Welfare Rationales for Conditionality of 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: March 8, 2021

Abstract

We study efficiency and distributional effects of conditioning transfers on educational investments by parents, in an OLG model with missing financial markets and heterogeneity of learning ability. Conditional cash transfers (CCT) can be designed to generate a Pareto improvement relative to either laissez faire, or unconditional transfers such as universal basic income proposals. This applies irrespective of whether the status quo involves underinvestment or overinvestment in education from a first-best perspective, or the nature and extent of parental altruism towards children. The CCT corrects a market failure of insurance and lack of consumption smoothing for parents with respect to random realizations of ability of their offspring.

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

Declarations of interest: none

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

**Corresponding author.

1 Introduction

In developing countries, cash transfer programs are becoming increasingly widespread in government safety nets and related policy discussions. Prominent among these are conditional cash transfer (CCT) programs, where transfers are conditioned on school enrollment of children. Unconditional cash transfer (UCT) programs are less common, though there have been a number of policy experiments, and are generating increased interest in policy discussions on Universal Basic Income (UBI) proposals.[1] UCTs and CCTs are sometimes bracketed together in evaluations of cash transfer programs based on RCT experiments (e.g., Banerjee et al. 2017). While the value of conditionality of the transfers has received less attention in RCT experiments, a number of recent empirical policy evaluations in Mexico have assessed their effects on labor force participation and income.[2] These studies indicate larger effects of CCTs on education, labor force participation and income over longer time spans, particularly for women, indicating the role of conditionality in achieving long term poverty reduction. On the other hand, a number of possible disadvantages of CCTs have been mentioned in policy discussions: narrower coverage, greater paternalism and higher enforcement burdens (Mundle 2017), in some contexts also lack of progressive impact resulting from uneven takeup (Das, Do, and Özler 2004). Hence policy makers face complex tradeoffs in evaluating the merits and disadvantages of conditional transfers.[3]

The empirical evaluations of CCTs cited above have naturally focused on their partial equilibrium (PE) labor market impacts. They do not address impacts on other relevant outcomes such as parental consumption or welfare of taxpayers, which depend on how transfers are financed. These also matter in determining the overall welfare impact of CCTs, whether they are fundamentally redistributive or also have a role in improving efficiency by correcting a market failure. From a conceptual viewpoint, welfare objectives of efficiency or redistribution have traditionally provided the foundations for design of systems of social security, taxation and government welfare programs. The efficiency objective pertains to a market failure (or Pareto inefficiency) that a safety net program is designed to correct, while redistributive goals are incorporated in utilitarian or Rawlsian notions of justice. An explicit argument for either of these rationales requires a model clearly articulating underlying economic fundamentals (tastes, technology, market structure, information) and financing mechanisms needed to assess efficiency and distribution effects of an intervention. The latter are not just abstract criteria used by academic theorists but often play an important role in policy discussions that evaluate the trade-offs involved. For example a World

[1]See 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 (2020).

[2]See Behrman, Parker, and Todd (2011), Parker and Todd (2017), and Parker and Vogl (2018).

[3]To be clear, our focus is on the 'conditionality' dimension rather than the 'cash' dimension of transfers.

Bank discussion paper on CCTs (Das et al. 2004, p. 2f) posed the following questions:

> "A rationale for conditionality must then lie in the ability of such schemes to address underlying market failures. ... conditional cash transfers seek to restore efficiency in the economy. ... The argument developed in this paper is that evidence of an externality, though compelling, may be insufficient grounds for conditional cash transfer schemes without additional information on its extent. This argument relies on the observation that such schemes have historically been used for an entirely different purpose, that of targeting resources and pro-poor redistribution. ... These two very different rationales for conditional transfers result in a tension. When used to increase investment in human capital, such schemes could have adverse redistributive impacts. Conversely, used as targeting or redistributive mechanisms, they could decrease efficiency. One way for policy-makers to then decide on the overall benefits would be to obtain information on both sides of the coin. How do efficiency gains compare to adverse redistributive impacts when conditional cash transfers are implemented? Similarly, when used for targeting purposes, how successful was the targeting given the associated efficiency loss? ... Careful analysis and information on the gains and losses is then critical for the overall evaluation of the program."

A systematic micro-founded welfare analysis can help evaluate the different advantages and disadvantages of conditionality mentioned in policy discussions. Moreover, it can highlight other dimensions hitherto ignored in such discussions (such as implications for consumption smoothing, financing costs and general equilibrium (GE) effects). This constitutes the motivation of this paper, which studies the welfare rationale for CCTs and the related question of design from a normative standpoint. We examine the welfare rationale for CCTs both relative to a laissez faire economy with missing financial markets, and to unconditional transfers such as a UBI financed by a progressive income tax.

The first question is the comparison with a laissez faire benchmark: is there a market failure that CCTs can help overcome? This requires demonstration of the existence of a CCT which if suitably designed would achieve a Pareto improvement. In policy discussions of CCTs, it is frequently asserted that missing markets for credit and insurance, and/or parent-child externalities, create market failure (in the form of 'underinvestment') in education.[4] The notion of underinvestment usually invoked relates to a first-best benchmark, e.g., whether the rate of return to education among beneficiaries exceeded the social cost. In a world of missing markets or asymmetric information, first-best criteria of Pareto efficiency are not relevant and need to be replaced by 'constrained' Pareto efficiency, incorporating borrowing and informational constraints faced

[4]See e.g. López-Calva and Lustig (2010, p. 15), Kahhat (2010, p. 32) and Das et al. (2004, p. 8).

by governments. Matters are further complicated in an OLG setting, where standard theorems of welfare economics do not necessarily apply.

Indeed, it often turns out that steady states of a model with dynastic households and missing financial markets are constrained Pareto efficient, i.e., there is no market failure. A Ramsey neoclassical growth model provides a ready illustration: despite lack of access to any borrowing opportunity, an autarkic agent can accumulate own savings to converge to a steady state which is fully Pareto efficient. Mookherjee and Ray (2003) show this extends to a 'standard' OLG model of human capital accumulation where households lack access to financial markets. If there are indivisibilities in investment (e.g., in the form of a limited set of occupational choices), they show efficient as well as inefficient steady states co-exist. Moreover, when such indivisibilities become negligible (with a rich set of occupational options), the set of steady states shrinks to the one that is fully efficient. This implies that the common view that missing financial markets *necessarily* imply a market failure is incorrect.

The first main result of this paper is that the common intuition of a market failure is actually correct, once the 'standard' model is extended to incorporate heterogeneity of education cost across children.[5] Such heterogeneity may arise from variation in either cognitive learning ability or non-cognitive personality traits, besides family characteristics such as proximity to schools or parental capacity to assist and monitor child learning. Moreover, many of these characteristics are difficult to predict *ex ante*, while their *ex post* realization is known by parents (but maybe not the government). A large literature in the economics of education provides extensive empirical evidence of such heterogeneity.[6] We show that once this heterogeneity (referred to as 'ability heterogeneity' hereafter) is incorporated into the model, any laissez faire dynamic competitive equilibrium is interim-Pareto dominated by a suitably designed CCT policy. The interim Pareto domination requirement is very demanding: the expected utility of every parent (with a given realization of own income, but prior to learning the ability of her child) must rise, irrespective of occupation or generation.

The underlying argument is the following. Critical to the welfare effects is the way that conditional transfers are financed. The CCT mechanism we consider finances education subsidies for households in any given income class by general income taxes paid by households in the same class. *Ex post* ability heterogeneity implies some parents will decide not to invest in their children's education, while others will invest. The mechanism hence effectively resembles an insurance program that redistributes from families that do not invest in education of their children, to families earn-

[5]The effects of heterogeneous fertility will also be similar, as will become evident in the discussion below.

[6]See, e.g., Card (2001), Heckman, Stixrud, and Urzua (2006), Tsai and Xie (2011) and Henderson, Polachek, and Wang (2011).

ing the same income that do invest.[7] The efficiency improvement generated thereby is driven by a combination of greater investment in education (which raises welfare of succeeding generations), and superior smoothing of parental consumption (since the parents that earn the same and invest more end up consuming less under laissez faire). These welfare benefits arise only in the presence of ability heterogeneity – because with homogenous ability, parents of a certain income will make exactly the same education decisions, so there is no variation of parental consumption, conditional on their income. This also explains why the result does not apply in a static deterministic setting, where neither consumption smoothing nor investment play any role – conditionality of transfers then just ends up lowering the welfare of recipients, a manifestation of the welfare loss associated with increased paternalism. These static 'paternalistic' welfare losses get transformed into welfare gains in the dynamic setting with heterogeneity.

We show the same logic ensures the welfare dominance of CCTs over UCTs in which taxes and transfers may be conditioned on income or occupation, but not on educational decisions made by parents on behalf of their children.[8] This pertains to the debate on CCTs versus UBI, since the latter constitutes a uniform, unconditional level of financial support provided by the government to all citizens.

