Poverty Dynamics and Efficiency Effects of Conditional Cash Transfers*

Dilip Mookherjee
Institute for Economic Development
Boston University
264 Bay State Road
Boston, MA 02215, USA
dilipm@bu.edu

Stefan Napel
Dept. of Economics
University of Bayreuth
Universitätsstr. 30
95440 Bayreuth, Germany
stefan.napel@uni-bayreuth.de

Version: October 9, 2019

Abstract

We study effects of conditioning cash transfers (on child education) on human capital accumulation, income distribution dynamics and welfare, in an OLG model with missing financial markets and heterogeneity of learning ability. The model incorporates effects on consumption smoothing, financing of transfers, and dynamic general equilibrium effects both in the short and long run. The main result is that starting with any set of unconditional transfers, a conditional transfer system can be designed to reduce the proportion of unskilled agents, raise per capita income, and create a Pareto improvement in every generation, irrespective of initial conditions of the economy.

Keywords: human capital incentives, conditional cash transfers, universal basic income

*We acknowledge help and feedback from Brant Abbott, Roland Bénabou, Marcello d'Amato, Maitreesh Ghatak and Debraj Ray, besides participants in various seminars and conferences.

1 Introduction

Whether cash transfers should be conditioned on parental investments in child education and health is an important question in designing a welfare system. Conditional cash transfer (CCT) schemes have been in place for over two decades in many developing and middle income countries, with *Progresa* in Mexico and *Bolsa Familia* in Brazil being the best known. The contrasting idea of a Universal Basic Income (UBI) has also been actively debated in developed and developing countries alike: e.g., versions of UBI have been implemented in Alaska, introduced on an experimental basis in Finland, rejected in a referendum in Switzerland, and constitute the topic of recent policy debates in India.[1] The two key features of cash transfers in UBI are that they are *universal* and *unconditional*: everyone in a given country receives the same amount, irrespective of their circumstances or behavior. CCTs on the other hand tend to be targeted to poor households (identified both geographically and by asset ownership), and condition cash transfers to parents on school participation of their children.

In this paper, we focus on the question of the conditionality of the transfers from a normative standpoint. This is closely related to classic debates in public economics between cash and in-kind transfers, since CCTs effectively amount to subsidies for child education. Arguments advanced in recent debates regarding UBI or CCTs have involved familiar issues of efficiency, paternalism, incentives, fiscal and political feasibility (e.g., see Hanna and Olken (2018), Ghatak and Maniquet (2019)). Most of these pros and cons pertain to a static setting, where the distribution of underlying determinants of poverty such as human capital are taken as given. This ignores the potential role of the welfare system in perpetuating or eliminating poverty in the future. Santiago Levy, one of the principal designers of the *Progresa* program in Mexico describes dynamic poverty reduction as one of the main motivations of the program:

> "It was also important, finally, to avoid generating lasting dependence on income transfers. Experience from other countries had shown that making pure income transfers just because the recipients were poor could reduce their incentives to work and invest, inadvertently leading a subset of able and productive citizens to permanent dependence on public welfare. To avoid that outcome, income

[1] See e.g., van Parijs and Vanderborght (2017), Bidadanure (2019), Government of India 2017 Economic Survey (Chapter 9), *Indian Journal of Human Development* (2017) Symposium Issue on UBI, Ideas for India 2019 E-Symposium on UBI, and Ghatak and Muralidharan (2019).

transfers should be designed to be transitory investments in the human capital of the poor. They should take a life cycle approach, helping poor households in the more critical aspects of each stage of their lives but always with the view that they should have incentives to earn a sufficient level of income through their own efforts to eventually pull themselves out of poverty." (Levy (2006, p. 13))

Empirical evidence from CCTs in various developing countries indicate that they have generated significant positive effects on investments in education and health, with impacts on future labor market participation and earnings of beneficiaries (Parker and Todd (2017)). The evidence also shows these effects would have been smaller in the absence of the conditionalities. Hence the argument in favor of CCTs is strengthened if reductions in future poverty is an important policy goal.

This still leaves open many additional questions pertaining to a welfare comparison, such as utility consequences for different generations and households with varying realizations of ability. These depend on the specific way that the transfers are designed and financed. If human capital investments are affected, so would consumption of parents which matters for their welfare. Wages in different occupations in the future would also be affected, creating dynamic pecuniary externalities via fiscal and general equilibrium (GE) effects. There are relatively few papers that provide a complete welfare analysis. While a detailed review of related literature is provided in Section 5, the two most relevant papers are Fender and Wang (2003) and Mookherjee and Ray (2008), both of which provide normative justifications for CCTs. However, both papers focus on the 'long run' by concentrating exclusively on comparison of steady states, ignoring transitional dynamics that are relevant in the short to medium term. They also evaluate effects on social welfare using specific distributional weights, rather than Pareto comparisons. This leaves open the possibility that CCTs may leave some agents worse off, which may be normatively undesirable, besides creating possible problems of political feasibility.

In this paper we provide a stronger justification for conditional transfers, based on Pareto comparisons and incorporating short run as well as long run considerations. We study an overlapping generations model where parents with non-paternalistic (Barro-Becker) altruism towards their descendants make investment decisions in their children's human capital, (learning) ability of children is heterogenous, and credit and insurance markets are missing.[2] The government is assumed to observe only occupations and incomes of parents, besides investments in child education. Transfers can be conditioned on parental occupation/income (with marginal tax rates between 0 and 1); the key issue is whether they should additionally be conditioned on education investment. We show that *given arbitrary initial conditions, and* **any** *dynamic competitive equilibrium in an economy with unconditional transfers, there exists a corresponding system involving conditional transfers which would generate an (interim) Pareto improvement.* Utility is improved for every agent in every household in every occupation and every generation, and is assessed at the interim stage when they do not yet know the ability realization of their child. For high enough parental altruism, we show that an ex post Pareto improvement results, i.e., every parent is better off irrespective of the child's realized ability.

The generality of this result is striking. It applies irrespective of initial conditions of the economy, e.g., whether it is underdeveloped or developed, or whether it is a steady state or not. Even if we start at a steady state, we incorporate short and long run effects by assessing welfare of every successive generation, rather than any new steady state the economy may eventually converge to. Whether or not the economy eventually converges to a steady state also does not matter. The result also applies irrespective of any specific functional forms for utility or production functions, apart from standard monotonicity, concavity and smoothness properties. It shows that the Pareto comparison that obtains in a static setting with rational agents[3] is reversed in a dynamic setting: conditional transfers Pareto dominate unconditional transfers rather than the other way around.

The key idea is illustrated as follows. In the simplest case, education is a zero-one decision, and there are two occupations, skilled and unskilled, where education is necessary to enter the skilled occupation. To avoid redistributive effects across occupations, we can construct education subsidies for each occupation in the initial generation (generation 0) that are financed by a supplementary income tax on the same occupation. In effect, this provides insurance against the 'risk' that one's child will have high learning ability and the parent will be motivated to invest in the child's education. Parents whose children have low ability end up not investing in their education, and hence obtain a lower net transfer

[2]Plausible reasons for the latter are that parents cannot borrow against their children's future earnings and are privately informed about ability, so adverse selection undermines provision of education insurance.

[3]For any given expenditure on transfers to the poor (implying a given cost incurred by the taxpaying non-poor that finance the transfers), the latter are better off if they are unconstrained in how to spend the money.

from the government. Since these parents have the same income as others in the same occupation but invest less in their child's education, they end up with higher consumption. The education subsidy therefore smooths parental consumption at the same time that it provides superior education incentives. This is an instance where insurance and incentives move in the same direction. This is the root cause of the Pareto improvement.

The next generation (generation 1) would be characterized by a higher fraction of skilled adults. This has two GE implications, one through the government's fiscal revenues, and the other through effects on wage rates. If the difference between skilled and unskilled marginal income tax rates is nonnegative, the effect on the fiscal balance is also nonnegative and hence benign (as budget surpluses can always be distributed back to agents in a Pareto improving manner). Skilled wages would be lower, implying that skilled agents in the next generation are worse off. Moreover, the lowered skill premium in wages in generation 1 would lower education incentives in generation 0. These market wage effects can however be 'sterilized' with a regressive variation in income taxes which ensure after-tax incomes remain unchanged for every occupation in generation 1.

Since there are more skilled adults in generation 1, and skilled parents (being richer than unskilled parents) are more likely to invest in their children, there would be further ripple effects into subsequent generations: there would be more skilled adults in generations $2, 3, \ldots$ But these can be dealt with in a similar manner as for generation 1. The eventual fiscal intervention overlays all these different interventions at different dates together, in a way that raises interim utilities of parents in every occupation-generation pair, while leaving after-tax wages unchanged. *Ex post*, parents who do not invest in their children's education end up cross-subsidizing the consumption of those who do. With sufficient altruism, the costs of this are outweighed by the gain in expected utilities of their descendants, also resulting in an *ex post* Pareto improvement.

We subsequently show the argument is robust to extensions of the model that incorporate endogenous labor supply, paternalistic altruism, and an arbitrary number of occupations. However, the results 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 the realization of children's ability risk – the educational subsidy policy described above would now impose additional consumption risk, and thereby create a welfare loss. The nature of a Pareto improving policy is reversed in this case: requiring education for the wealthy (as defined above and corresponding to the top 5% in the US wealth distribution according to the calibrations by Abbott, Gallipoli, Meghir and Violante (2013)) to be taxed, and these taxes to fund income transfers to the same class. 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. Hence Pareto efficiency still requires fiscal transfers to be conditioned on education investment, in a manner that may depend on financial bequests.

It is important to clarify that the argument does rest on the assumption of heterogeneity of learning ability. In the absence of such heterogeneity, the model collapses to the one studied in Mookherjee and Ray (2003), which showed there always exist steady states under laissez faire which are Pareto efficient. Intuitively, the heterogeneity of ability creates the scope for education subsidies to be entirely self-financing within any given occupation-generation pair, whereby parents of children with low ability cross-subsidize those with children with high ability.

To summarize, the distinguishing characteristic of this paper is a simple but important qualitative and robust result concerning the efficiency role of conditional transfers, when ability heterogeneity and financial market frictions coexist. This conceptual insight helps explain findings of detailed quantitative models of various real economies that educational subsidies raise aggregate welfare (e.g., Berriel and Zilberman (2011), Cespedes (2014), Peruffo and Ferreira (2017)). Our paper complements this literature by providing qualitative results on the superiority of combining income transfers with education subsidies, that apply irrespective of the specific welfare function, technology or preferences. In particular, they illustrate the benefits of conditional cash transfer (CCT) programs adopted in various developing countries.

