\Big], \tag{26}$$

and substituting this into (24) we obtain

$$\frac{\partial x_{ct}(0)}{\partial \epsilon} > \frac{u'(w_{ct}^* + \tau_{ct})}{u'(w_{ct}^* + \tau_{ct} + e_{ct} - x_{ct}^*)} > 0. \tag{27}$$

So assuming that the policy change leaves after-tax wages unchanged for each occupation, it produces a first-order increase in investment thresholds for parents in occupation c and a zero first-order effect on thresholds for parents in occupation d at every $t \geq 1$. Also it generates a zero first-order effect on dynastic utilities at every $t = 0, 1, 2, \ldots$ if after-tax wages are unchanged.

Step 3: In order to ensure that after-tax wages remain at their original levels, we introduce a wage neutralization policy at each t. First, for any $\epsilon \geq 0$ and $t \geq 0$, recursively define the skill ratio that would be induced in period $t+1$ by the investment thresholds $x_{ct}(\epsilon), x_{dt}(\epsilon)$

$$\lambda_{t+1}(\epsilon) = F(x_{1t}(\epsilon))\lambda_t(\epsilon) + F(x_{0t}(\epsilon))(1 - \lambda_t(\epsilon)) \tag{28}$$

with $\lambda_0(\epsilon) = \lambda_0$ given. Since $x_{1t}(\epsilon), x_{0t}(\epsilon)$ are C^1 functions of ϵ in a neighborhood of 0, so – by induction – must be $\lambda_{t+1}(\epsilon)$ for all t. Also note that (27) combined with $x_{1t}(\epsilon) > x_{0t}(\epsilon)$ at all t implies that $\lambda'_{t+1}(0)$ is positive and uniformly bounded away from 0.

Now switch to the following modified policy $\big(\tilde{\tau}_{ct}(\epsilon), \tilde{\tau}_{dt}(\epsilon), \tilde{e}_{ct}(\epsilon), \tilde{e}_{dt}(\epsilon)\big)$ for each period $t \geq 1$

$$\tilde{\tau}_{ct}(\epsilon) = w_{ct}^* - w_{ct}(\epsilon) + \tau'_{ct}(\epsilon) = w_{ct}^* - w_{ct}(\epsilon) + \tau_{ct} - \epsilon F_{ct}^* \tag{29}$$

$$\tilde{e}_{ct}(\epsilon) = e'_{ct}(\epsilon) = e_{ct} + \epsilon(1 - \mu_t) \tag{30}$$

$$\tilde{\tau}_{dt}(\epsilon) = w_{dt}^* - w_{dt}(\epsilon) + \tau'_{dt} = w_{dt}^* - w_{dt}(\epsilon) + \tau_{dt} \tag{31}$$

$$\tilde{e}_{dt}(\epsilon) = e'_{dt}(\epsilon) = e_{dt} \tag{32}$$

where $w_{ot}(\epsilon) = g_o(\lambda_t(\epsilon))$, $o = c, d$.

This modified policy induces a DCE with skill ratios $\{\lambda_t(\epsilon)\}_{t=1,2,\ldots}$, investment thresholds $\{x_{ct}(\epsilon), x_{dt}(\epsilon)\}_{t=0,1,2\ldots}$ and the interim utilities $\{\mathcal{U}_{ct}(x_{ct}(\epsilon), \epsilon), \mathcal{U}_{dt}(x_{dt}(\epsilon), \epsilon)\}_{t=0,1,2,\ldots}$

which were constructed in Step 2 under the assumption of unchanged after-tax wages in each occupation at each date. Given investment thresholds $x_{ct}(\epsilon), x_{dt}(\epsilon)$ the resulting skill ratio is $\lambda_{t+1}(\epsilon)$ and hence pre-tax wages are $g_c(\lambda_{t+1}(\epsilon)), g_d(\lambda_{t+1}(\epsilon))$. The transfers defined by (29)–(32), therefore, ensure that the household's optimization problem in each period corresponds to the one under original wages $\{w^*_{1t}, w^*_{0t}\}_{t=0,1,2,\ldots}$ and the policy $\{\tau'_c(\epsilon), \tau'_d(\epsilon), e'_c(\epsilon), e'_d(\epsilon)\}_{t=0,1,2,\ldots}$.