The welfare dominance of CCTs turns out to be robust to many extensions of our base model, pertaining to the nature of parental altruism, divisibility of educational investments, elastic labor supply or general equilibrium effects on wages. For instance, it applies irrespective of whether parental altruism is paternalistic or non-paternalistic[9], the intensity of altruism and accordingly whether or not there is under-investment or over-investment in a first-best sense. If parents are non-paternalistic and altruistic enough, the interim-Pareto improving CCT ends up achieving an ex post Pareto improvement as well. In that case, even parents that do not themselves invest in education end up benefitting from the policy in anticipation of the resulting utility benefit to succeeding generations.

Moreover, the resulting equilibrium allocations achieve superior macroeconomic performance on many relevant dimensions: in every generation, per capita income and skill rises, income gaps between skilled and unskilled become smaller, and there is greater upward mobility. Hence utility improvements do not accrue only to households that are better off to start with. If parental altruism is paternalistic, redistribution and efficiency objectives can be separated. If it is non-paternalistic, this is no longer the case, but the CCT can be designed to alter interim utilities

[7]A similar argument applies to fluctuations in total parental investment arising from variation in the number of children arising from fertility shocks.

[8]We impose very mild conditions on the progressivity of income taxes, that (effective) marginal tax rates are positive but less than 100%.

[9]Parental altruism is said to be non-paternalistic or paternalistic depending on whether parental utility functions are increasing in their children's utility, or in their children's future income.

of agents at different income levels in exactly the same way, thereby avoiding any redistribution *across* income classes. This prevents a dilution of parental incentives to invest as a means of achieving upward mobility of their children. With a CCT of this kind the efficiency improvement does not come at the cost of raising inequality across income classes, thereby addressing the concern raised by Das et al. (2004).

Another lesson from our analysis is the necessity to pay attention to dimensions that tend to be overlooked in policy discussions: mechanisms for financing transfers, and effects on parental consumption smoothing, over and above effects of transfer conditionality on education, upward mobility or labor force participation. There is an underlying market failure that the CCT corrects, but it is not a market failure in education. Education could be 'over-provided' under laissez faire in the standard first-best sense, and yet the CCT would generate a Pareto improvement at the same time that it induces even higher educational investments. The root source of market failure that is being corrected is one of missing markets for insurance (rather than credit) to parents against uncertain costs of investing in their children's human capital. Our analysis indicates the need for empirical analyses of CCTs to focus on this hitherto neglected dimension.[10] To the extent that CCTs in practice have benefitted better-off parents more, our results suggest the possibility of redesigning the subsidy and financing methods to avoid this problem.

In order to achieve the demanding requirement of a Pareto improvement, the CCT has to be designed carefully to realize the required improvements in efficiency and incentives. Are the informational or enforcement requirements of the resulting policy too demanding, thereby rendering it impractical? Is it likely to be politically feasible? These questions are difficult to answer in the abstract, and are likely to depend on details of the specific context. Theoretical arguments for efficiency or inefficiency of certain policies (such as free trade, or Pigouvian pollution taxes) however have traditionally been based on the notion of a potential Pareto improvement, as embodied in Kaldor-Hicks welfare criteria where 'in principle' losers from the policy could be compensated by the gainers. This is as far as one can go on the basis of theory alone: establishing 'proof of concept', a necessary first step before embarking on empirical analyses needed to design and evaluate specific versions of policies, or political feasibility. However, the design of the CCT that we consider in the paper does not seem any more complicated than the design of any insurance program in terms of choosing premiums and benefits to achieve both self-financing and significant take-up objectives.

On the other hand, we do not 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

[10] Some authors such as de Janvry et al. (2006) have studied the role of CCTs in providing parents with insurance against other external shocks, but not with respect to 'ability' risk of their children.

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 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, which is relevant to longer term considerations in which enforcement of conditionality is a lower-order concern.

The paper is structured as follows. Section 2 illustrates the market failure that CCTs can correct and how they improve on a status quo involving laissez faire or rather arbitrary combinations of UCTs. To clarify the exposition, we start with a simple setting with two occupations which abstracts from labor market GE effects, endogenous labor supply or the possibility of financial bequests. Section 3 discusses robustness of our results to these simplifying assumptions. Section 4 relates our analysis to existing literature, while Section 5 concludes.

2 Baseline Model and Results

2.1 Paternalistic Altruism

Consider an economy with multiple dates and a continuum of households, each comprising a parent and a child at any given date. There are two types of occupations, skilled ($c = 1$) and unskilled ($c = 0$); work in the former requires an indivisible educational investment when the agent is young. Parental earnings depend on their occupation, but are subject to exogenous shocks. Conditional on realized household income y, the parental occupation does not matter.

Let the cumulative distribution function of y be denoted G_c for parents in occupation c. Abstract initially from labor market GE effects, by assuming that G_c is not affected by the proportion of agents in occupation c. We can think of G_c as reflecting a skill-specific probability distribution of numbers of efficiency units of labor, which households supply to firms with a constant returns to scale production function on a competitive labor market. To simplify exposition, assume a common finite support $Y \equiv \{y_1, \ldots, y_n\}$ for both G_1 and G_0, with $y_1 > 0$ and $y_i < y_{i+1}$ for all $i = 1, \ldots, n-1$. The probability of income realization y_i in occupation $c \in \{0, 1\}$ is $\pi_{ic} > 0$. We assume $G_{j1} \equiv \sum_{i=1}^{j} \pi_{i1} < G_{j0} \equiv \sum_{i=1}^{j} \pi_{i0}$ for all $j = 1, \ldots, n-1$, so income distribution G_1 among the skilled (strongly) first order stochastically dominates distribution G_0 among unskilled households.

Every parent privately observes the idiosyncratic cost $\tilde{x}$ of educating its child; this represents the heterogeneity of learning abilities in the population: realization $\tilde{x}$ of the education cost of any child is drawn randomly and independently according to a cumulative distribution function F defined on $[0, \infty)$. F is C^2 and strictly increasing. Parental income is divided between consumption and education. Households cannot borrow against children's future earnings to finance $\tilde{x}$; and they can neither insure against income shocks nor the risk that the child's learning ability is high or low. So a parent with income realization y and a child of type $\tilde{x}$ takes education decision $e \in \{0, 1\}$ to maximize

$$u(y - e\tilde{x}) + [eV_1 + (1 - e)V_0] \tag{1}$$

where $V_c \equiv \sum_{i=1}^{n} \pi_{ic} V(y_i)$. Function u is strictly increasing, strictly concave and smooth on $(0, \infty)$ with $\lim_{c \to 0} u(c) = -\infty$. No restriction is imposed on the function V that reflects parental altruism towards the child, except that it is strictly increasing. Parents may under-value the benefits of higher earnings of their children, resulting in a large parent-child externality and 'underinvestment' in a first-best sense (based on pecuniary rate of return on education).[11] Or they could over-value it resulting in 'over-investment'. The benefit to the parent of educating the child is $B \equiv [V_1 - V_0] > 0$. The combination of paternalism and absence of labor market GE effects implies B is exogenous and stationary.

We consider a standard OLG model in which a child whose parent chose $e \in \{0, 1\}$ works in occupation $c = e$ in the next generation and maximizes (1) for new draws of income and child ability.[12] The proportion of population in the skilled occupation, λ, is then the dynamic state variable of interest. Still, each parent in any given occupation at any date will face an independent and stationary environment; its optimization decision (1) is unaffected by decisions of any other household in the economy.

The solution to (1) is the following: $e(y, \tilde{x}) = 1$ iff $\tilde{x} \leq x^*(y)$ where

$$u(y) - u(y - x^*(y)) = B. \tag{2}$$

This results in interim parental welfare $W_y^* \equiv \mathcal{U}_y^* + F(x^*(y))B + V_0$, where

$$\mathcal{U}_y^* = [1 - F(x^*(y))]u(y) + F(x^*(y))\mathsf{E}[u(y - \tilde{x})|\tilde{x} \leq x^*(y)] \tag{3}$$

denotes interim consumption utility before the parent observes $\tilde{x}$.

A **dynamic competitive equilibrium (DCE)** in this economy with an initial skill proportion λ_0 at date 0 consists of a sequence of subsequent skill proportions λ_k, $k = 1, 2, \ldots$ such

[11] Among many possibilities, the parent might care about its offspring's earnings only for the prospective aid received after retirement; or high skilled wages might be subjectively discounted because corresponding work by the child increases geographic distance to the family.