We do not, however, address 'third-best' considerations (Ghatak and Maniquet (2019)) pertaining to the administration and enforcement of transfer conditionality. Governments have to verify school participation of children and deny transfers to parents if their children do not meet the required conditions. The widespread adoption of CCTs in many countries suggests this is not an overwhelming problem, though in some countries with poor state capacity it could pose an important barrier. In any case, our analysis helps identify the welfare benefits from transfer conditionality, which have to be traded off against the accompanying administration and enforcement costs. It is also worth mentioning that similar problems would arise in implementation of UBI in societies with low levels of financial inclusion, which create problems for direct transfers from the state to citizens outside the formal financial sector. Our results thus accord with the broad assessment of Ghatak and Maniquet (2019) that it is difficult to provide a convincing rationale for UBI in a second-best environment. Moreover, we show that this continues to hold when we incorporate dynamic GE effects.

The remainder of 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

In this section we present the simplest version of the model with two occupations, zero-one education decisions, inelastic labor supply, and no financial bequests. In later sections we discuss how the analysis would be altered as these assumptions are relaxed.

2.1 Laissez Faire

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

There are two occupations: unskilled and skilled (denoted 0 and 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 measure of the intensity of parental altruism, 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 needs to spend in order 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. 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.[4]

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 high or 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 owe to privacy of information amongst parents regarding the realization of their children's ability, and the adverse selection this would generate.

Household earnings are defined by occupational wages: $y_{it} = w_{0t} + I_{i,t-1} \cdot (w_{1t} - w_{0t})$, where w_{ct} denotes the wage in occupation c in generation t obtaining 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 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 strictly 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

[4]Findings would not change if we assumed that parents receive a noisy signal $\hat{x}_{it}$ of true education costs. Very mild conditions on the noise structure and risk attitudes would imply that uncertainty in ex post returns to education does not change the pattern in parental consumptions that is implied by ability heterogeneity. Same would apply if, e.g., the child stayed unskilled with a fixed probability between 0 and 1 despite a parental investment of x_{it}.

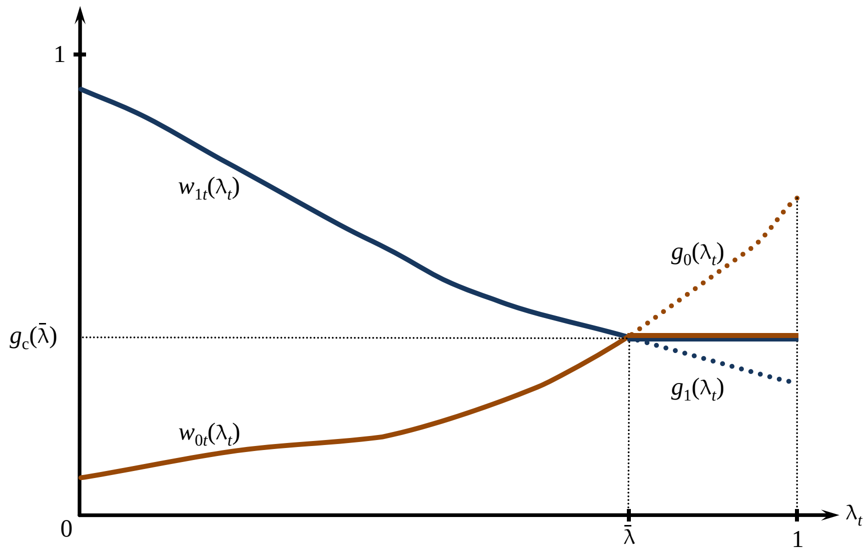

Figure 1: Labor Market

is bounded above even as its proportion in the economy becomes vanishingly small.[5]

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 $w_{1t} = g_1(\lambda_t)$ 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 that 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

[5]When the production function satisfies Inada conditions, i.e., marginal products are unbounded, 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 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 with an 'almost everywhere' proviso.

ratio in the economy in generation t.

2.2 Dynamic Competitive Equilibrium under Laissez Faire

Definition 1 *Given a skill ratio* $\lambda_0 \in (0, \bar{\lambda})$ *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) $\lambda_t < \bar{\lambda}$*,* $w_{1t} > w_{0t}$*.*

(iii) $I_{ct}(x) = 1$ *iff* $x < x_{ct}$*, where threshold* 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 agents 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 receive an education, while some untalented children born to skilled parents fail to receive an education. Finally, equilibrium consumptions and utility differences are bounded, which will be useful in our subsequent analysis.

2.3 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 in any generation t: income transfers τ_{1t}, τ_{0t} based on parental occupation, and transfers e_{1t}, e_{0t} based additionally on the parent's education investment decision. In particular, the government does not observe directly nor indirectly the ability realization of any given child.[6] This is 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 and schemes which condition on a family's transfer or decision history. Similar to private agents, the government will also not be able to lend or borrow across generations, and will hence have to balance its budget within each generation.

Definition 2 *Given a skill ratio* $\lambda_0 \in (0, \bar{\lambda})$ *in generation* 0*, a* dynamic competitive equilibrium (DCE) given fiscal policy $\{\tau_{0t}, \tau_{1t}, e_{0t}, e_{1t}\}_{t=0,1,2,...}$ *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*

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

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

(b)

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

[6] Indirect observability of children's abilities from the parental education expenses or test results would allow policy to realize efficiency gains from explicit improvements in the talent composition of investors. We think of education costs x as having a major unverifiable component, possibly also reflecting parental time that would be dedicated to a child's education and training in a more sophisticated model of labor supply.

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

The government has a balanced budget *if at every t it is the case that*

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

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

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

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 is then described by a sequence $\{\lambda_t, w_{1t}, w_{0t}, x_{1t}, x_{0t}, \mathcal{U}_{1t}, \mathcal{U}_{0t}\}_{t=0,1,2,\ldots}$ which satisfies equalities (5) and (7)–(9).

3 The Main Result for the Baseline Model

We now introduce our main result, pertaining to a comparison of transfers that do and do not condition on child education.

Theorem 1 *Consider any DCE given an initial skill ratio $\lambda_0 \in (0, \bar{\lambda})$ 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 with a higher

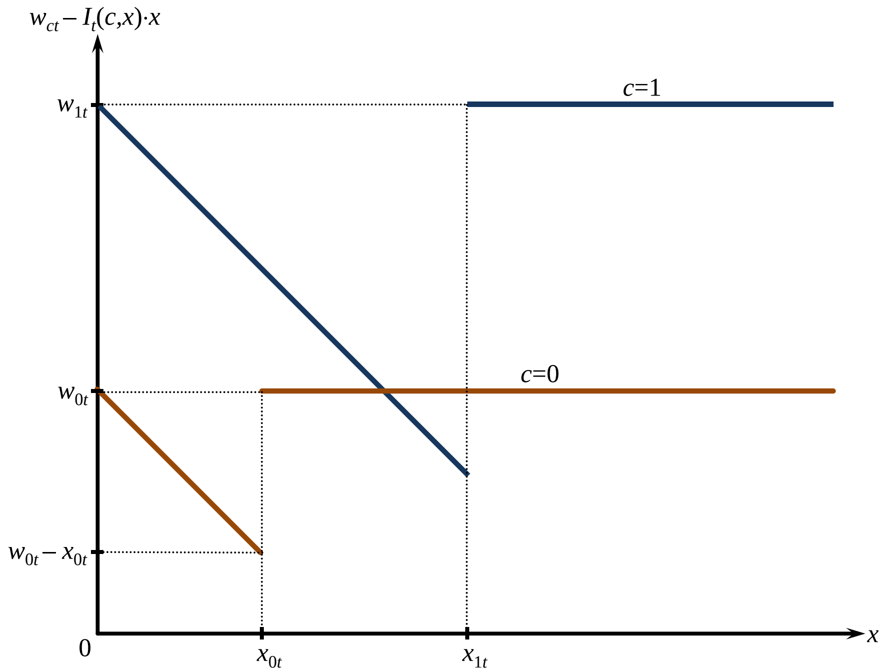

Figure 2: Variation of Parental Consumption with Education Cost

skill ratio λ_t and lower investment threshold x_{ct} for every c, t, which results in an interim Pareto improvement ($W_{ct} \equiv \mathsf{E}_x V_t(c, x)$ is higher for every c, t).

Condition (a) of Theorem 1 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 mild technical restriction ensuring that competitive equilibria always involve bounded consumptions and investment thresholds. The theorem states that given any policy of unconditional transfers meeting these conditions, it is possible to construct a corresponding policy involving conditional transfers which raises education investment and interim utility for every household in every generation. The new policy provides educational subsidies to a given occupation, which are funded by higher income taxes *levied on the same occupation*. In this sense, the policy does not redistribute across occupations. Instead, it redistributes across parents within each occupation class, between those that do and do not invest in their children's education.

The key observation is that with this mode of financing, the education subsidies simultaneously boost education incentives and provide insurance against the uncertain realizations of children's ability. The 'accident' in question is that one's child is born with enough talent that the parent invests in education, resulting in lower consumption than other parents in the same occupation who do not invest (i.e., whose children are not so talented). The resulting variation of parental consumption with the realization of their child's ability is illustrated in Figure 2. If the child is a genius and can be costlessly educated, the parent's consumption equals his earning. The same is true at the opposite extreme when the child has low enough ability that the parent is not motivated to invest in education. For intermediate abilities where the parent is motivated to provide an education, the parent invests a positive amount, lowering consumption. Hence parental consumption varies non-monotonically with respect to the cost necessary to educate the child.

The policy variation involves an increase $F(x_{ct})\epsilon > 0$ in (income) tax on all parents in a given occupation c in generation t, accompanied by an education subsidy of $\epsilon(1 - \mu_t)$. For $\mu_t > 0$, the total education subsidy paid to the fraction $F(x_{ct})$ of parents in occupation c who invest is smaller than the corresponding increase in tax revenue. This will raise the fiscal budget surplus. Fraction μ_t will be sufficiently small such that parents in occupation c receive a positive net transfer if they invest in their child's education. So those in occupation c who invest in education will receive a higher net transfer and thus consume more than before the variation, which will be funded by a fall in net transfer and consumption of those that do not invest. Figure 3 illustrates the construction of such a policy for unskilled parents ($c = 0$) in a given generation.