Step 4: We next check that there is a first-order improvement in government revenues at every t. Supposing that $c = 1, d = 0$ (an analogous argument works for the opposite case), the budget surplus for $t \geq 0$ is

$$\begin{aligned}
B_t(\epsilon) &= -\lambda_t(\epsilon)[\tilde{\tau}_{1t}(\epsilon) + F(x_{1t}(\epsilon))\tilde{e}_{1t}(\epsilon)] \\
&\quad -[1-\lambda_t(\epsilon)][\tilde{\tau}_{0t}(\epsilon) + F(x_{0t}(\epsilon))\tilde{e}_{0t}(\epsilon)] \\
&= -\lambda_t(\epsilon)[w^*_{1t} - w_{1t}(\epsilon) + \tau_{1t} - \epsilon F^*_{1t} + F_{1t}(\epsilon)\{e_{1t} + \epsilon(1-\mu_t)\}] \\
&\quad -[1-\lambda_t(\epsilon)][w^*_{0t} - w_{0t}(\epsilon) + \tau_{0t} + F_{0t}(\epsilon)e_{0t}] \\
&= -\lambda_t(\epsilon)F_{1t}(\epsilon)\epsilon\mu_t - \lambda_t(\epsilon)\epsilon[F_{1t}(\epsilon) - F^*_{1t}] \\
&\quad -\lambda_t(\epsilon)[w^*_{1t} - w_{1t}(\epsilon) + \tau_{1t} + F_{1t}(\epsilon)e_{1t}] \\
&\quad -[1-\lambda_t(\epsilon)][w^*_{0t} - w_{0t}(\epsilon) + \tau_{0t} + F_{0t}(\epsilon)e_{0t}]
\end{aligned}$$

where $F_{ct}(\epsilon) \equiv F(x_{ct}(\epsilon))$.

Hence, recalling that $\lambda_t g'_1(\lambda_t) + (1-\lambda_t)g'_0(\lambda_t) = 0$ for every λ_t owing to the CRS property of the production function, we obtain

$$\begin{aligned}
B'_t(0) &= F^*_{1t}\mu_t\lambda^*_t \\
&\quad -\lambda_t{}'(0)[\tau_{1t} + F^*_{1t}e_{1t}] + \lambda_t{}'(0)[\tau_{0t} + F^*_{0t}e_{0t}] \\
&\quad -\lambda^*_t f(x^*_{1t})x'_{1t}(0)e_{1t} \\
&\geq F^*_{1t}\mu_t\lambda^*_t - \lambda_t{}'(0)[(\tau_{1t} + F^*_{1t}e_{1t}) - (\tau_{0t} + F^*_{0t}e_{0t})] \\
&\geq F^*_{1t}\mu_t\lambda^*_t
\end{aligned}$$

where the first inequality uses the assumption that $e_{1t} \leq 0$ and the second that the original fiscal policy is progressive in each period t.[10]

From Step 1, $F_{1t}^{*}\mu_t\lambda_t^{*}$ and hence $B_t'(0)$ is bounded away from 0 uniformly. Hence a sufficiently small $\epsilon > 0$ will generate a positive budget surplus at every date.

Step 5: Finally, apply Lemma 2 in order to dispose of the resulting budget surplus in an interim Pareto-improving way.

[10] Since $\lambda_0'(0) = 0$, the second term in the penultimate line is zero rather than positive for $t = 0$: the skill ratio is unchanged and a budget surplus arises only from $\mu_0 > 0$. At all succeeding dates, the skill ratio rises and generates a higher budget improvement owing to the weak progressivity of the tax system.