[12] The assumptions could, however, also pertain to a static two-period model as in Jacobs et al. (2012) and other public economics literature (see Section 4.2): the world ends when the child becomes an adult and consumes her entire earnings.

that at each date $k \geq 0$: (i) a fraction λ_k of the adult population is in the skilled occupation; (ii) incomes of adults in occupation $c = 0, 1$ are drawn from the distribution G_c; (iii) every household with income realization y then learns its child's education cost realization $\tilde{x}$, and chooses e to maximize utility (1); (iv) these choices give rise to a proportion λ_{k+1} of the population having a skilled occupation at date $k + 1$.[13]

This definition can be extended to an economy with a *stationary fiscal policy*, in which Y denotes after-tax income levels $y_i \equiv \tilde{y}_i + \tau_i$ that result from market incomes $\tilde{y}_i \in \tilde{Y} = \{\tilde{y}_1, \ldots, \tilde{y}_n\}$ with $\tilde{y}_i < \tilde{y}_{i+1}$, supplemented by progressive net transfers τ_i that satisfy $\tau_{i-1} \geq \tau_i$ and $\tau_i > -\tilde{y}_i$ for all i. A negative net transfer corresponds to an income tax payment. For instance, a universal basic income of $b > 0$ financed by proportional or progressive income taxes $\hat{t}_i = \sum_{j=1}^{i} \alpha_j (\tilde{y}_j - \tilde{y}_{j-1})$ would correspond to $\tau_i = b - \hat{t}_i$ (with marginal tax rates $0 \leq \alpha_i \leq \alpha_{i+1} < 1$ and $\tilde{y}_0 \equiv 0$). Unconditional cash transfers to poor households with pre-tax incomes $\tilde{y}$ below a threshold $\tilde{y}_l$ that are financed by households with $\tilde{y} \geq \tilde{y}_h > \tilde{y}_l$ would amount to $\tau_1 \geq \ldots \geq \tau_l > 0 > \tau_h \geq \ldots \geq \tau_n$, and $\tau_i = 0$ otherwise. Laissez faire obviously corresponds to $\tau_i \equiv 0$ for all i.

In a DCE, successive generations of every household transit between skilled and unskilled occupations according to a time-homogeneous Markov chain, with transition probabilities $F_1^* > F_0^*$ (where $F_c^* \equiv \sum_{i=1}^{n} \pi_{ic} F(x^*(y_i)))$ from occupation c to the skilled occupation. This stochastic process converges to a limiting distribution which is a steady state, where upward mobility flows from the unskilled to the skilled occupation equal downward flows in the opposite direction. Our analysis does *not* presume the economy starts at a steady state; hence the status quo DCE may well involve a skill proportion that increases or decreases over time.

Now consider an income-specific CCT program where a household with income y_i pays a tax of

$$t_i = \epsilon_i \cdot \frac{F(x_i)}{1 - F(x_i)} \tag{4}$$

if it does not invest in education, and receives a subsidy of

$$s_i = \epsilon_i \tag{5}$$

if it does invest; $\epsilon_i > 0$ and $x_i > x^*(y_i)$ are parameters to be chosen. Intuitively, the policy seeks to raise the education threshold for parents with income y_i from $x^*(y_i)$ to x_i, and ϵ_i will represent the scale of the intervention, i.e., the size of the subsidy. If the policy is successful in inducing parents with income y_i to raise their education threshold to x_i, the policy will not affect the public budget surplus. So it is essentially an insurance scheme (breaking even within income class y_i) where parents with less able children ($\tilde{x}$ above the education cost threshold) do not invest and thus enjoy higher parental consumption, will end up paying taxes that finance subsidies to parents that end up consuming less owing to investing in their children's education.

[13] It is evident from the definition that there is a unique DCE corresponding to any initial skill proportion.

Proposition 1 *Consider a DCE involving unconditional taxes/transfers with* $\tau_{j-1} \geq \tau_j$ *and* $\tau_j > -\tilde{y}_j$*, or laissez faire. Take any* i*. There exist* $x_i > x_i^* \equiv x^*(y_i)$ *and* $\epsilon_i > 0$ *such that introducing the corresponding CCT will generate an interim Pareto improvement. At every date: interim welfare of parents with income* y_i *increases strictly, their education cost threshold rises from* x_i^* *to* x_i *(implying the skill proportion rises at every subsequent date), ex post welfare and education decisions of all other income classes remain unchanged, and the public budget surplus improves.*

Proof. Consider any date k, and suppress the notation for k in what follows. Given any $x > x^*(y_i)$ and any $\epsilon \geq 0$, define $x_i(\epsilon, x)$ by the condition

$$u\Big(y_i - \epsilon\frac{F(x)}{1-F(x)}\Big) - u(y_i + \epsilon - x_i(\epsilon, x)) = B. \tag{6}$$

Our conditions on u ensure this is well defined. It is evident that $x_i(0, x) = x^*(y_i)$, and x_i rises in ϵ (holding x fixed) with a slope exceeding 1. Hence there exists $\epsilon_i(x) \in (0, x)$ such that $x_i(\epsilon_i(x), x) = x$, i.e., a parent with income realization y_i will select the threshold x under the CCT corresponding to x and $\epsilon_i = \epsilon_i(x)$. Moreover, $\epsilon_i(x)$ is strictly increasing in x with $\epsilon_i(x^*(y_i)) = 0$.

Next, we claim that we can select $x_i > x^*(y_i)$ such that

$$\frac{\mathsf{E}[u'(y_i + \epsilon_i(x_i) - \tilde{x})|\tilde{x} \leq x_i]}{u'\big(y_i - \epsilon_i(x_i)\frac{F(x_i)}{1-F(x_i)}\big)} > \frac{F(x_i)/(1-F(x_i))}{F(x^*(y_i))/(1-F(x^*(y_i)))}. \tag{7}$$

As $x_i \to x^*(y_i)$, the LHS approaches

$$\frac{\mathsf{E}[u'(y_i - \tilde{x})|\tilde{x} \leq x^*(y_i)]}{u'(y_i)} \tag{8}$$

which strictly exceeds 1 (since $x^*(y_i) > 0$), while the RHS approaches 1. Hence by continuity of all relevant functions, condition (7) holds for some x_i in a right neighborhood of $x^*(y_i)$, thereby establishing the claim.

Suppose the CCT corresponding to x_i and an intermediate scale $\epsilon \in (0, \epsilon_i(x_i))$ is introduced. This would induce a cost threshold $\hat{x}_i(\epsilon, x_i)) \in (x^*(y_i), x_i)$ for parental education decisions, satisfying

$$u\Big(y_i - \epsilon\frac{F(x_i)}{1-F(x_i)}\Big) - u(y_i + \epsilon - \hat{x}_i(\epsilon, x_i))) = B \tag{9}$$

and would generate interim consumption utility

$$\mathcal{U}_i(\epsilon, x_i) = [1 - F(\hat{x}_i(\epsilon, x_i))]u\Big(y_i - \epsilon\frac{F(x_i)}{1-F(x_i)}\Big) + F(\hat{x}_i(\epsilon, x_i))\mathsf{E}[u(y_i + \epsilon - \tilde{x})|\tilde{x} \leq \hat{x}_i(\epsilon, x_i)]. \tag{10}$$

Using the Envelope Theorem, the change in interim welfare of a parent with ex post income y_i from a small rise in the scale ϵ of this CCT equals

$$\frac{\partial \mathcal{U}_i(\epsilon, x_i)}{\partial \epsilon} = F(\hat{x}_i(\epsilon, x_i)))\mathsf{E}[u'(y_i+\epsilon-\tilde{x})|\tilde{x} \leq \hat{x}_i(\epsilon, x_i)))]-[1-F(\hat{x}_i(\epsilon, x_i)))]\frac{F(x_i)}{1-F(x_i)}u'\Big(y_i-\epsilon\frac{F(x_i)}{1-F(x_i)}\Big) \tag{11}$$

which is strictly positive (using $\hat{x}_i(\epsilon, x_i) \in (x^*(y_i), x_i)$, $\epsilon < \epsilon_i(x_i)$ and (7)). This implies that interim welfare of the parent is strictly higher than in status quo when the scale ϵ is set at its maximum value $\epsilon_i(x_i)$. This induces threshold $x_i > x_i^*$ without changing ex post welfare or education decisions at other income levels; nor having a direct effect on the public budget.