If μ_t were zero, average consumption would be unaffected but differences in consumption associated with heterogeneity of the children's education costs would be reduced (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 of the variation in parental consumption, thus raising the interim expected utility of consumption of unskilled adults of the given generation.

The parameter μ_t is set above zero so as to reduce the mean unskilled parental consumption enough that there is no change in the expected (contemporaneous) utility of these parents at date t. Assuming wages are unchanged, this implies that dynastic utilities of both occupations are unchanged. Hence investment incentives for generations prior to t are unchanged. The subsidization of education in the unskilled occupation on the other hand lowers the sacrifice these parents must endure to educate their children. Hence unskilled households in generation t will be motivated to invest more often (i.e., the threshold ability for their children will fall).

Aggregate investment in the economy will then rise, which will tend to lower skilled

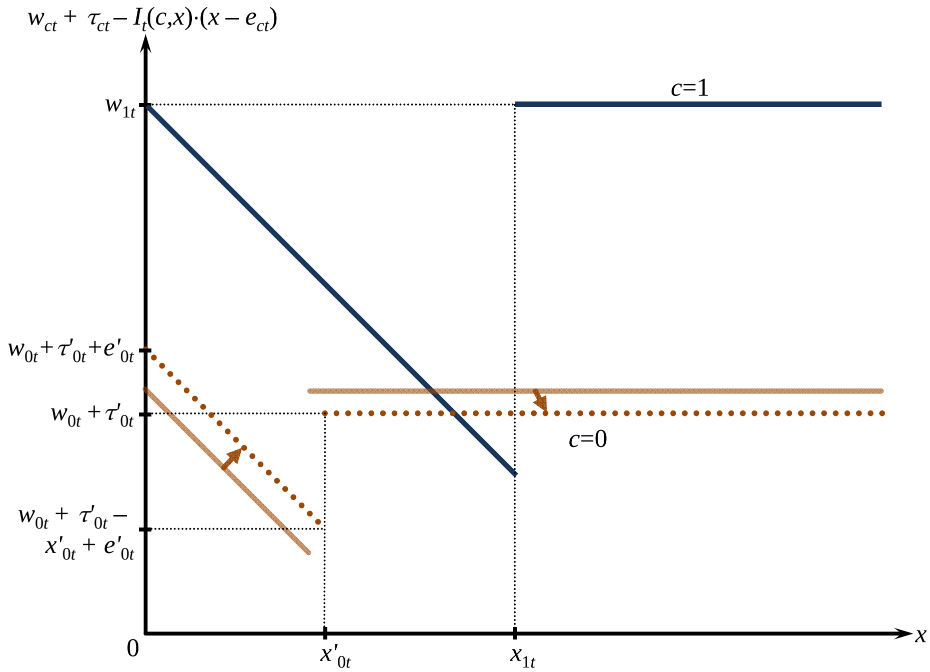

Figure 3: Effects of Steps 1 and 2 of Fiscal Policy Variation on Parental Consumption

wages and raise unskilled wages. These general equilibrium changes would reduce the benefits of investment and therefore fiscal policy is adjusted further to neutralize the wage changes. This results in a new competitive equilibrium 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 extraction of resources from households by setting $\mu_t > 0$. The progressivity of the original fiscal policy implies that the government budget surplus also improves as a result of the decline in the proportion of unskilled households.

In the last step of the argument the government constructs another variation in its tax-subsidy policy. It distributes the additional revenues so as to achieve a strict interim Pareto improvement, while preserving investment incentives. Note that by construction the dispersion in utility of consumption between occupations is unchanged, while a fraction of agents move up from the unskilled to the skilled occupation in every generation.

The consumption losses which the policy imposes on non-investing parents stay bounded while the dynastic gains which are created for all parents grow without bound as $\delta \to 1$. So with a sufficiently high degree of parental altruism, parameterized by δ, the policy-induced gains in expected utility of descendants outweigh any loss in own consumption relative to laissez faire. The constructed policy then achieves an ex post Pareto improvement. Formally, we can show:

Theorem 2 *Let a collection of economies with identical consumption utility function u, production function G and ability distribution F but different parental discount factors $\delta \in (0,1)$ be given. For each corresponding DCELF that starts from skill ratio $\lambda_0 \in (0,\bar{\lambda})$ at $t=0$, consider fiscal policies $\left\{\tilde{\tau}^{\delta}_{0t}(\epsilon), \tilde{\tau}^{\delta}_{1t}(\epsilon), \tilde{e}^{\delta}_{0t}(\epsilon), \tilde{e}^{\delta}_{1t}(\epsilon)\right\}_{t=0,1,2,\ldots}$ which induce an interim Pareto improvement according to Theorem 1. Then there exist $\underline{\delta} \in (0,1)$ and $\bar{\epsilon} > 0$ such that for any $\epsilon \in (0,\bar{\epsilon})$ and $\delta \in (\underline{\delta},1)$ the fiscal policy also ex post Pareto dominates the respective DCELF for all t.*

4 Extensions

We provide an informal discussion of how the preceding results would extend to variations of the model in different directions.

4.1 Endogenous Labor Supply

A first extension of the baseline model could be to consider households who choose how many hours of labor they supply, together with the binary decision whether to invest in education or not. That is, each household in occupation c selects $I_{ct}(x) \in \{0,1\}$ and $l_{ct}(x) \geq 0$ to maximize

$$u\big(l_{ct}\, w_{ct} - I_{ct}\, x\big) - d(l_{ct}) + \delta \mathsf{E}_{\tilde{x}} V_{t+1}(I_{ct}, \tilde{x}) \tag{10}$$

for strictly increasing and convex disutility of labor d. The optimal investment strategy $I_{ct}(x)$ in this case is of the same threshold form as in the baseline model. Namely, if we define

$$v(w_{ct}, x, I_{ct}) \equiv \max_{l_{ct}} \Big[u\big(l_{ct}\, w_{ct} - I_{ct}\, x\big) - d\big(l_{ct}\big)\Big] \tag{11}$$

then a parent in occupation c in period t who faces education cost x will invest iff $x < x_{ct}$, where threshold x_{ct} is defined by

$$v(w_{ct}, x_{ct}, 0) - v(w_{ct}, x_{ct}, 1) = \delta[W_{1,t+1} - W_{0,t+1}] \tag{12}$$

and $W_{ct} \equiv \mathsf{E}_{\tilde{x}} V_t(c, \tilde{x})$. Parents in occupation c with cost $x = 0$ or cost $x \geq x_{ct}$ have identical (indirect) utilities of consumption $v(w_{ct}, 0, 1) = v(w_{ct}, x, 0)$, while those with cost

$x \in (0, x_{ct})$ consume less. In particular, from (11) and the envelope theorem, we have

$$\frac{\partial v(w_{ct}, x, I_{ct}(x))}{\partial x} = -u'\big(l_{ct}(x)w_{ct} - x\big) < 0 \text{ for each } x \in (0, x_{ct}). \tag{13}$$

It follows that consumption utilities $v(w_{ct}, x, I_{ct}(x))$ are decreasing on $[0, x_{ct})$, jump back to $v(w_{ct}, 0, 1)$, and then stay at this level. That is, they exhibit a non-monotonic pattern with respect to education cost x just like in the baseline model (cf. Figure 2). A variation of the baseline policy intervention can therefore be applied in order to raise interim utility.

4.2 Paternalistic Altruism

Next suppose parents do not have Barro-Becker dynastic preferences. Instead, they 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) – perhaps incorporating parental concern for their own old age security. 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 $u\big(w_{ct} - I\,x\big) + I\,Y(w_{1,t+1}) + (1-I)\,Y(w_{0,t+1})\big)$. Theorem 1 extends 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 the variation of parental consumption is lowered. So investment incentives continue to rise, while enhancing interim expected utilities.

4.3 Continuous Investment Choices

What if educational investments can be varied continuously, rather than being indivisible? Our results extend straightforwardly to this context, too, 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 remuneration 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. The latter varies according to a continuous 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

$$V(e|x) \equiv \max_{0 \leq e' \leq \bar{e}} \left[u(w(e) - I(e';x)) + \delta W(e')\right] \tag{14}$$

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.[7] 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 $\frac{\partial I(e';x)}{\partial e'}\Big|_{e'=0} = \frac{\delta W'(0)}{u'(w(e))}$. Then the optimal policy function takes the form $e'(e;x) = 0$ if $x \geq x^*(e)$ and positive otherwise.[8] 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 subsidies funded by income taxes on all parents with the same education will reduce the riskiness of parental consumption, and thereby permit a Pareto improvement.

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

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

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

arise from the 'ability lottery' of their children, the policy increases interim expected utilities of each occupation. The preceding analytical details were needed to ensure that there is a positive mass of investors and non-investors respectively, so as to allow a strict Pareto improvement.

4.4 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 leave 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 here 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 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

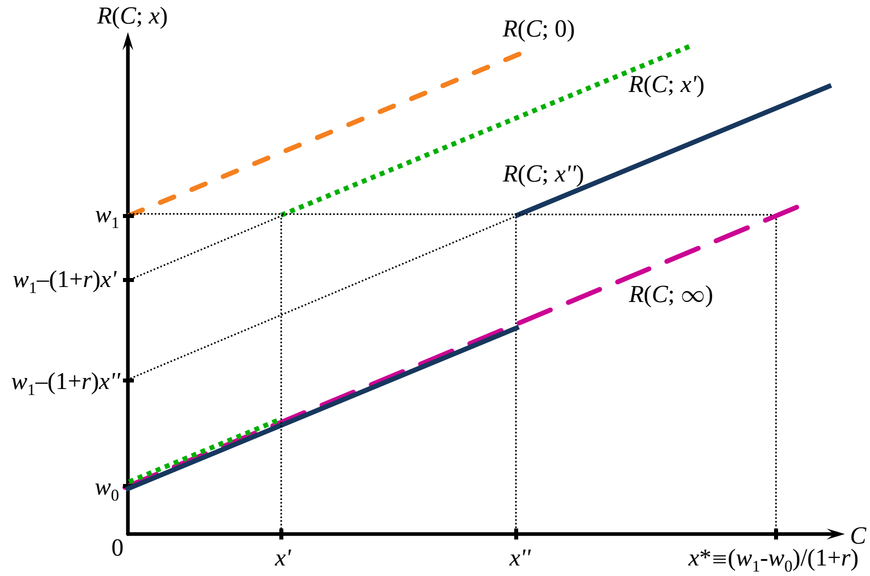

Figure 4: Child wealth as function of total investment expenditure C, given cost x

$b = C - x$. Then the child ends up with wealth $W' \equiv R(C; x)$ given by

$$R(C;x) = \begin{cases} (1+r)C + w_0 & \text{if } C < x, \text{ or } C \geq x \text{ and } \frac{w_1 - w_0}{x} \leq 1 + r, \\ (1+r)C + w_1 - (1+r)x & \text{if } C > x \text{ and } \frac{w_1 - w_0}{x} > 1 + r. \end{cases} \tag{15}$$

It is illustrated in Figure 4.

