Optimal tax when children's abilities depend on parents' resources

Alexander Gelber and Matthew Weinzierl*

February 10, 2009

Preliminary and incomplete

Abstract

Micro and macro data suggest that the distribution of resources among families affects children's future earnings abilities. As a result, an optimal tax policy will treat future ability distributions as endogenous. In this paper, we analytically characterize how making children's abilities a function of parental resources affects optimal tax policy. We also numerically simulate optimal policy in this setting, using evidence on the distribution and heritability of natural abilities and the elasticity of children's ability with respect to parental resources to parameterize the model. Preliminary results show that optimal policy raises the earnings abilities of all children relative to a policy that neglects the endogeneity of ability, and that the welfare gains from optimal policy are substantial.

This paper is motivated by two empirical observations. First, recent evidence has shown that increased disposable incomes for low-income parents raises the performance of their children on tests of cognitive ability. Paxson and Schady (2007) and Milligan and Stabile (2008) study programs in Ecuador and Canada and find evidence that increasing parental resources, in particular through transfers to poor families, generates improved performance by their children. Second, as shown below, developed countries with more redistributive policies in the 1980s had statistically significantly lower pre-tax income inequality approximately one generation later, controlling for redistribution at the later date and other natural explanatory variables.

Together, these observations suggest a potentially important (positive and normative) role for tax policy in helping to determine the distribution of abilities in the taxpayer population through its effects on parental resources. The modern study of taxation, starting with Mirrlees (1971), makes clear that the distribution of abilities is a key determinant of how difficult a tradeoff between efficiency and equality society will face. If parental resources matter for children's abilities, the optimal tax treatment of parents' incomes can make the tax problem easier for future generations.

In principle, the characteristics of optimal tax policy in such an environment are ambiguous. Consider one natural proposal based on the evidence mentioned above: increase the disposable incomes of low earners. This will increase the abilities of their children, making the tax problem for the children's generation easier. But, it also will weaken the incentives to work in the parents' generation, making today's tax problem harder. The opposite policy, increasing the disposable incomes of high earners, avoids that tension because it raises the abilities of high-income children while relaxing today's incentive problems. However, it reduces the welfare of low-income parents, and its effects on the high-income children may be small.

This paper addresses this ambiguity by embedding in a formal model of optimal taxation the effects of parental resources on children's ability. We incorporate endogenous future generations' ability distributions into a dynamic Mirrleesian tax framework and show how doing so affects analytical results on optimal policy.

*Wharton and Harvard Business School, contact mweinzierl@hbs.edu

Then, we choose values for the key parameters in the model to match recent evidence on the effects of parental resources on child achievement, the heritability of natural cognitive ability, and the ability distribution in the United States. We numerically simulate and characterize optimal policy, and we detail the welfare gains from taking this source of endogeneity into account.

This paper is part of a surge of current research[1] that relaxes standard optimal tax models' division of the determinants of income into either exogenous ability or endogenous effort. Parental resources, uniquely, are both exogenous (to children) and endogenous (to parents), so they cannot be classified as either ability or effort. Modeling parental resources appropriately requires a dynamic approach to the optimal tax problem that is flexible enough to treat a variable as endogenous to one generation and exogenous to the next.[2]

Finally, this paper also provides a methodological contribution to the literature. To compare the optimal policy with one where the effect of parental resources is ignored, we specify a "mistaken" planner's problem. The planner is mistaken in that it has false beliefs about the structure of the economy and the wage process. The methodological challenge is that this mistaken planner must be modeled as ignorant of the mistakes it makes but still bound by them, so that its task is no easier than a correct planner's problem. Our approach to this mistaken planner starts by having it solve its perceived optimization problem, including the mistaken incentive compatibility and feasibility constraints. We then have agents optimize among the choices offered by the mistaken planner. This leads to allocations that are generally not incentive compatible (i.e., private information is not revealed) and that either leave resources unspent or that violate feasibility. The resulting planner's deficit or surplus is treated as a windfall relaxing or tightening (respectively) the feasibility constraints of the correct planner's problem, making its outcome comparable to the mistaken planner's. This technique may prove useful in other contexts requiring comparisons between imperfect planners and the optimum.