The increased proportion of skilled in the population has a beneficial indirect effect on the budget if the status quo involves progressive fiscal policies where some inequality $\tau_{j-1} \geq \tau_j$, $j = 2, \ldots, n$, is strict. Namely, suppose the dynamic sequence of skill proportions in the status quo DCE is $\lambda_0^*, \lambda_1^*, \lambda_2^*, \ldots$ and introduce the intervention at date $k = 0$ (w.l.o.g.). Then the post-intervention investment thresholds are $x(y) = x^*(y)$ for all $y \neq y_i$ and $x(y_i) = x_i > x^*(y_i)$. This strictly increases the transition probability from occupation c to the skilled occupation: $F_c = \sum_{i=1}^n \pi_{ic} F(x(y_i)) > F_c^* = \sum_{i=1}^n \pi_{ic} F(x^*(y_i))$; and induces a skill proportion $\lambda_1 > \lambda_1^*$. Moreover, $\lambda_k > \lambda_k^*$ implies

$$\lambda_{k+1} = \lambda_k^* F_1 + (1-\lambda_k^*)F_0 + (\lambda_k - \lambda_k^*)[F_1 - F_0] > \lambda_k^* F_1^* + (1-\lambda_k^*)F_0^* = \lambda_{k+1}^* \tag{12}$$

using $F_1 > F_0$. So skill proportions in the population stay above their status quo benchmarks in every period of the CCT intervention. Stochastic dominance of the income distribution among the skilled and $\tau_{j-1} > \tau_j$ for some j then imply that public expenditure $\lambda^* \bar{\tau}_1 + (1-\lambda^*)\bar{\tau}_0$ with $\bar{\tau}_c = \sum_{i=1}^n \pi_{ic}\tau_i$ decreases:

$$\frac{\partial[\lambda\bar{\tau}_1 + (1-\lambda)\bar{\tau}_0]}{\partial \lambda} = \bar{\tau}_1 - \bar{\tau}_0 < 0. \tag{13}$$

■

The idea behind the interim Pareto improvement compared to laissez faire or any unconditional transfers is the following. The CCT induces greater educational investment, as education is being subsidized. And at a small scale, the scheme offers a first order improvement in consumption smoothing. This is illustrated in Figure 1: consumption of a parent with income y_i generally varies non-monotonically in education cost realization $\tilde{x}$. In status quo (bold line), only parents whose child can costlessly be educated ($\tilde{x} = 0$) and non-investing parents consume their full income y_i; all those with $\tilde{x} \in (0, x_i^*)$ consume less. The CCT reduces this variation (dotted line): parents who would have invested in status quo do still invest and enjoy a consumption increase of s_i; parents who invest neither in status quo nor with the CCT see their consumption lowered by $t_i = s_i \cdot F(x_i)/[1 - F(x_i)]$ in return. A small share $[F(x_i) - F(x^*)]$ of parents 'avoid' the latter: they reduce own consumption further in favor of obtaining paternalistic benefit B. If these

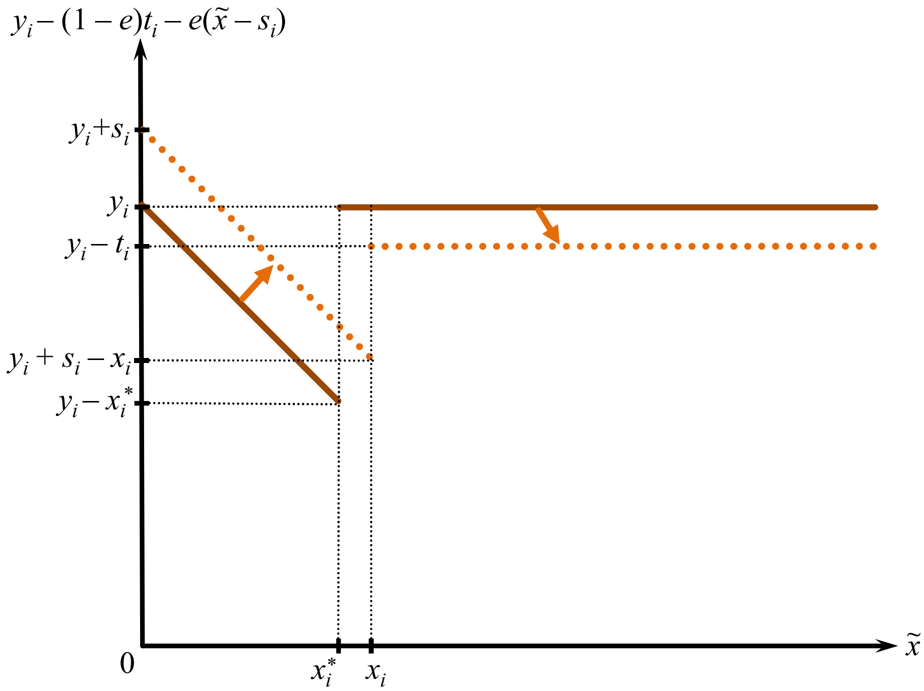

Figure 1: Effects of CCT program on parental consumption

parents did not change their behavior, the consumption distribution with CCT would constitute a mean-preserving compression and second-order stochastically dominate the status quo (provided x_i is not too far above x_i^*, as ensured by (7)). They do, however, change their behavior and switch to investing in their children's education, and become better off as a result. So interim utility of parents with income y_i increases by an even greater amount than that implied by the respective mean-preserving compression of consumption.

Note that the scheme can be offered independently for one, some, or all incomes y_i. So the efficiency improvement is orthogonal to effects on inequality of welfare at different income levels. If the government wants to reduce inequality, it can introduce it only for low incomes. Or it can offer it for rich and poor, so that all gain equally. It can also be restricted to higher incomes, should the government want to raise inequality.

Observe also that the result applies irrespective of whether there is underinvestment or over-investment in education in the conventional sense. The market failure is in insurance, and that is being corrected by the CCT which is a form of consumption insurance. This is a novel insight into the welfare role of CCTs, operating partly via consumption smoothing, and partly via enhanced investment. The former ensures all parents are better off (at the interim stage). The latter guarantees that future generations benefit, even if the CCT intervention should be temporary.[14] To

[14] If CCTs are phased out at some date k', the skill shares $\lambda_{k'}, \lambda_{k'+1}, \ldots$ converge to status quo levels $\lambda^*_{k'}, \lambda^*_{k'+1}, \ldots$

ensure the improvement in consumption smoothing, the education subsidy is financed in a specific way: by taxing incomes of those in the same income class. The policy substitutes for missing insurance markets in the status quo (ostensibly owing to adverse selection or other transaction cost / enforcement problems), and functions like an insurance policy. There is a separate policy for each income level, so education subsidies to parents at any income level are funded by others with the same income who do not invest in education, owing to low ability realizations of their children. This ensures that the program does not result in any redistribution between the poor and rich, beyond the fiscal policy that may apply in status quo (such as UBI financed by a proportional income tax, progressive UCTs, or combinations).

Of course if the government wants to additionally redistribute in favor of the poor, the subsidies could be restricted only to the poor, and funded by taxes paid by the rich. In practice, governments often have such a redistributive goal and do fund welfare benefits in this way. But such interventions are not Pareto improving. What Proposition 1 shows is that if the government wants to avoid (additional) redistribution, it is possible in principle to design a CCT that succeeds in doing so and to generate welfare improvements for both rich and poor.

We show below that these results are robust to different extensions of the model.

2.2 Non-Paternalistic Altruism

To demonstrate that the possibility of a Pareto improvement is not limited to contexts where the perceived benefits of education are paternalistic, let us move to a more refined form of parental altruism, where households are dynasties and parents are non-paternalistic *à la* Barro-Becker. In this specification, parents internalize the utility consequences of education decisions for their offspring, albeit scaled down by a discount factor $\delta \in (0, 1)$. If δ is close to 1, the parent-child externality tends to vanish. One might guess that market failure in education is now less likely to occur. But as we have argued above, the key market failure is in insurance, and that is unaffected by the magnitude of δ. In fact we will show below in Proposition 3 that δ large enough ensures that the CCT program generates an ex post Pareto improvement, so the efficiency improvement is if anything enhanced when δ is high.

We continue to assume a finite support Y of the after-tax income distribution (resulting either from stationary fiscal policies that do not condition on education investment, or laissez faire), a well-behaved consumption utility function u, and that the income distribution given $c = 1$ first order stochastically dominates the distribution for $c = 0$ strongly. The definition of DCE is modified in the obvious way. To illustrate, the DCE in the status quo is characterized as follows. Let x_i^* denote the investment threshold for a parent with income y_i in the status quo DCE. The

from above, noting that $F_c \geq F_c^*$ is sufficient for concluding (12).

threshold is uniquely determined by[15]

$$u(y_i) - u(y_i - x_i^*) = B^* \equiv \delta[W_1^* - W_0^*] \tag{14}$$

where

$$W_c^* \equiv \bar{\mathcal{U}}_c^* + F_c^* B^* + \delta W_0^* \tag{15}$$

is the expected dynastic utility of a household in occupation c based on its expected consumption utility

$$\bar{\mathcal{U}}_c^* = \sum_{i=1}^{n} \pi_{ic} \mathcal{U}_i^* \tag{16}$$

where

$$\mathcal{U}_i^* = [1 - F(x_i^*)]u(y_i) + F(x_i^*)\mathsf{E}[u(y_i - \tilde{x})|\tilde{x} \leq x_i^*] \tag{17}$$

is the interim consumption utility of income type i (with income realization y_i), and we abbreviate

$$F_c^* = \sum_{i=1}^{n} \pi_{ic} F(x_i^*). \tag{18}$$

We first verify that Proposition 1 extends to dynastic Barro-Becker preferences. The respective policy involves a new education threshold x_i, expected consumption utility $\mathcal{U}_i$ for income type i, $\bar{\mathcal{U}}_c$ for occupation c, and investment probability F_c for occupation c satisfying analogous conditions:

$$u(y_i - t_i) - u(y_i + s_i - x_i) = B \equiv \delta[W_1 - W_0] \tag{19}$$

$$W_c = \bar{\mathcal{U}}_c + F_c B + \delta W_0 \tag{20}$$

$$\bar{\mathcal{U}}_c = \sum_{i=1}^{n} \pi_{ic} \mathcal{U}_i \tag{21}$$

$$\mathcal{U}_i = [1 - F(x_i)]u(y_i - t_i) + F(x_i)\mathsf{E}[u(y_i + s_i - \tilde{x})|\tilde{x} \leq x_i] \tag{22}$$

$$F_c = \sum_{i=1}^{n} \pi_{ic} F(x_i). \tag{23}$$

With Barro-Becker preferences, the post-CCT benefits of investing in education, B, are jointly determined by *all* new investment thresholds x_i via (20)–(23). This renders an intervention that affects welfare of just one income type i infeasible. But it is still possible to achieve an interim Pareto improvement – for instance, by introducing CCTs simultaneously for all income classes.