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

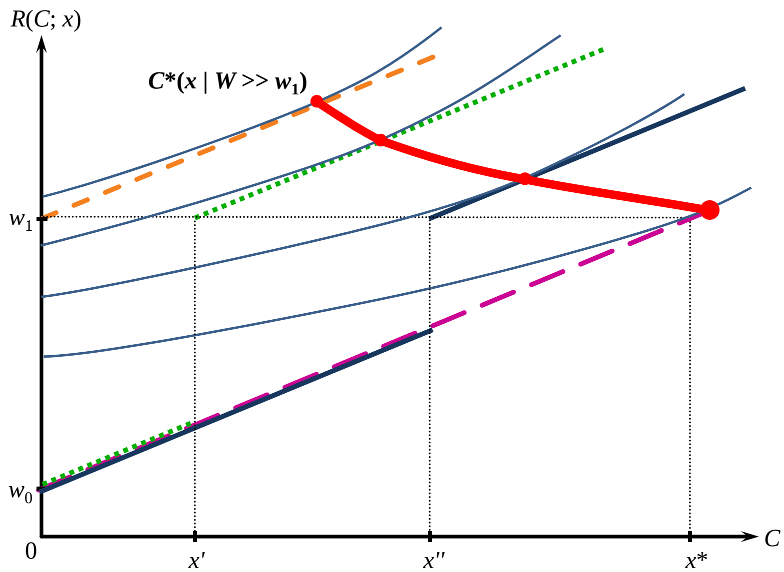

Figure 5: Investment expenditures of sufficiently wealthy parents (case A)

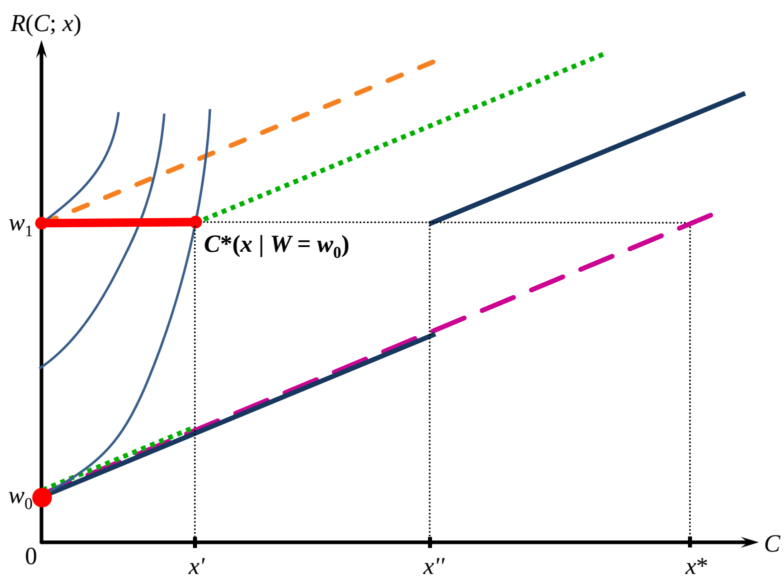

Figure 6: Investment expenditures of poor parents (case B)

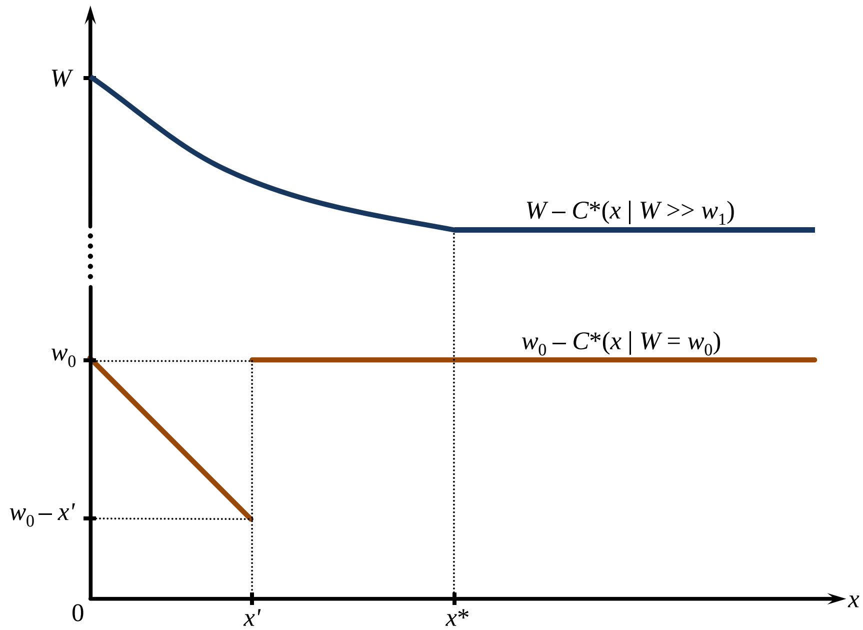

Figure 7: Consumption of sufficiently wealthy and poor parents

Case A. *W sufficiently large:* Suppose W is large enough that $C^{BT}(W; w_1 - (1+r)x) > x$ for all $x \leq x^*$.[9] 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 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.

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

The argument that educational subsidies (financed by income or wealth taxes) lower consumption risk no longer applies to wealthy households falling under case 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 households, therefore, our previous results remain unchanged: educational subsidies funded by income taxes would be Pareto improving.

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, Gallipoli, Meghir and Violante (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.

The additional heterogeneity among parents that comes with different accumulated wealth (compared to the baseline in which only the occupation varies) diminishes the target group for positive transfers that condition on educational investment but leaves the key mechanisms for making everyone better off untouched. The same would remain true if we added ex post heterogeneity resulting from other sources to the model, such as earnings uncertainty or i.i.d. wealth shocks. In fact, as long as part of total dynastic capital – human, financial, or physical – is indivisible and fully depreciates in one period, extensions along the indicated lines would also permit an entrepreneurial reinterpretation of the model: the dynasties might be risk-averse small businesses that decide on short-term indivisible investments, like planting a high-yield crop, with idiosyncratic variations in returns. The analogous intervention would then consist of investment subsidies financed by a tax on all businesses of comparable size.

5 Related Literature

Our paper is related to literatures in development and occupational choice, public economics and macroeconomics. We discuss these in turn.

5.1 Development and Occupational Choice

The closest connection is with the literature on occupational choice with credit market imperfections.[10] With few exceptions, this literature focuses on poverty dynamics under laissez faire, rather than normative properties of laissez faire or effects of fiscal policy. Mookherjee and Ray (2003) study a model which is a special case of the one we consider here, which abstracts from ability heterogeneity and fiscal policy interventions. In this framework Mookherjee and Ray (2008) compare properties (such as per capita output and social welfare corresponding to differing degrees of inequality aversion) of (suitably selected) steady states resulting from conditional and unconditional transfers. As explained in the

[10]See, e.g., Banerjee and Newman (1993), Galor and Zeira (1993), Ljungqvist (1993), Freeman (1996), Aghion and Bolton (1997), Bandopadhyay (1997), Maoz and Moav (1999), Lloyd-Ellis and Bernhardt (2000), Matsuyama (2000, 2006), Ghatak and Jiang (2002), Fender and Wang (2003), Mookherjee and Ray (2002, 2003, 2008, 2010), and Mookherjee and Napel (2007).

Introduction, their analysis is subject to a number of problems which we overcome in the current paper: they ignore short run effects, ability heterogeneity and (Pareto) efficiency comparisons.[11]

Fender and Wang (2003) incorporate ability heterogeneity in what is, essentially, a two-period model of occupational choice with credit rationing arising owing to moral hazard. Their model is relevant to higher education by young adults rather than education of children: there is no parental altruism; agents finance their own education and consumption utility is linear. They evaluate effects of public provision of education according to different methods of financing. Interventions that improve utilitarian welfare are shown to generally exist, but the tax burdens on those who remain uneducated make part of the population worse off. An exception arises when additional education investments prompt interest rates to increase so much (assuming there is no access to world capital markets) that this could dominate the direct effects for some parameter values. By contrast our model focuses on investments in children by their parents, incorporates consumption smoothing preferences, transition dynamics and identifies a general and robust source of Pareto improvements resulting from CCTs.

Finally, D'Amato and Mookherjee (2013) investigate the efficiency role of a different policy instrument: public provision of education loans, rather than CCTs. They focus on a two-skill OLG model with paternalistic altruism, ability heterogeneity and missing financial markets. Similar to this paper, they show that ex post Pareto improving education interventions exist. They additionally show the result is robust when education signals unobserved productivity of workers to employers.

5.2 Public Economics

Sinn (1995, 1996) and Varian (1980) evaluate incentive and insurance effects of social insurance provided by a progressive fiscal policy in a setting with ex ante representative households and missing credit and/or insurance markets. Interim or ex post Pareto improvements do not arise in those settings. Subsequent literature in public economics has

[11] In fact, Mookherjee and Ray (2003) showed existence of Pareto efficient steady states under laissez faire, which indicates the key role of ability heterogeneity in our model (since laissez faire constitutes a special case of unconditional transfers). The role of ability heterogeneity was investigated in an earlier paper of ours (Mookherjee and Napel (2007)) on uniqueness and stability of steady states under laissez faire, in the presence of paternalistic (instead of non-paternalistic) parental altruism; welfare effects of fiscal policy were not addressed.

examined implications of redistributive tax distortions for education subsidies.[12] Bovenberg and Jacobs (2005) have argued in a static model without any borrowing constraints or income risk that redistributive taxes and education subsidies are 'Siamese twins': the latter are needed to counter the effects of the former in dulling educational incentives. Jacobs, Schindler and Yang (2012) show the same result obtains when the model is extended to a context with uninsurable income risk. Unlike our paper, these arguments for educational subsidies arise from pre-existing income tax distortions, which disappear in the case of a laissez faire status quo. None of these models incorporate ability heterogeneity and missing credit markets, which create an efficiency role for educational subsidies in our model, even in the absence of any progressive income taxes.