References

Abbott B., Gallipoli G., Meghir C. and Violante G.L. (2013). "Education policy and intergenerational transfers in equilibrium," Discussion Paper No. 1887, Cowles Foundation.

Aghion P. and Bolton P. (1997). "A Theory of Trickle-Down Growth and Development." *Review of Economic Studies* 64(2), 151-172.

Aiyagari S.R. (1994), "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics* 109(3), 659–684.

Aiyagari S.R., Greenwood J. and Seshadri A. (2002), "Efficient Investment in Children," *Journal of Economic Theory* 102(2), 290–321.

Atkinson A. and Stiglitz J. (1976),"The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics* 6(1–2), 55–75.

Bandyopadhyay D. (1997). "Distribution of Human Capital and Economic Growth," Working Paper No. 173, Department of Economics, Auckland Business School.

Banerjee, A. and Newman, A. (1993). "Occupational Choice and the Process of Development." *Journal of Political Economy* 101(2), 274–298.

Barro R. and Becker G. (1989) "Fertility Choice in a Model of Economic Growth," *Econometrica* 57(2), 481–501.

Becker G. and Tomes, N. (1979), "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy* 87(6), 1153–89.

Bénabou R. (2002), "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70(2), 481–517.

Berriel T.C. and Zilberman E. (2011) "Targeting the Poor: A Macroeconomic Analysis of Cash Transfer Programs," Working Paper No. 726, Fundação Getulio Vargas/EPGE.

Cespedes N. (2011) "General Equilibrium Analysis of Conditional Cash Transfers," Working Paper, University of Rochester.

D'Amato M. and Mookherjee D. (2013), "Welfare Economics of Educational Policy When Signaling and Credit Constraints Co-exist," Working Paper, University of Salerno and Boston University.

Freeman, S. (1996), "Equilibrium Income Inequality among Identical Agents." *Journal of Political Economy* 104(5), 1047–1064.

Galor, O. and Zeira, J. (1993), "Income Distribution and Macroeconomics." *Review of Economic Studies* 60(1), 35–52.

Ghatak, M. and N. Jiang (2002), "A Simple Model of Inequality, Occupational Choice and Development" *Journal of Development Economics* 69(1), 205–226.

Heathcote J. (2005), "Fiscal Policy with Heterogeneous Agents and Incomplete Markets" *Review of Economic Studies* 72(1), 161-188.

Ljungqvist, L. (1993), "Economic Underdevelopment: The Case of Missing Market for Human Capital," *Journal of Development Economics* 40(2), 219–239.

Lloyd-Ellis, H. and Bernhardt D. (2000), "Enterprise, Inequality and Economic Development," *Review of Economic Studies* 67(1), 147–168.

Loury, G. (1981), "Intergenerational Transfers and the Distribution of Earnings." *Econometrica* 49(4), pp. 843–867.

Matsuyama, K. (2000), "Endogenous Inequality", *Review of Economic Studies* 67(4), 743–759.

Matsuyama, K. (2006), "The 2005 Lawrence Klein Lecture: Emergent Class Structure," *International Economic Review* 47(2), 327–260.

Mirrlees J. (1971), "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies* 38(2), 175–208.

Mookherjee D. and Napel S. (2007), "Intergenerational Mobility and Macroeconomic History Dependence," *Journal of Economic Theory* 137(1), 49–78

Mookherjee D. and Ray D. (2002), "Is Equality Stable?," *American Economic Review (Papers and Proceedings)* 92(2), 253–259.

Mookherjee D. and Ray D. (2003), "Persistent Inequality," *Review of Economic Studies* 70(2), 369–394.

Mookherjee D. and Ray D. (2008), "A Dynamic Incentive-Based Argument for Conditional Transfers," *Economic Record* 84(S1): S2–16.

Mookherjee D. and Ray D.(2010), "Inequality and Markets: Some Implications of Occupational Diversity," *American Economic Journal: Microeconomics* 2(4), 38-76.