Proposition 2 *Consider a DCE involving unconditional taxes/transfers with $\tau_{j-1} \geq \tau_j$ and $\tau_j > -\tilde{y}_j$, or laissez faire. There exist $t_i = \epsilon_i \frac{F(\bar{x}_i)}{1-F(\bar{x}_i)}$ and $s_i = \epsilon_i$ for some $\epsilon_i > 0$ and $\bar{x}_i > x_i^*$*

[15] A household's value function can be bounded by $u(y_1)/(1-\delta)$ and $u(y_n)/(1-\delta)$ from below and above, given $\delta \in (0, 1)$. Blackwell's sufficient conditions then hold, thereby guaranteeing a unique solution.

for all $i = 1, \ldots, n$ such that introducing the corresponding CCTs will generate an interim Pareto improvement. At every subsequent date, interim welfare of parents in every income class increases by the same positive amount, the education cost threshold for income y_i households rises from x_i^* to some $x_i \in (x_i^*, \bar{x}_i)$, and the public budget surplus improves.

Proof. As in Proposition 1, select $\bar{x}_i > x_i^*$ such that

$$\frac{\mathsf{E}[u'(y_i + \bar{\epsilon}_i - \tilde{x})|\tilde{x} \leq \bar{x}_i]}{u'(y_i - \bar{\epsilon}_i \frac{F(\bar{x}_i)}{1-F(\bar{x}_i)})} > \frac{F(\bar{x}_i)/[1 - F(\bar{x}_i)]}{F(x_i^*)/[1 - F(x_i^*)]} \tag{24}$$

where $\bar{\epsilon}_i$ is defined by

$$u\big(y_i - \bar{\epsilon}_i \frac{F(\bar{x}_i)}{1 - F(\bar{x}_i)}\big) - u(y_i + \bar{\epsilon}_i - \bar{x}_i) = B^*. \tag{25}$$

Choose $\epsilon_i \in [0, \bar{\epsilon}_i]$ and CCT with $s_i = \epsilon_i, t_i = \epsilon_i \frac{F(\bar{x}_i)}{1-F(\bar{x}_i)}$. Let $x_i = x_i(\epsilon_i)$ denote the corresponding investment threshold for type i with the same education return B^* as in the status quo:

$$u\Big(y_i - \epsilon_i \frac{F(\bar{x}_i)}{1 - F(\bar{x}_i)}\Big) - u(y_i + \epsilon_i - x_i) = B^*. \tag{26}$$

Previous arguments imply $x_i \in [x_i^*, \bar{x}_i]$ with $x_i = x_i^*$ if $\epsilon_i = 0$ at the status quo, and $x_i = \bar{x}_i$ if $\epsilon_i = \bar{\epsilon}_i$. By construction, the CCT generates the same budget surplus as the status quo in the latter case, and improves the budget surplus if ϵ_i, x_i are respectively smaller than $\bar{\epsilon}_i, \bar{x}_i$. Interim consumption utility

$$\mathcal{U}_i(\epsilon_i) \equiv [1 - F(x_i)]u\Big(y_i - \epsilon_i \frac{F(\bar{x}_i)}{1 - F(\bar{x}_i)}\Big) + F(x_i)\mathsf{E}[u(y_i + \epsilon_i - \tilde{x})|\tilde{x} \leq x_i] \tag{27}$$

is strictly increasing in ϵ_i over this range. It remains to show that we can select $\epsilon_i \in (0, \bar{\epsilon}_i)$ for all $i = 1, \ldots, n$ such that the gross return from education is unchanged compared to the status quo:

$$B^* = \frac{\delta \sum_{i=1}^n [\pi_{i1} - \pi_{i0}]\mathcal{U}_i(\epsilon_i)}{1 - \delta \sum_{i=1}^n [\pi_{i1} - \pi_{i0}]F(x_i)}. \tag{28}$$

This is because (28) is a necessary and sufficient condition for investment thresholds x_i to constitute a DCE following the chosen CCT.

This condition can also be written as

$$\delta \sum_{i=1}^n (\pi_{i1} - \pi_{i0})[\mathcal{U}_i(\epsilon_i) + B^* F(x_i)] = B^*, \tag{29}$$

and it holds at the status quo:

$$\delta \sum_{i=1}^n (\pi_{i1} - \pi_{i0})[\mathcal{U}_i(0) + B^* F(x_i^*)] = B^*. \tag{30}$$

Hence (28) reduces to

$$\sum_{i=1}^{n}(\pi_{i1}-\pi_{i0})\psi_i(\epsilon_i)=0 \tag{31}$$

where

$$\psi_i(\epsilon_i)\equiv[\mathcal{U}_i(\epsilon_i)+B^*F(x_i)]-[\mathcal{U}_i(0)+B^*F(x_i^*)] \tag{32}$$

is a measure of the relative interim welfare improvement for income type i, as the actual improvement is $\psi_i(\epsilon_i)+\delta[W_0-W_0^*]$ and $\delta[W_0-W_0^*]$ does not vary with i. $\psi_i(\epsilon_i)$ is strictly increasing in ϵ_i over the range $[0,\bar{\epsilon}_i]$. Hence there exists a small $\eta>0$ such that $\epsilon_i=\psi_i^{-1}(\eta)\in(0,\bar{\epsilon}_i)$ for all $i=1,\ldots,n$. It follows from $\sum_{i=1}^{n}\pi_{ic}=1$ that (31) and hence (28) hold with these choices for ϵ_i. The public budget surplus, already improved because $\epsilon_i\in(0,\bar{\epsilon}_i)$, is additionally aided if the status quo involves progressive fiscal policies where $\tau_{j-1}\geq\tau_j$ holds strictly for some j, by the same argument as in Proposition 1.

■

In this construction, the CCT intervention generates an equal welfare improvement for all income classes. It is possible to modify it to ensure that lower income groups attain a higher welfare improvement, with an exception at the very top. Specifically, we can choose $\epsilon_i\in[0,\bar{\epsilon}_i], i=1,\ldots,n-1$ such that $0<\psi_i(\epsilon_i)<\psi_{i-1}(\epsilon_{i-1})$ for all $i=n-1,n-2,\ldots,2$. Now (31) requires

$$\begin{aligned}(\pi_{n1}-\pi_{n0})\psi_n(\epsilon_n) &= -\sum_{i=1}^{n-1}(\pi_{i1}-\pi_{i0})\psi_i(\epsilon_i)\\ &= (\pi_{n1}-\pi_{n0})\psi_{n-1}(\epsilon_{n-1})+\sum_{i=1}^{n-2}[G_{i1}-G_{i0}][\psi_{i+1}(\epsilon_{i+1})-\psi_i(\epsilon_i)]. \qquad (33)\end{aligned}$$

Since $\psi_i(\epsilon_i)$ is decreasing in i by construction, the strong first order stochastic dominance property implies that $(\pi_{n1}-\pi_{n0})$ and the RHS of (33) are both positive. Hence the required value of $\psi_n(\epsilon_n)$ is positive. If $\epsilon_i, i=1,\ldots,n-1$, are chosen sufficiently close to 0, this required value is close to 0. Then there exists $\epsilon_n\in(0,\bar{\epsilon}_n)$ such that (33) holds.

However, it is also clear that this kind of CCT cannot be designed to generate redistribution across the entire income scale, as that would require ψ_i to be decreasing in i throughout, and stochastic dominance would then imply that the sign of the LHS of (31) is negative. Conversely it cannot be designed to be throughout regressive. But as shown, it is possible to construct it to ensure that every income class attains the same welfare improvement $\eta>0$. This improvement η can be chosen by the policymaker from an interval $(0,\hat{\eta})$, where $\hat{\eta}>0$ is determined by the status quo DCE and economic fundamentals. Among the latter, the discount factor δ represents the extent to which a parent internalizes the utility of its child and succeeding generations of offspring. We prove in the appendix that if δ is large enough, the scheme results in an ex post Pareto improvement, since parents' valuation of the benefit of increased education among their descendants outweighs their own tax burden even if they themselves do not invest and avail of the subsidy.