5.3 Macroeconomics

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 and fiscal policy.[13] Most of these papers examine dynamic properties of competitive equilibria, and show that redistributive policies could raise aggregate output and welfare, but do not explore the possibility of Pareto improving fiscal policy. An exception is Bénabou (1996), who shows that collective financing of education can be ex post Pareto improving in a sufficiently patient society, similar to our Theorem 2.

Versions of these models have been calibrated to fit data of real economies in order to evaluate the welfare and macroeconomic effects of various fiscal policies in numerical simulations.[14] These studies rely on specific functional forms for technology and preferences,

---

[12] A large part of the recent dynamic public finance literature (e.g., Golosov, Tsyvinski and Werning (2006), or Golosov, Troshkin and Tsyvinski (2016)) is unrelated insofar as it abstracts from human capital investments and assumes that skills follow an exogenous Markov process. Its focus is to extend the Mirrlees (1971) optimal income tax model to a dynamic setting and examine consequences for optimal taxation of labor and savings. Other strands of literature on public education address its political economy. For instance, Glomm and Ravikumar (1992) study an endogenous growth model and compare human capital accumulation under laissez faire to public schooling funded by a linear income tax, where the tax rate is determined by a political majority. The focus there is on macroeconomic implications (per capita income, inequality and resulting tradeoffs), rather than the scope for efficiency enhancing interventions.

[13] See, e.g., Loury (1981), Aiyagari (1994), Aiyagari, Greenwood and Sheshadri (2002), Bénabou (1996, 2002).

[14] See Heathcote (2005), Bohacek and Kapicka (2008), Cespedes (2014), Berriel and Zilberman (2011), Abbott, Gallipoli, Meghir and Violante (2013), Findeisen and Sachs (2016, 2017), and Peruffo and Ferreira

and focus on aggregate measures of welfare. These papers leave open the question whether there may exist other policies which could have resulted in a Pareto improvement, or what the effects might be in economies with different preferences and technology. Our paper complements this literature by providing purely qualitative results concerning Pareto improving fiscal policies which apply irrespective of the specific welfare function, technology or preferences.

6 Concluding Observations

We have provided theoretical arguments for Pareto-superiority of cash transfers that condition on investments in child education, in a second-best environment with imperfect financial markets, and privately observed learning ability. Pareto-improvements arise when the implicit education subsidy is funded by income taxes imposed on the same income/occupational class, thereby avoiding redistribution across classes. The results apply generally, irrespective of specific assumptions on preferences or technology, initial conditions, equilibrium selection, and incorporate short as well as long run effects. We have argued the results also apply irrespective of labor supply elasticity or investment divisibility. However, the results concerning desirability of subsidizing education apply only if parents do not supplement education investments with financial bequests, which seems plausible for poor households. For wealthy households that leave financial bequests, Pareto optimality requires an opposite policy involving educational taxes or fees which fund unconditional transfers within the same class. Hence conditional transfers, broadly defined, continue to be Pareto efficient in the second-best world with financial market imperfections, and heterogeneity of child learning ability observed privately by parents.

Further generalizations are desirable but left for future research. For instance, a child's future wage income and financial inheritance could be subject to random shocks. Analysis of the corresponding extension of the scenario considered in Section 4.4 would be complicated by gains from diversifying the risky payoffs to financial vs. educational investment. For the very poor households who do not leave financial bequests, the identified pattern of parental consumption however should prevail, suggesting that a scheme along the indicated lines could still make everyone better off interim.

One question we did not address is the underlying source of missing markets for credit

(2017).

or insurance. Couldn't members of, say, the unskilled occupation – or profit-maximizing companies – organize a similar kind of scheme as does the government in our model? Why is public intervention needed? Mutual aid and benefit societies, fraternal lodges, trade unions and guilds have historically provided many private insurance services that have been taken over – and to some extent crowded out – by the welfare state (see Beito (2000)). Such societies usually have better social monitoring and enforcement possibilities than commercial companies. Still, collective education financing at more than a very localized scale seems to have been the exception rather than the rule. One can only speculate what the underlying reasons may have been – adverse selection (associated with opportunistic non-participation of parents who do not expect to benefit from it ex post), or the general equilibrium effects of such schemes (which lower the profitability of private insurance firms and households owing to the induced changes in the skill premium in the labor market) that are neutralized by the government in our construction.

References

Abbott B., Gallipoli G., Meghir C. and Violante G.L. (2013). "Education Policy and Intergenerational Transfers in Equilibrium," NBER Working Paper No. 18782.

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.

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.

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.

Beito D.T. (2000), *From Mutual Aid to the Welfare State – Fraternal Societies and Social Service, 1890–1967*, Chapel Hill, NC: University of North Carolina Press.

Bénabou R. (1996), "Heterogeneity, Stratification, and Growth: Macroeconomic implications of Community Structure and School Finance," *American Economic Review* 86(3), 584–609.

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.

Bidadanure J.U. (2019), "The Political Theory of Universal Basic Income" *Annual Review of Political Science* 22, 481–501.

Bohacek R. and Kapicka M. (2008), "Optimal Human Capital Policies," *Journal of Monetary Economics* 55(1), 1–16.

Bovenberg A.L. and Jacobs B. (2005), "Redistribution and Education Subsidies are Siamese Twins," *Journal of Public Economics* 89(11–12), 2005–2035.

Cespedes N. (2014) "General Equilibrium Analysis of Conditional Cash Transfers," Working Paper No. 2014-25, Peruvian Economic Association.

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.

Fender J. and Wang P. (2003), "Educational Policy in a Credit Constrained Economy with Skill Heterogeneity," *International Economic Review* 44(3), 939–964.

Findeisen S. and Sachs D. (2016), "Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans," *Journal of Public Economics* 138, 1-21.

Findeisen S. and Sachs D. (2017), "Optimal Need-Based Financial Aid," Working Paper, University of Mannheim and European University Institute.

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

Ghatak M. and Maniquet F. (2019), " Some Theoretical Aspects of a Universal Basic Income Proposal," *Annual Review of Economics* 11, forthcoming.

Ghatak M. and Muralidharan K. (2019), "An Inclusive Growth Dividend: Reforming the Role of Income Transfers in India's Anti-Poverty Strategy," forthcoming, 2018/19 India Policy Forum, National Council of Applied Economic Research, New Delhi.

Glomm G. and Ravikumar B. (1992), "Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality," *Journal of Political Economy* 100(4), 818–834.

Golosov M., Troshkin M. and Tsyvinski A. (2016) "Redistribution And Social Insurance," *American Economic Review* 106(2), 359–386.

Golosov M., Tsyvinski A. and Werning I. (2006), "New Dynamic Public Finance: A User's Guide," *NBER Macroeconomics Annual* 21, 317–388.

Government of India 2017 Economic Survey, Ministry of Finance, Government of India, New Delhi.

Hanna R. and Olken B.A. (2018), "Universal Basic Incomes versus Targeted Transfers: Anti-Poverty Programs in Developing Countries," *Journal of Economic Perspectives* 32(4), 201–226.

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

Ideas for India (2019), E-Symposium on Universal Basic Income in the Indian Context, https://www.ideasforindia.in/topics/poverty-inequality/introduction-to-e-symposium-the-idea-of-a-universal-basic-income-in-the-indian-context1.html.

Indian Journal of Human Development (2017), Symposium Issue on Universal Basic Income, vol. 11(2).

Jacobs J., Schindler D. and Yang H. (2012), "Optimal Taxation of Risky Human Capital," *Scandinavian Journal of Economics* 114(3), 908–931.

Levy S. (2006), *Against Poverty – Sustaining Mexico's Progresa-Oportunidades Program*, Washington, DC: Brookings Institution Press.

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.

Maoz Y.D. and Moav O. (1999), "Intergenerational Mobility and the Process of Development," *Economic Journal* 109(458), 677–697.

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.

Parker S. and Todd P.E. (2017), "Conditional Cash Transfers: The Case of Progresa/Oportunidades," *Journal of Economic Literature* 55(3), 866–915.

Peruffo, M. and Ferreira P.C. (2017), "The Long-term Effects of Conditional Cash Transfers on Child Labor and School Enrollment," *Economic Inquiry* 55(4), 2008–2030.

Sinn H.W. (1995), "A Theory of the Welfare State," *Scandinavian Journal of Economics* 97(4), 495–526.

Sinn H.W. (1996), "Social Insurance, Incentives and Risk Taking," *International Tax and Public Finance* 3(3), 259–280.

Van Parijs P. and Vanderborght Y. (2017), *Basic Income – A Radical Proposal for a Free Society and a Sane Economy*, Cambridge, MA: Harvard University Press.

Varian H.R. (1980), "Redistributive Taxation as Social Insurance," *Journal of Public Economics* 14(1), 49–68.

Appendix: Proofs

Proof of Lemma 1: Part (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 (i) and (ii) above. To show the next claim in (iv), suppose it is not true. Then we can find a subsequence $\{x_{c,t_n}\}_{n=1,2,\ldots}$ 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} must tend to 0 for *both* occupations $d = 0,1$, 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} must tend to ∞ for both occupations $d = 0,1$ by virtue of (3). Equation (2) then implies λ_{t_n+1} 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 which ruled out that sequence $\{x_{ct}\}_{t=1,2,\ldots}$ has a cluster point at 0 also ensures λ_t is bounded away from $\bar{\lambda}$. The bounds on consumption follow from the bounds on wages and on investment thresholds. ■

The next 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 non-negative budgetary surpluses $\{R_t\}_{t=0,1,\ldots}$ resulting from a fiscal policy $\{\tau_{ct}, e_{ct}\}_{c;t}$ and an associated DCE $\{\lambda_t, w_{ct}, x_{ct}, \mathcal{U}_{ct}\}_{c;\,t}$, suppose that the surplus is strictly positive at some date. Then 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:

$$\begin{aligned} R_{ct} &= \alpha_{ct}[\Delta\tau_{ct} + F(x_{ct})\Delta e_{ct}] \\ u(w_{ct}+\tau_{ct}) - u(w_{ct}+\tau_{ct}+e_{ct}-x_{ct}) &= u(w_{ct}+\tau_{ct}+\Delta\tau_{ct}) \\ &\quad - \; u(w_{ct}+\tau_{ct}+\Delta\tau_{ct}+e_{ct}+\Delta e_{ct}-x_{ct}) \end{aligned} \tag{16}$$

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 (16) 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. ∎

Proof of Theorem 1:

The proof of Theorem 1 proceeds in five steps.

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

(i) λ_t is uniformly bounded away from 0 and 1;

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

(iii) consumptions of all agents are uniformly bounded.

To see this note that the bounds on income transfers and on marginal products g_c over $[0, 1]$

imply that post-tax incomes are uniformly bounded. These imply existence of: a uniform upper bound on consumption (since consumption is bounded above by post-tax income); a uniform upper bound on W_{ct} (given the upper bound on consumption); a uniform lower bound on W_{ct} (from the option of always consuming all post-tax income); and, consequently, a uniform upper bound on $\Delta W_t = W_{1t} - W_{0t}$.

The latter also is a uniform upper bound on the utility sacrifice of investing parents. Combined with the uniform bounds on post-tax incomes, we infer that investment thresholds x_{ct} are uniformly bounded above, which in turn implies equilibrium consumption is uniformly bounded from below. So (iii) holds. Condition (b) implies post-tax income differences between the skilled and unskilled occupation are bounded away from zero. Hence ΔW_t is uniformly bounded away from zero, implying the same for investment thresholds. This establishes (ii).

The first part of (ii) implies that λ_t is uniformly bounded away from 1 because the distribution of education costs has full support on $\mathbb{R}_+$. Moreover, full support and uniform positive lower bound on investment thresholds ensure that λ_t is uniformly bounded away from 0. So (i) holds.

Step 2: For arbitrary $\epsilon > 0$, construct the following policy change. Denote the status quo DCE by a * superscript. Choose any occupation c and select alternative fiscal policy $\tau'_{ct}(\epsilon) = \tau_{ct} - \epsilon F(x^*_{ct}), e'_{ct}(\epsilon) = \epsilon(1 - \mu_t)$ for this occupation, while leaving that for the other occupation $d \neq c$ unchanged, where

$$\mu_t \equiv (1 - F(x^*_{ct}))\left[1 - \frac{F(x^*_{ct})u'(w^*_{ct} + \tau_{ct})}{\int_0^{x^*_{ct}} u'(w^*_{ct} + \tau_{ct} - x)dF(x)}\right] \tag{17}$$

independently of ϵ. 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.

For either occupation $i \in \{c, d\}$, define post-reform post-tax wages: $w^\epsilon_{it} = w^*_{ct} + \tau_{ct} - \epsilon F(x^*_{ct})$ if $i = c$, and $w^*_{dt} + \tau_{dt}$ otherwise, and education subsidies: $e^\epsilon_{it} = \epsilon(1 - \mu_t)$ if $i = c$ and 0 otherwise. Let the corresponding investment thresholds be denoted x^ϵ_{it} and dynastic utilities be denoted W^ϵ_{it}. In the following we refer to the effect of a policy reform where the value of ϵ is raised slightly above 0.

Claim: Assuming that the policy change leaves after-tax wages unchanged for each occupation, it generates: (i) a positive first order increase in investment thresholds for parents in occupation c at every t; (ii) a zero first order effect on thresholds for parents in occupation d at every $t \geq 1$, and (iii) a zero first order effect on the dynastic utilities at every

$t = 0, 1, 2, \ldots$. Specifically, $\frac{\partial x^{\epsilon}_{ct}(0)}{\partial \epsilon}$ is positive and uniformly (with respect to t) bounded away from zero, while $\frac{\partial x^{\epsilon}_{dt}(\epsilon)}{\partial \epsilon}$ and $\frac{\partial W^{\epsilon}_{it}(\epsilon)}{\partial \epsilon}, i = c, d$ converge uniformly (with respect to t) to zero as $\epsilon \downarrow 0$.

To prove this, we proceed as follows. For arbitrary thresholds $x_{it}, i \in \{c, d\}$ define the C^2 function

$$\mathcal{U}_{it}(x_{it}, \epsilon) \equiv u\big(w^{\epsilon}_{it})\big)[1 - F(x_{it})] + \int_0^{x_{it}} u\big(w^{\epsilon}_{it} + e^{\epsilon}_{it} - x\big) dF(x). \tag{18}$$

And define for $i \in \{c, d\}$:

$$W^{\epsilon}_{it} = \mathsf{E}_x V^{\epsilon}_t(i, x) \tag{19}$$

where

$$V^{\epsilon}_t(i, x) \equiv \max_{I \in \{0,1\}} \Big[u\big(w^{\epsilon}_{it} + I\{e^{\epsilon}_{it} - x\}\big) + \delta \mathsf{E}_x V^{\epsilon}_{t+1}(I, x)\Big]. \tag{20}$$

We restrict $\epsilon \leq \bar{\epsilon}$ for some bound $\bar{\epsilon} < \infty$. Given any such bound, it is evident that $w^{\epsilon}_{it}, e^{\epsilon}_{it}$ and therefore consumptions are uniformly bounded above (i.e., for all $i \in \{c, d\}$, all t and all $\epsilon \leq \bar{\epsilon}$). (In what follows, statements concerning uniform bounds will be taken to mean this.) Hence W^{ϵ}_{it} is also uniformly bounded above.

Next, note that W^{ϵ}_{it} is also uniformly bounded below, since a parent at any t and in any occupation i always has the option of not investing in education for its child and consuming its (uniformly bounded) post-tax income. So (given the intertemporal consistency of dynastic utility) the present value dynastic payoff associated with the parent as well as all subsequent descendants never investing forms a lower bound to W^{ϵ}_{it}.

Define $\Delta W^{\epsilon}_t \equiv W^{\epsilon}_{1t} - W^{\epsilon}_{0t}$, which constitutes the post-reform return to any dynasty to investing at $t-1$. The preceding arguments imply that this return ΔW^{ϵ}_t is uniformly bounded above. So parents' sacrifices associated with investment must be uniformly bounded above, implying that post-reform consumptions will be uniformly bounded from below.

In what follows, we shall say that a family of real-valued functions $\{y^{\epsilon}_t; \epsilon \in [0, \bar{\epsilon}), t \geq 0\}$ *satisfies the Cauchy (C) property* if given any $\eta > 0$, there exists $\zeta > 0$ such that $\sup_t |y^{\epsilon_1}_t - y^{\epsilon_2}_t| < \eta$ whenever $\epsilon_1, \epsilon_2 < \zeta$.

Since $|w^{\epsilon_1}_{it} - w^{\epsilon_2}_{it}| \leq (\epsilon_1 + \epsilon_2) F(x^*_{it}) \leq 2\bar{\epsilon}$, it follows that w^{ϵ}_{it} satisfies the C-property. By a similar argument, e^{ϵ}_{it} also satisfies the C-property.

We claim that W^{ϵ}_{it} satisfies the C-property. To prove this, let us define consumptions conditional on investment I and ability draw x as follows: $c^{\epsilon}_{it}(I, x) \equiv w^{\epsilon}_{it} + I[e^{\epsilon}_{it} - x]$. Then $|c^{\epsilon_1}_{it}(I, x) - c^{\epsilon_2}_{it}(I, x)| = \big|w^{\epsilon_1}_{it} - w^{\epsilon_2}_{it} + I[e^{\epsilon_1}_{it} - e^{\epsilon_2}_{it}]\big| < 2\eta$ for all i, t if $\epsilon_1, \epsilon_2 < \zeta$. Since consumptions are uniformly bounded, marginal utility is uniformly bounded above and away from zero.

Hence for any given η^*, there exists $\bar{\epsilon}$ such that $|u(c_{it}^{\epsilon_1}(I,x)) - u(c_{it}^{\epsilon_2}(I,x))| < \eta^*$ for all I, t, x, i if $\epsilon_1, \epsilon_2 < \bar{\epsilon}$. This implies that $|V_{it}^{\epsilon_1}(i,x) - V_{it}^{\epsilon_2}(i,x)| < \frac{\eta^*}{1-\delta}$ if $\epsilon_1, \epsilon_2 < \bar{\epsilon}$ for all i, x. Hence W_{it}^{ϵ} satisfies the C-property.

Next, we claim for suitable choice of $\bar{\epsilon}$, ΔW_t^{ϵ} is uniformly bounded away from zero. The argument used in Step 1 established existence of $b > 0$ such that $\Delta W_t^0 > b$ for all t. The claim in the previous paragraph implies that ΔW_t^{ϵ} satisfies the C-property. Hence there exists $\zeta > 0$ such that $|\Delta W_t^{\epsilon} - \Delta W_t^0| < \frac{b}{2}$ for all t if $\epsilon < \zeta$. This implies $\Delta W_t^{\epsilon} > \frac{b}{2}$ for all t if $\epsilon < \zeta$.

Now investment threshold x_{it}^{ϵ} solves:

$$u(w_{it}^{\epsilon}) - u(w_{it}^{\epsilon} + e_{it}^{\epsilon} - x_{it}^{\epsilon}) = \delta \Delta W_{t+1}^{\epsilon} \tag{21}$$

which is a well-defined positive real number owing to the Implicit Function Theorem. Next, observe that x_{it}^{ϵ} satisfies the C-property, since $w_{it}^{\epsilon}, e_{it}^{\epsilon}, \Delta W_t^{\epsilon}$ all satisfy the C-property, and since marginal utility is uniformly bounded away from zero (since consumption is uniformly bounded above). Since x_{it}^0 is uniformly bounded away from zero, the same must be true for x_{it}^{ϵ}.

The preceding results imply that $x_{it}^{\epsilon}, \Delta W_t^{\epsilon}, w_{it}^{\epsilon}, e_{it}^{\epsilon}$ converge uniformly to $x_{it}^*, \Delta W_t^0, w_{it}^*, e_{it}$ respectively as ϵ goes to zero. Hence $\mathcal{U}_{it}(x_{it}^{\epsilon}, \epsilon)$ converges uniformly to $\mathcal{U}_{it}(x_{it}^*, 0)$, i.e., for every $\eta > 0$ there exists a $\zeta > 0$ such that $|\mathcal{U}_{it}(x_{it}^{\epsilon}, \epsilon) - \mathcal{U}_{it}(x_{it}^*, 0)| < \eta$ if $\epsilon < \zeta$.