Proposition 3 *Let a collection of economies with identical consumption utility function u and probability distributions F, G_0 and G_1, but different parental discount factors $\delta \in (0,1)$ be given. For each corresponding DCE, consider CCTs $\{t_{\delta,i}(\eta), s_{\delta,i}(\eta)\}_{i=1,\ldots,n}$ that induce an interim Pareto improvement according to Proposition 2. Then there exist $\underline{\delta} \in (0,1)$ and $\bar{\eta} > 0$ such that for any $\eta \in (0,\bar{\eta})$ and $\delta \in (\underline{\delta},1)$ the intervention also generates an ex post Pareto improvement, i.e., the welfare of every agent in the economy at every subsequent date is higher, irrespective of income or child's learning ability.*

3 Extensions

We provide an informal discussion of how the preceding results are modified when the model is extended in different directions.

3.1 Endogenous Labor Supply

A first extension of the baseline model allows labor supply to vary. Interpret $y \in Y$ as the wage rate available to a given household, and let households choose how many hours of labor they supply, together with the binary decision whether to invest in education or not. Consider for simplicity the case of paternalistic altruism. Then each household facing wage rate y and education cost $\tilde{x}$ selects $e \in \{0,1\}$ and $l \geq 0$ to maximize

$$u\big(ly - e\tilde{x}\big) - d(l) + eV_1 + (1-e)V_0 \tag{34}$$

for strictly increasing and convex disutility of labor d given $V_c \equiv \sum_{i=1}^{n} \pi_{ic} V(y_i)$. Here $V(y)$ denotes the benefit perceived by the parent from the child's future when the latter would be able to earn a wage rate y, and we naturally assume that V is strictly increasing.

The optimal investment strategy $e(y,\tilde{x})$ in this case is of the same threshold form as in the baseline model. Namely, if we define

$$v(y_i,\tilde{x},e) \equiv \max_{l_i} \Big[u\big(l_i y_i - e\tilde{x}\big) - d(l_i)\Big] \tag{35}$$

then a parent with wage rate y_i who faces education cost $\tilde{x}$ will invest iff $\tilde{x} < x_i$, where threshold x_i is defined by

$$v(y_i,x_i,0) - v(y_i,x_i,1) = V_1 - V_0. \tag{36}$$

Parents with wage rate y_i and cost $\tilde{x} = 0$ or cost $\tilde{x} \geq x_i$ have identical (indirect) utilities of consumption $v(y_i,0,1) = v(y_i,\tilde{x},0)$, while those with cost $\tilde{x} \in (0,x_i)$ consume less. In particular, from (35) and the Envelope Theorem, we have

$$\frac{\partial v(y_i, \tilde{x}, e(y, \tilde{x}))}{\partial \tilde{x}} = -u'\big(l(y_i, \tilde{x})y_i - \tilde{x}\big) < 0 \text{ for each } \tilde{x} \in (0, x_i). \tag{37}$$

It follows that consumption utilities $v(y_i, \tilde{x}, e(y_i, \tilde{x}))$ are decreasing on $[0, x_i)$, jump back to $v(y_i, 0, 1)$, and then stay at this level. That is, they exhibit a non-monotonic pattern with respect to education cost $\tilde{x}$ just like in the baseline model. A variation of the baseline policy intervention can therefore be applied in order to create an interim Pareto improvement.

3.2 Continuous Education Choices

What if educational investments can be varied continuously, rather than being indivisible? CCTs are designed to subsidize only variations in education on the extensive margin rather than the intensive margin – i.e., parents are eligible for the subsidy provided their children are enrolled in school; the size of the subsidy does not vary with the extent of educational achievement.[16] It is presumably for this reason that they are typically offered for enrollment of children in secondary schooling in countries with significant dropout rates in secondary but not primary schooling. So we consider an extension of our model consistent with non-universal enrollment in the status quo situation, and show that the CCT can continue to be designed on the basis of enrollment decisions.

Let the extent of education be described by a compact interval $E \equiv [0, \bar{e}]$ of the real line. Enrollment corresponds to a positive choice of e. Conditional on education $e \in E$, the distribution of earnings is given by a cdf G_e, where $e' > e$ implies $G_{e'}$ strongly first order stochastically dominates G_e. To simplify the exposition, we assume that $\tilde{V}(e) \equiv \int_Y V(y)dG_e(y)$ is a concave C^2 function with $0 < \frac{\partial \tilde{V}(e)}{\partial e} < \infty$ for all $e \in E$. Next, let $I(e; \tilde{x})$ 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 $\tilde{x}$. 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 $I(0; \tilde{x}) = 0$ for all $\tilde{x}$, and for any given $e \geq 0$ the marginal cost $\frac{\partial I(e;\tilde{x})}{\partial e}$ is increasing in $\tilde{x}$, zero at $\tilde{x} = 0$ and approaches ∞ as $\tilde{x} \to \infty$.

A parent with income y and a child with learning cost $\tilde{x}$ then solves

$$\max_{0 \leq e' \leq \bar{e}} \big[u(y - I(e'; \tilde{x})) + \tilde{V}(e')\big]. \tag{38}$$

Let the corresponding policy function be $e'(y; \tilde{x})$.

[16] Of course the extent of enrollment as measured by proportion of classes attended can also vary continuously. We refer to enrollment as the achievement of a minimum target for the proportion of classes attended, as commonly required in most CCTs as a precondition for subsidy eligibility. We implicitly assume that the basis for setting this threshold is that it refers to a minimum required attendance for the student to receive a passing grade.

Under these assumptions $x^*(y) > 0$ is well-defined as the solution for x in the equation $u'(y)\frac{\partial I(0;x)}{\partial e'} = \frac{\partial \tilde{V}(0)}{\partial e'}$, and the optimal policy function takes the form $e'(y;\tilde{x}) = 0$ if $\tilde{x} \geq x^*(y)$ and positive otherwise. 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 $x^*(y)$. These 'non-investors' consume their entire earnings y – just like those parents with the same income y whose children have learning cost parameter $\tilde{x} = 0$. 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. Parents whose children do not enroll therefore consume more than parents earning the same whose children do enroll. The educational subsidy funded by the income tax in this group then redistributes consumption away from those consuming high amounts to those consuming less. Since these consumption variations arise from the 'ability lottery' of their children, the policy increases interim expected utilities of each income class.

3.3 Financial Bequests

In the baseline model 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 or her 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.

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).[17] To simplify the exposition, assume incomes are non-stochastic and depend only on occupation: w_c now denotes wage earnings in occupation c with $w_1 > w_0$. The idiosyncratic education cost needed for working in the skilled occupation is again $\tilde{x}$, that for the unskilled occupation equals zero. Parental altruism is paternalistic, where a parent with lifetime wealth W and education cost $\tilde{x}$ chooses financial bequest $b \geq 0$ and education investment $e \in \{0, 1\}$ to maximize $u(W - b - e\tilde{x}) + \delta V(W')$ where V is a strictly increasing and strictly concave function of the child's future wealth W' which equals $(1 + r)b + ew_1 + (1 - e)w_0$.

[17]This corresponds 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.

The details of the analysis are provided in the working paper version (Mookherjee and Napel 2019). We highlight here the solutions for two ranges of parental wealths.

Case A. *W sufficiently large:* For W large enough, the parent will always make a financial bequest that is possibly supplemented by an education investment. The sum of expenditures on education and financial bequests is lowest – and parental consumption highest – for the most talented education cost type $\tilde{x} = 0$. From there, total spending for the child increases in $\tilde{x}$ until some threshold x^*_W, and then it becomes optimal to transfer a constant amount of wealth purely via financial bequests.

Case B. *W sufficiently small:* Suppose $W = w_0$, $\delta(1+r) \leq 1$ and $V \equiv u$. Then the parent never makes a financial bequest. If however the child learning cost $\tilde{x}$ is below a positive threshold level $x^*_{w_0}$, the parent will invest in education.

Parents in case B behave exactly as described in previous sections and their consumption varies with cost $\tilde{x}$ exactly as in Figure 1. So our previous arguments continue to apply for poor households in case B, who never make any financial bequests. Offering educational subsidies for them, funded by corresponding income taxes, would be interim Pareto improving. The model of Abbott et al. (2019), calibrated to fit NLSY 1997 data, suggests that case A applies to the top 5% of the US population and case B applies to the bottom third. We speculate that the respective share of population described by case B is even bigger in most developing countries.

3.4 General equilibrium wage effects; non-stationary fiscal policy

Finally consider a setting where skilled and unskilled wages depend on the skill composition in the population. Then increases in skill composition induced by a CCT would lower the skill premium in wages: skilled wages would fall while unskilled wages would rise. This would lower educational investment incentives. Moreover, the outcome would lower the welfare of skilled households, so would not be Pareto improving. This necessitates further modification to the design of the CCT. In particular, it needs to be accompanied by an offsetting regressive change in fiscal policy which ‘neutralizes’ these GE effects, lowering taxes on high incomes and raising them on low incomes, so as to keep inter-occupation wage and welfare differences the same as in the status quo. In the working paper version (Mookherjee and Napel 2019), we consider the case of two occupations, non-stochastic income and non-paternalistic utility, and show that the CCT design can indeed be modified in this manner to ensure that a Pareto improvement results in which welfare of skilled and unskilled households within each generation rise by exactly the same extent. The extension also includes the case where the status quo involves a non-stationary fiscal policy, and the government is required to balance its budget at every date.

4 Related Literature

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

4.1 Development and Occupational Choice

The closest connection is with the literature on occupational choice with credit market imperfections.[18] 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. Their analysis is subject to a number of problems which we overcome in the current paper: by focusing on the long run they ignore impacts in the short run and the transition to a new steady state. They ignore ability heterogeneity and do not investigate the possibility of efficiency improvements resulting from CCTs.

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 altruism. However, welfare effects of fiscal policy were not addressed, so this paper is a natural complement of the earlier one. 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.[19] 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 in-

[18] See, e.g., Banerjee and Newman (1993), Galor and Zeira (1993), Ljungqvist (1993), Freeman (1996), Aghion and Bolton (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).

[19] 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.

strument: public provision of education, 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 Pareto improving interventions exist. However, they focus on a different policy instrument: public provision of education, where children receiving a public schooling are required to pay back to the government when they become adults. The nature of the efficiency improvement in that paper is also different, consisting of reducing misallocation of education between children in rich and poor families, while leaving unchanged the aggregate proportion educated. They additionally show the result is robust when education signals unobserved productivity of workers to employers.

4.2 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 examined implications of redistributive tax distortions for education subsidies. For instance, Bovenberg and Jacobs (2005) argue in a static context 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.

4.3 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.[20] 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 Proposition 3.

Versions of these models have been calibrated to fit data of real economies in order to evaluate

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

the welfare and macroeconomic effects of various fiscal policies in numerical simulations.[21] These studies rely on specific functional forms for technology and preferences, 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.

5 Concluding Observations

We have provided a theoretical argument 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 CCTs are funded by income taxes imposed on the same income / occupational class, thereby avoiding redistribution across income groups. The results hold irrespective of specific assumptions on preferences or technology, initial conditions, general equilibrium effects, and incorporate short as well as long run effects. We have argued the results also apply irrespective of labor supply elasticity or investment divisibility. When parents have the additional option of leaving financial bequests to their children, subsidizing education is still desirable for parents in income classes who do not supplement education investments with financial bequests, which seems plausible for most poor households.

Normative discussions of CCTs usually argue that inefficient underinvestment in human capital is the relevant market failure and the key rationale for conditional transfers. Addressing this can, however, involve tension with other rationales for CCTs, e.g., using them to achieve pro-poor redistribution (see, e.g., Das et al. 2004). Our analysis identifies a market failure in education that differs from what is widely supposed among applied economists and policy-makers. It is unrelated to any notion of underinvestment or biased valuation of children's education by their parents. Instead it relates to a failure of insurance markets, manifested in family consumptions that vary inefficiently with respect to ability realizations of children. This is a dimension in the welfare evaluation of CCTs which has been overlooked so far both in theoretical or empirical research. While the design we propose addresses directly an insurance rather than underinvestment problem, it nevertheless results in increased investments, per capita income and upward intergenerational mobility.

Comparatively little attention has also been devoted so far on how a CCT program should

[21] See Heathcote (2005), Bohacek and Kapicka (2008), Berriel and Zilberman (2011), Céspedes (2014), Findeisen and Sachs (2016), Peruffo and Ferreira (2017), Abbott, Gallipoli, Meghir, and Violante (2019), and Colas, Findeisen, and Sachs (2021).

be financed. In practice it is typically financed by general income taxes, collected primarily from better-off households who do not qualify for the pertinent education subsidies. A CCT program is then explicitly redistributive, involving trade-offs between efficiency and distribution objectives, and potentially vulnerable to political opposition from better off households that end up paying for the program. We have argued there is an alternative funding mechanism for CCTs in which efficiency can be enhanced without adverse redistributive impacts, which could avoid such political opposition. So our paper may help promote consideration of a new form of CCT policy proposal.

Our analysis also casts a different perspective on arguments in the debate of universal basic income as an alternative to CCT schemes, by showing that any UBI scheme would be Pareto dominated by a CCT. This addresses two common criticisms of CCTs concerning narrow coverage and greater paternalism that have arisen in these debates. The 'narrow coverage' concern is an articulation of an ex post perspective, where some households end up ineligible for the subsidy. A similar concern could be raised about *any* insurance program, where some agents (often the vast majority) end up worse off ex post as a result of having paid premiums but not received any payout on account of not having experienced an accident. This indicates the need to adopt an ex ante or interim perspective instead. And most concerns of paternalism are based on a static riskless perspective where there are no investment or insurance considerations at play. It therefore seems that the only credible argument against CCTs is that transfer conditionalities entail higher costs of monitoring and enforcement. But given CCT adoption and experience of many middle and low income developing countries, this does not seem to be very widely applicable. At any rate, if weak state capacity happens to be a binding constraint, such countries should aspire to adopt CCTs as they enhance their capacity over time.

Appendix – Proof of Proposition 3

It is sufficient to show that, for any realized income y_i, parents who do not invest but pay tax $t_{\delta,i}(\eta)$ to finance the respective CCT, are still rendered better off when η is small and δ is large enough. The parents who invest only under the CCT reveal to be even better off. Those who already invested in status quo are rendered better off by subsidy $s_{\delta,i}(\eta)$ and higher dynastic welfare $\delta(B + W_0)$. We show that W_0 actually increases at an unbounded rate, as $\delta \to 1$, while non-investors' losses of consumption utility are bounded.

Let us indicate variables and policy parameters that vary in δ with a corresponding subscript. Note that η fixes $\epsilon_{\delta,i} = \psi_{\delta,i}^{-1}(\eta)$ and investment thresholds $x_{\delta,i}(\epsilon_{\delta,i})$, and implicitly induces bounds for $\bar{x}_{\delta,i}$, which scales taxation and associated budget gains. Strong stochastic dominance of the skilled income distribution G_1 ensures that expected welfare of skilled households is strictly higher than that of unskilled households, independently of δ. This bounds the benefits of investing in education, B_δ^*, away from zero. Since u is increasing, thresholds $x_{\delta,i}^*$ are also bounded away from zero. Moreover, our conditions on u ensure $x_{\delta,i}^* < y_i$.

Raising the scale η of the CCT policy from status quo $\eta = 0$ lowers a non-investing parent's consumption utility at a rate of

$$\frac{\partial}{\partial \eta}\Big[u(y_i) - u\Big(y_i - \psi_{\delta,i}^{-1}(\eta)\frac{F(\bar{x}_{\delta,i})}{1-F(\bar{x}_{\delta,i})}\Big)\Big]\Bigg|_{\eta=0} = u'(y_i)\cdot\frac{F(x_{\delta,i}^*)}{1-F(x_{\delta,i}^*)}\cdot\frac{\partial \psi_{\delta,i}^{-1}(0)}{\partial \eta}. \tag{39}$$

The first two factors are bounded, respectively, by $u'(y_1)$ and $F(y_n)/[1-F(y_n)]$, noting that $\bar{x}_{\delta,i}$ can be chosen arbitrarily close to $x_{\delta,i}^*$ as $\eta \to 0$. To see that also $\frac{\partial \psi_{\delta,i}^{-1}(0)}{\partial \eta}$ is bounded as $\delta \to 1$, recall that

$$\psi_{\delta,i}(\epsilon_i) = [\mathcal{U}_{\delta,i}(\epsilon_i) + B_\delta^* F(x_{\delta,i}(\epsilon_i))] - [\mathcal{U}_{\delta,i}^* + B_\delta^* F(x_{\delta,i}^*)] \tag{40}$$

and

$$\mathcal{U}_{\delta,i}(\epsilon_i) = [1 - F(x_{\delta,i}(\epsilon_i))]u\Big(y_i - \epsilon_i\frac{F(\bar{x}_{\delta,i})}{1-F(\bar{x}_{\delta,i})}\Big) + F(x_{\delta,i}(\epsilon_i))\mathsf{E}[u(y_i + \epsilon_i - \tilde{x})|\tilde{x} \le x_{\delta,i}(\epsilon_i)] \tag{41}$$

with $\psi_{\delta,i}(0) = 0 = \psi_{\delta,i}^{-1}(0)$. So, using the Envelope Theorem,

$$\begin{aligned}
\frac{\partial \psi_{\delta,i}^{-1}(0)}{\partial \eta} &= \left[\frac{\partial \psi_{\delta,i}(0)}{\partial \epsilon_i}\right]^{-1} = \left[\frac{\partial \mathcal{U}_{\delta,i}(0)}{\partial \epsilon_i} + B_\delta^* f(x_{\delta,i}^*)\cdot\frac{\partial x_{\delta,i}(0)}{\partial \epsilon_i}\right]^{-1} \\
&= \Big[F(x_{\delta,i}^*)\mathsf{E}[u'(y_i - \tilde{x})|\tilde{x} \le x_{\delta,i}^*] - [1 - F(x_{\delta,i}^*)]u'(y_i)\frac{F(x_{\delta,i}^*)}{1-F(x_{\delta,i}^*)}\Big]^{-1} \\
&= \frac{1}{F(x_{\delta,i}^*)}\Big[\mathsf{E}[u'(y_i - \tilde{x})|\tilde{x} \le x_{\delta,i}^*] - u'(y_i)\Big]^{-1} < L
\end{aligned} \tag{42}$$

for some positive constant L, given that $x_{\delta,i}^*$ is bounded away from zero.