Next, note that for any y

$$\frac{\partial \mathcal{U}_{ct}(y,0)}{\partial \epsilon} = -[1 - F(y)]u'(w_{ct}^{\epsilon})F(x_{ct}^*) + [1 - \mu_t - F(x_{ct}^*)] \int_0^y u'(w_{it}^{\epsilon} + e_{it}^{\epsilon} - x) dF(x) \tag{22}$$

while $\frac{\partial \mathcal{U}_{dt}(y,0)}{\partial \epsilon} = 0$. So the preceding results and the definition of μ_t imply that $\frac{\partial \mathcal{U}_{it}(x_{it}^{\epsilon},\epsilon)}{\partial \epsilon}$ converges uniformly to 0.

Moreover

$$\Delta W_t^{\epsilon} \equiv \sum_{k=0}^{\infty} \nu_k^{\epsilon} \Big[\mathcal{U}_{1,t+k}(x_{1,t+k}^{\epsilon}, \epsilon) - \mathcal{U}_{0,t+k}(x_{0,t+k}^{\epsilon}, \epsilon) \Big] \tag{23}$$

where $\nu_0^{\epsilon} \equiv 1$ and $\nu_k^{\epsilon} \equiv \delta^k \Pi_{l=0}^{k-1} [F(x_{1,t+l}^{\epsilon}) - F(x_{0,t+l}^{\epsilon})]$. Since $\nu_k^{\epsilon} \leq \delta^k < \delta < 1$, the preceding results allow the order of summation and differentiation operators to be interchanged in the following expression:

$$\frac{\partial \Delta W_t^{\epsilon}}{\partial \epsilon} = \sum_{k=0}^{\infty} \nu_k^{\epsilon} \Big[\frac{\partial \mathcal{U}_{1,t+k}(x_{1,t+k}^{\epsilon}, \epsilon)}{\partial \epsilon} - \frac{\partial \mathcal{U}_{0,t+k}(x_{0,t+k}^{\epsilon}, \epsilon)}{\partial \epsilon} \Big] \tag{24}$$

where we use the Envelope Theorem to ignore the effects of changes in ϵ on the optimally chosen (interior) investment thresholds. It follows that $\frac{\partial \Delta W_t^\epsilon}{\partial \epsilon}$ converges uniformly to 0 as ϵ goes to 0. Moreover, by a similar argument, since $\frac{\partial \mathcal{U}_{it}(x_{it}^\epsilon,\epsilon)}{\partial \epsilon}$ converges uniformly to 0, the same is true of $\frac{\partial \mathcal{W}_{it}^\epsilon}{\partial \epsilon}$.

As we vary ϵ from 0, we claim that the threshold x_{ct} undergoes a first order increase, while the first order change in x_{dt} is zero. To see this, differentiating (21) for $i = d$ yields $\frac{\partial x_{dt}^\epsilon(0)}{\partial \epsilon} = 0$ because $\frac{\partial \Delta W_t^\epsilon(0)}{\partial \epsilon} = 0$. Moreover, from the uniform convergence of $\frac{\partial \Delta W_t^\epsilon}{\partial \epsilon}$, we can conclude that $\frac{\partial x_{dt}^\epsilon(\epsilon)}{\partial \epsilon}$ converges uniformly to 0 as ϵ goes to 0. In contrast, for $i = c$ we obtain

$$\frac{\partial x_{ct}^\epsilon(0)}{\partial \epsilon} = F(x_{ct}^*)\frac{u'(w_{ct}^* + \tau_{ct})}{u'(w_{ct}^* + \tau_{ct} - x_{ct}^*)} + (1 - \mu_t - F(x_{ct}^*)). \tag{25}$$

The concavity of u implies

$$\int_0^{x_{ct}^*} u'(w_{ct}^* + \tau_{ct} - x)dF(x) < F(x_{ct}^*)u'(w_{ct}^* + \tau_{ct} - x_{ct}^*). \tag{26}$$

Hence recalling the definition (17) of μ_t,

$$\mu_t < (1 - F(x_{ct}^*))\Big[1 - \frac{u'(w_{ct}^* + \tau_{ct})}{u'(w_{ct}^* + \tau_{ct} - x_{ct}^*)}\Big], \tag{27}$$

and substituting this into (25) we obtain

$$\frac{\partial x_{ct}^\epsilon(0)}{\partial \epsilon} > \frac{u'(w_{ct}^* + \tau_{ct})}{u'(w_{ct}^* + \tau_{ct} - x_{ct}^*)} \tag{28}$$

which is uniformly bounded away from 0. This concludes the proof of the Claim.

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 $x_{ct}^\epsilon(\epsilon), x_{dt}^\epsilon(\epsilon)$

$$\lambda_{t+1}(\epsilon) = F(x_{1t}^\epsilon(\epsilon))\lambda_t(\epsilon) + F(x_{0t}^\epsilon(\epsilon))(1 - \lambda_t(\epsilon)) \tag{29}$$

with $\lambda_0(\epsilon) = \lambda_0$ given. Note that (28) combined with $x_{1t}^\epsilon(\epsilon) > x_{0t}^\epsilon(\epsilon)$ 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 $t \geq 1$

$$\tilde{\tau}_{ct}(\epsilon) = w_{ct}^* - w_{ct}(\epsilon) + \tau_{ct}'(\epsilon) \equiv w_{ct}^* - w_{ct}(\epsilon) + \tau_{ct} - \epsilon F_{ct}^* \tag{30}$$

$$\tilde{e}_{ct}(\epsilon) = e_{ct}'(\epsilon) \equiv \epsilon(1 - \mu_t) \tag{31}$$

$$\tilde{\tau}_{dt}(\epsilon) = w_{dt}^* - w_{dt}(\epsilon) + \tau_{dt}' \equiv w_{dt}^* - w_{dt}(\epsilon) + \tau_{dt} \tag{32}$$

$$\tilde{e}_{dt}(\epsilon) = e_{dt}'(\epsilon) \equiv 0 \tag{33}$$

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 $\{x^\epsilon_{ct}(\epsilon), x^\epsilon_{dt}(\epsilon)\}_{t=0,1,2\ldots}$ and the interim utilities $\{\mathcal{U}_{ct}(x^\epsilon_{ct}(\epsilon),\epsilon), \mathcal{U}_{dt}(x^\epsilon_{dt}(\epsilon),\epsilon)\}_{t=0,1,2,\ldots}$ which were constructed in Step 2 under the assumption of unchanged after-tax wages in each occupation at each date. Given investment thresholds $x^\epsilon_{ct}(\epsilon), x^\epsilon_{dt}(\epsilon)$ 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 (30)–(33), 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

$$\begin{aligned} B_t(\epsilon) &= w_{0t} - w^*_{0t} - \tau_{0t} \\ &\quad + \lambda_t\big[(\tau_{0t} - \tau_{1t}) + (w^*_{0t} - w_{0t}) - (w^*_{1t} - w_{1t}) - \epsilon\{F_{1t}(1-\mu) - F^*_{1t}\}\big] \end{aligned} \tag{34}$$

where we abbreviate $F(x^\epsilon_{it})$ and w^ϵ_{it} by F_{it} and w_{it}, respectively. Indicating the corresponding derivatives w.r.t. ϵ by F'_{it} and w'_{it}, we then have

$$\begin{aligned} B'_t(\epsilon) &= w'_{0t} + \lambda'_t\big[(\tau_{0t} - \tau_{1t}) + (w^*_{0t} - w_{0t}) - (w^*_{1t} - w_{1t}) - \epsilon\{F_{1t}(1-\mu_t) - F^*_{1t}\}\big] \\ &\quad - \lambda_t\big[w'_{0t} - w'_{1t} + \{F_{1t}(1-\mu_t) - F^*_{1t}\} + \epsilon F'_{1t}(1-\mu_t)\big]. \end{aligned} \tag{35}$$

We can use Euler's theorem to obtain $\lambda_t w'_{1t} + (1-\lambda_t) w'_{0t} \equiv 0$, i.e., the corresponding terms in (35) cancel. The progressivity assumption $\tau_{0t} \geq \tau_{1t}$ in condition (a) and $\lambda'_t \geq 0$ for all ϵ sufficiently close to zero imply we may drop the corresponding term to bound (35) from below:

$$\begin{aligned} B'_t(\epsilon) &\geq \lambda_t F_{1t}\mu_t + \lambda'_t\big[(w^*_{0t} - w_{0t}) - (w^*_{1t} - w_{1t}) - \epsilon\{F_{1t} - F^*_{1t}\} + \epsilon F_{1t}\mu_t\big] \\ &\quad - \lambda_t\big[\{F_{1t} - F^*_{1t}\} + \epsilon F'_{1t}(1-\mu_t)\big]. \end{aligned} \tag{36}$$

The first term, $\lambda_t F_{1t}\mu_t$, is positive and uniformly bounded away from zero. We want to show that the rest of the RHS can be brought arbitrarily close to zero by choosing $\underline{\epsilon} > 0$ sufficiently small, whence we could conclude that $B'_t(\epsilon) > 0$ for all t and all $\epsilon \in [0, \underline{\epsilon})$.

We claim that λ'_t is uniformly bounded. Apply $\lambda'_{t+1} = F'_{1t}\lambda_t + F'_{0t}(1-\lambda_t) + (F_{1t} - F_{0t})\lambda'_t$ recursively to obtain

$$\lambda'_{t+1} = F'_{1t}\lambda_t + F'_{0t}(1-\lambda_t) + \sum_{k=1}^{t}(F_{1t} - F_{0t})\cdot\ldots\cdot(F_{1,t-k+1} - F_{0,t-k+1})[F'_{1,t-k}\lambda_{t-k} + F'_{0,t-k}(1-\lambda_{t-k})]. \tag{37}$$

Now F'_{1t} and F'_{0t} are uniformly bounded since the density of the ability distribution is bounded, and preceding arguments imply that $\frac{\partial x^{\epsilon}_{it}}{\partial \epsilon}$ is uniformly bounded. Hence any convex combination of F'_{1t} and F'_{0t} is uniformly bounded. So there exists $m > 0$ such that $[F'_{1,t-k}\lambda_{t-k} + F'_{0,t-k}(1-\lambda_{t-k})] < m$ for all $t \geq k$ if $\epsilon \in [0, \underline{\epsilon})$. There also exists $\alpha \in (0,1)$ such that $(F_{1,t-k+1} - F_{0,t-k+1}) < \alpha$ for all $t \geq k$ and all $\epsilon \in [0, \bar{\epsilon}]$. Hence $\lambda'_{t+1} < \frac{m}{1-\alpha}$, i.e., is uniformly bounded.

Choosing $\underline{\epsilon} > 0$ small enough therefore allows to make $\lambda'_t[(w^*_{0t} - w_{0t}) - (w^*_{1t} - w_{1t})]$ arbitrarily small: the skill change $\lambda^*_t - \lambda_t(\epsilon)$ can be made arbitrarily small; this extends to wage changes.

Given that $\lambda'_t F_{1t}\mu_t$ and $F'_{1t}(1-\mu_t)$ are uniformly bounded, products of ϵ and these terms in (36) can be made arbitrarily small for $\epsilon \in [0, \underline{\epsilon})$ through an appropriate choice of $\underline{\epsilon} > 0$. Because $\lambda_t F_{1t}\mu_t$ is uniformly bounded away from zero, this in summary means we can find $\underline{\epsilon} > 0$ such that

$$B'_t(\epsilon) \geq \frac{1}{2} \cdot \lambda_t F_{1t}\mu_t \tag{38}$$

for all t and $\epsilon \in [0, \underline{\epsilon})$. As $\lambda_t F_{1t}\mu_t$ is uniformly bounded away from zero, a small policy reform starting from $\epsilon = 0$ therefore generates a strictly positive budget surplus.

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

Proof of Theorem 2

We denote the respective investment thresholds, skill ratios, wages, etc. in the DCELF associated with a given discount factor δ by x^{δ}_{ct}, λ^{δ}_t, w^{δ}_{ct}, etc. Recall that the induced skill ratio λ^{δ}_t must be strictly smaller than $\bar{\lambda}$ for every δ and t (cf. Lemma 1). From this follows that if we fix an arbitrary $\delta' > 0$ there exist $\underline{x}$ and $\bar{x}$ such that $0 < \underline{x} \leq x^{\delta}_{0t} \leq x^{\delta}_{1t} \leq \bar{x} < \infty$ for all t and all $\delta \in (\delta', 1)$. To see this, suppose otherwise, i.e., that there exists a sequence $\{t_n, \delta_n\}_{n=1,2,\ldots}$ such that (i) $x^{\delta_n}_{c,t_n} \to 0$ or (ii) $x^{\delta_n}_{c,t_n} \to \infty$. In case (i), vanishing investment by occupation c in period t_n requires that the benefit of having a skilled child in $t_n + 1$ vanishes. Then no parent in occupation $d \neq c$ would have an incentive to invest in t_n either, implying $\lambda_{t_n+1} \approx 0$. The consequent gap between skilled and unskilled equilibrium wages in period $t_n + 1$ and $\delta_n > \delta' > 0$ would then, however, induce a non-vanishing benefit of one's child to be skilled in $t_n + 1$ – a contradiction. In case (ii), benefits of having a skilled child in $t_n + 1$ would need to grow without bound. This implies that parents in occupation

$d \neq c$ will also find it optimal to invest in t_n for arbitrarily large x. But $x^\delta_{0,t_n}, x^\delta_{1,t_n} \to \infty$ implies $\lambda^\delta_{t_n+1} \to 1$, in contradiction to $\lambda^\delta_t < \bar{\lambda}$ for all t.

We next establish that there exist $\delta' > 0$ and $\underline{b}$ such that

$$F(x^\delta_{1t})\lambda^\delta_t \mu^\delta_t \geq \underline{b} > 0 \text{ for all } t \text{ and } \delta \in (\delta', 1) \tag{39}$$

with

$$\mu^\delta_t \equiv (1 - F(x^\delta_{ct}))\left[1 - \frac{F(x^\delta_{ct})u'(w^\delta_{ct})}{\int_0^{x^\delta_{ct}} u'(w^\delta_{ct} - x)dF(x)}\right] \tag{40}$$

for a given $c \in \{0, 1\}$. To see (39) note, first, that $F(x^\delta_{1t}) \geq F(\underline{x}) > 0$ because F is strictly increasing. Second, $\lambda^\delta_t < \bar{\lambda}$ implies $w^\delta_{1t} > w^\delta_{0t}$ and $x^\delta_{1t} > x^\delta_{0t}$. From this follows $\lambda^\delta_t \geq F(x^\delta_{0t}) \geq F(\underline{x}) > 0$. Finally, note that $(1 - F(x^\delta_{ct})) \geq (1 - F(\bar{x})) > 0$ in (40). Now suppose that

$$\frac{u'(w^{\delta_n}_{ct_n})}{\frac{1}{F(x^{\delta_n}_{ct_n})}\int_0^{x^{\delta_n}_{ct_n}} u'(w^{\delta_n}_{ct_n} - x)dF(x)} \to 1$$

for some sequence $\{t_n, \delta_n\}_{n=1,2,\ldots}$. This would require $x^{\delta_n}_{ct_n} \to 0$, in contradiction to $x^\delta_{ct} \geq \underline{x} > 0$ for all t and $\delta \in (\delta', 1)$. Hence μ^δ_t is bounded away from zero.

Equation (39) implies that, by choosing $\epsilon \in (0, \bar{\epsilon})$ for a small $\bar{\epsilon} > 0$, a strictly positive budget surplus $B^\delta_t(\epsilon)$ is created for all t and $\delta \in (\delta', 1)$ in the first steps of the proof of Theorem 1. Since the full policy intervention is budget balancing, it must raise consumption in period t for at least one occupation $d \in \{0, 1\}$ by a non-vanishing amount. This must increase the respective interim consumption utility $\mathcal{U}^\delta_{dt}$ at a rate which is bounded away from zero, recalling that $g_1(0)$ is an upper bound to consumption in any laissez faire equilibrium and so every agent's marginal utility of consumption is bounded below by $u'(g_1(0)) > 0$. The final Step 5 of the proof of Theorem 1 makes sure that differences between $\mathcal{U}^\delta_{1t}$ and $\mathcal{U}^\delta_{0t}$ in the status quo are preserved. Therefore, also interim consumption utility in occupation $c \neq d$ must locally increase in ϵ at a rate that is bounded away from zero. Let $\underline{\nu} > 0$ denote the corresponding uniform lower bound on marginal improvements in expected utility associated with period t consumption.

Increases in period t's average consumption utility of the skilled and unskilled arise mainly because parents who educated their children already in the laissez-faire benchmark consume their net subsidy. For them, the consumption and the dynastic components of utility go up. The same applies to non-investors in occupation d if the first steps of the scheme are targeted at only one occupation c (because of transfers everyone receives in Step 5). However, non-investors in the targeted occupation are net contributors and have to reduce consumption. We will establish that the policy constitutes an ex post Pareto improvement by showing that, for δ high enough, their gain in dynastic utility outweighs the loss in consumption utility. Note here that newly investing parents in the targeted occupation suffer an even bigger drop in consumption – despite being net beneficiaries – than their non-investing peers. But they could have stayed non-investors and reveal to be yet better off by investing.

The key reason why also net contributors to the scheme are better off for high δ is that the rate at which $\epsilon > 0$ decreases these non-investing parents' consumption utility is bounded above. Namely, there exists $L < \infty$ such that

$$\frac{\partial}{\partial \epsilon}\Big[u(w^{\delta}_{ct}) - u\big(w^{\delta}_{ct} - \epsilon F(x^{\delta}_{ct}) + S^{\delta}_{ct}(\epsilon)\big)\Big]\bigg|_{\epsilon=0} = \Big(F(x^{\delta}_{ct}) - \frac{\partial S^{\delta}_{ct}(0)}{\partial \epsilon}\Big) \cdot u'(w^{\delta}_{ct}) \qquad (41)$$

$$\leq F(\bar{x}) \cdot u'(g_0(0)) < L \qquad (42)$$

where $S^{\delta}_{ct}(\epsilon)$ denotes the (increasing) budget surplus which is allocated to occupation c in Step 5 of the proof of Theorem 1.

In contrast, the rate at which $\epsilon > 0$ increases the non-investing parents' dynastic utility $\delta\, W^{\delta}_{0,t+1}$ is unbounded. Namely, for every $M < \infty$ there exists $\underline{\delta} \in (0,1)$ such that for all $\delta \in (\underline{\delta}, 1)$:

$$\frac{\partial}{\partial \epsilon}\Big\{\delta\Big[W^{\delta}_{0,t+1}(\epsilon) - W^{\delta}_{0,t+1}\Big]\Big\}\bigg|_{\epsilon=0} > M \qquad (43)$$

where $\mathcal{W}^{\delta}_{0,t+1} = \mathcal{W}^{\delta}_{0,t+1}(0)$ refers to interim welfare in the original DCELF. To see this, consider

$$\frac{\partial}{\partial \epsilon}\Big[W^{\delta}_{0,t+1}(\epsilon) - W^{\delta}_{0,t+1}\Big] = \frac{\partial}{\partial \epsilon}\left[\sum_{k=0}^{\infty}\delta^k \mathcal{U}^{\delta}_{0t}(\epsilon) + \delta\sum_{k=0}^{\infty}\delta^k F(x^{\delta}_{0,t+k}(\epsilon))\Delta W^{\delta}_{t+1+k}(\epsilon)\right] \qquad (44)$$

and note that the derivative at $\epsilon = 0$ of the second summand in the brackets is zero (cf. Step 2 in Proof of Theorem 1). The corresponding derivative of the first summand is

$$\sum_{k=0}^{\infty}\delta^k \frac{\partial \mathcal{U}^{\delta}_{0t}}{\partial \epsilon}\bigg|_{\epsilon=0} \geq \sum_{k=0}^{\infty}\delta^k \underline{\nu} = \frac{\underline{\nu}}{1-\delta}, \qquad (45)$$

and so the left-hand side of (43) grows without bound as $\delta \to 1$.

Combining (41) and (43), we can conclude that the total welfare change of non-investing parents satisfies

$$\frac{\partial}{\partial \epsilon}\left\{\Big[u(w^{\delta}_{ct} - \epsilon F(x^{\delta}_{ct}) + S^{\delta}_{ct}(\epsilon) + \delta\, W^{\delta}_{0,t+1}(\epsilon)\Big] - \Big[u(w^{\delta}_{ct}) + \delta\, W^{\delta}_{0,t+1}\Big]\right\}\Bigg|_{\epsilon=0} \geq \psi \qquad (46)$$

for all t and $\delta \in (\underline{\delta}, 1)$ for some $\psi > 0$. We can therefore choose $\bar{\epsilon} > 0$ such that each individual's ex post welfare change is positive for any $\epsilon \in (0, \bar{\epsilon})$ for every $\delta \in (\underline{\delta}, 1)$. $\square$