The increase in the dynastic component of a non-investing parent's welfare, $\delta[W_{\delta,0} - W^*_{\delta,0}]$, is such that for every $M < \infty$ there exist $\bar{\delta} \in (0,1)$ so that for all $\delta \in (\bar{\delta}, 1)$

$$\begin{aligned}\frac{\partial}{\partial \eta}\left\{\delta[W_{\delta,0}(\eta) - W^*_{\delta,0}]\right\} &= \frac{\partial}{\partial \eta}\left\{\frac{1}{1-\delta}\Big[\bar{\mathcal{U}}_{\delta,0}(\psi^{-1}_{\delta,i}(\eta)) + F_{\delta,0}(\psi^{-1}_{\delta,i}(\eta))B^*_{\delta})\Big]\right\} \\ &= \frac{\partial}{\partial \eta}\left\{\frac{1}{1-\delta}\sum_{i=1}^{n}\pi_{i0}\cdot\psi_{\delta,i}(\psi^{-1}_{\delta,i}(\eta))\right\} = \frac{1}{1-\delta} > M. \end{aligned} \tag{43}$$

Combining (42) and (43), we can conclude that the total welfare change of non-investing parents with income y_i satisfies

$$\frac{\partial}{\partial \eta}\left[\left\{u\Big(y_i - \psi^{-1}_{\delta,i}(\eta)\frac{F(\bar{x}_{\delta,i})}{1-F(\bar{x}_{\delta,i})}\Big) + \delta W_{\delta,0}(\eta)\right\} - \left\{u(y_i) - \delta W^*_{\delta,0}\right\}\right]\Bigg|_{\eta=0} \geq m \tag{44}$$

for all $\delta \in (\bar{\delta}, 1)$ for some $m > 0$. We can therefore choose $\bar{\eta} > 0$ such that each households's ex post welfare change is positive for any $\eta \in (0, \bar{\eta})$ for every $\delta \in (\bar{\delta}, 1)$.

■

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2019). Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy 127*(6), 2569–2624.

Aghion, P. and P. Bolton (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., J. Greenwood, and A. Seshadri (2002). Efficient investment in children. *Journal of Economic Theory 102*(2), 290–321.

Banerjee, A. V., R. Hanna, G. E. Kreindler, and B. A. Olken (2017). Debunking the stereotype of the lazy welfare recipient: Evidence from cash transfer programs. *World Bank Research Observer 32*(2), 155–184.

Banerjee, A. V. and A. Newman (1993). Occupational choice and the process of development. *Journal of Political Economy 101*(2), 274–298.

Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy 87*(6), 1153–1189.

Behrman, J. R., S. W. Parker, and P. E. Todd (2011). Do conditional cash transfers for schooling generate lasting benefits? A five-year followup of Progresa/Oportunidades. *Journal of Human Resources 46*(1), 93–122.

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 E. Zilberman (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 M. Kapicka (2008). Optimal human capital policies. *Journal of Monetary Economics 55*(1), 1–16.

Bovenberg, A. L. and B. Jacobs (2005). Redistribution and education subsidies are Siamese twins. *Journal of Public Economics 89*(11-12), 2005–2035.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127-1160.

Céspedes, N. (2014). General equilibrium analysis of conditional cash transfers. Working Paper No. 2014-25, Peruvian Economic Association.

Colas, M., Findeisen, S., and D. Sachs (2021). Optimal need-based financial aid. *Journal of Political Economy 129*(2), 492–533.

D'Amato, M. and D. Mookherjee (2013). Welfare economics of educational policy when signaling and credit constraints co-exist. Working Paper, University of Salerno and Boston University.

Das, J., Q.-T. Do, and B. Özler (2004). Conditional cash transfers and the equity-efficiency debate. Working Paper No. 3280, World Bank.

De Janvry, A., F. Finan, E. Sadoulet, and R. Vakis (2006). Can conditional cash transfer programs serve as safety nets in keeping children at school and from working when exposed to shocks? *Journal of Development Economics 79*(2), 349–373.

Fender, J. and P. Wang (2003). Educational policy in a credit constrained economy with skill heterogeneity. *International Economic Review 44*(3), 939–964.

Findeisen, S. and D. Sachs (2016). Education and optimal dynamic taxation: The role of income-contingent student loans. *Journal of Public Economics 138*, 1–21.

Freeman, S. (1996). Equilibrium income inequality among identical agents. *Journal of Political Economy 104*(5), 1047–1064.

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

Ghatak, M. and N. Jiang (2002). A simple model of inequality, occupational choice and development. *Journal of Development Economics 69*, 205–226.

Ghatak, M. and F. Maniquet (2019). Universal basic income: Some theoretical aspects. *Annual Review of Economics 11*, 895–928.

Ghatak, M. and K. Muralidharan (2020). An inclusive growth dividend: Reframing the role of income transfers in India's anti-poverty strategy. Forthcoming in *India Policy Forum* 2020, National Council of Applied Economic Research, New Delhi.

Government of India Economic Survey (2017). Technical report, Ministry of Finance, Government of India, New Delhi.

Heathcote, J. (2005). Fiscal policy with heterogeneous agents and incomplete markets. *Review of Economic Studies 72*(1), 161–188.

Heckman, J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. NBER Working Paper 12006, National Bureau of Economic Research, Cambridge MA.

Henderson, D., S. Polachek, and L. Wang (2011). Heterogeneity in schooling rates of return. *Economics of Education Review 30*(6), 1202-1214.

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-context.html.

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

Jacobs, B., D. Schindler, and H. Yang (2012). Optimal taxation of risky human capital. *Scandinavian Journal of Economics 114*(3), 908–931.

Kahhat, J. (2010). Labor earnings inequality: The demand for and supply of skills. In L. F. López-Calva and N. C. Lustig (Eds.), *Declining Inequality in Latin America: A Decade of Progress?*, pp. 25–38. Washington, DC: Brookings Institution Press.

Ljungqvist, L. (1993). Economic underdevelopment: The case of a missing market for human capital. *Journal of Development Economics 40*(2), 219–239.

Lloyd-Ellis, H. and D. Bernhardt (2000). Enterprise, inequality and economic development. *Review of Economic Studies 67*(1), 147–168.

López-Calva, L. F. and N. C. Lustig (Eds.) (2010). *Declining Inequality in Latin America: A Decade of Progress?* Washington, DC: Brookings Institution Press.

Loury, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica 49*(4), 843–867.

Maoz, Y. D. and O. Moav (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 R. Klein Lecture: Emergent class structure. *International Economic Review 47*(2), 327–360.

Mookherjee, D. and S. Napel (2007). Intergenerational mobility and macroeconomic history dependence. *Journal of Economic Theory 137*(1), 49–78.

Mookherjee, D. and S. Napel (2019). Poverty dynamics and efficiency effects of conditional cash transfers. Working Paper, Boston University and University of Bayreuth.

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

Mookherjee, D. and D. Ray (2003). Persistent inequality. *Review of Economic Studies 70*(2), 369–393.

Mookherjee, D. and D. Ray (2008). A dynamic incentive-based argument for conditional transfers. *Economic Record 84*, S2–S16.

Mookherjee, D. and D. Ray (2010). Inequality and markets: Some implications of occupational diversity. *American Economic Journal: Microeconomics 2*(4), 38–76.

Mundle, S. (2017). Symposium on Universal Basic Income: Overview of deliberations, *Indian Journal of Human Development 11*(2), 206–209.

Parker, S. W. and P. E. Todd (2017). Conditional cash transfers: The case of Progresa/Oportunidades. *Journal of Economic Literature 55*(3), 866–915.

Parker, S. W. and Vogl (2018). Do conditional cash transfers improve economic outcomes in the next generation? Evidence from Mexico. NBER Working Paper 24303, National Bureau of Economic Research, Cambridge MA.

Peruffo, M. and P. C. Ferreira (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.

Tsai, S. and Y. Xie (2011). Heterogeneity in returns to college education: Selection bias in contemporary Taiwan. *Social Science Research 40*(3), 796–810.

Van Parijs, P. and Y. Vanderborght (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.