The paper proceeds as follows. Section 1 describes the dynamic optimal tax model in which a child's ability is a function of his parent's disposable income and an exogenous component, i.e., the parent's ability. Section 2 analyzes the optimal tax policy in this model and compares the resulting design to that recommended in a model with fully exogenous ability. Section 3 calibrates the model to recent evidence on the effect of parental resources on children's ability, simulates the optimal policy for a realistic distribution of ability, and characterizes both the structure and welfare implications of optimal taxes. Section 4 turns to data on income redistribution and inequality in the OECD over the last few decades. These data show, consistent with the implications of the model, that countries with more redistributive policies exhibit lower pre-tax and pre-transfer income inequality approximately one generation later. Section 5 concludes.

1 The Model

Individuals are linked in families, with one individual per generation in each family. Generations are indexed by $t = \{1, ..., T\}$ and families are indexed by $f = \{1, ..., F\}$. Each individual is described by a "natural ability" level, indexed by $i = \{1, ..., I\}$, where $\omega_t^{f,i}$ denotes the natural ability level i for an individual in family f of generation t. Each individual's wage is a function of both this natural ability and the disposable income of the individual's parent. Let $w\left(\omega_t^{f,i}, c_{t-1}^f\left(\boldsymbol{\omega}_{t-1}^f\right)\right)$ denote the wage of the individual of generation t in family f with natural ability level i and parental disposable income $c_{t-1}^f\left(\boldsymbol{\omega}_{t-1}^f\right)$, where parental disposable

[1] See, among others, Anderberg (2009), Grochulski and Piskorski (2006), and Kapicka (2006).

[2] Contributions to that literature include Golosov, Kocherlakota, and Tsyvinski (2003), Kocherlakota (2005), Golosov and Tsyvinski (2006), Albanesi and Sleet (2006), Golosov, Tsyvinski, and Werning (2007), Farhi and Werning (2008), and Weinzierl (2009).

income is a function of $\boldsymbol{\omega}_{t-1}^{f}$, the history of natural abilities in the family through generation $t-1$. Natural ability may be correlated across generations.

The planner maximizes the ex ante ($t=0$) expected utility of the families in the population. As all families are identical ex ante, this is equivalent to maximizing the expected utility of a single family, so we omit the family notation for simplicity. The planner's objective is:

$$\max_{c,y} E_0 \left[\sum_{t=1}^{T} \beta^t \sum_{\boldsymbol{\omega}_t} \pi(\boldsymbol{\omega}_t) \left(u(c_t(\boldsymbol{\omega}_t)) - v\left(\frac{y_t(\boldsymbol{\omega}_t)}{w(\omega_t^i, c_{t-1}(\boldsymbol{\omega}_{t-1}))} \right) \right) \right],$$

where $c_0(\boldsymbol{\omega}_0)$ is exogenously given. The probability $\pi(\boldsymbol{\omega}_t)$ is the probability that the sequence $\boldsymbol{\omega}_t$ is realized for the representative family.

The planner's maximization is subject to generation-by-generation feasibility constraints:[3]

$$\sum_{\boldsymbol{\omega}_t} \pi(\boldsymbol{\omega}_t) \left[(y_t(\boldsymbol{\omega}_t) - c_t(\boldsymbol{\omega}_t)) \right] + X_t \geq 0,$$

where X_t is an exogenous shock to feasibility, a "windfall", received at time t. We include this feasibility shock to allow us to relate the optimall planner's policy to a mistaken planner's policy, discussed in more detail below.

Maximization is constrained by incentive compatibility, as well. Each individual maximizes her own well-being given the available allocations. In this model, incentive compatibility constraints take the form that each individual prefers its allocation, regardless of its ancestors' choices, to the allocation of any other individual of its generation. The incentive constraint for an individual of natural ability i and generation t is:

$$\begin{pmatrix} u\left(c_t\left(\left\{\omega_t^i, \boldsymbol{\omega}_{t-1}\right\}\right)\right) - v\left(\frac{y_t\left(\left\{\omega_t^i, \boldsymbol{\omega}_{t-1}\right\}\right)}{w\left(\omega_t^i, c_{t-1}(\boldsymbol{\omega}_{t-1})\right)}\right) + \\ \sum_{\tau=1}^{T-t} \beta^t \sum_{\boldsymbol{\omega}_{t+\tau}} \pi\left(\boldsymbol{\omega}_{t+\tau} | (\boldsymbol{\omega}_{t+\tau-1})\right) \left(u\left(c_{t+\tau}\left(\left\{\boldsymbol{\omega}_{t,t+\tau}^i, \omega_t^i, \boldsymbol{\omega}_{t-1}\right\}\right)\right) - v\left(\frac{y_{t+\tau}\left(\left\{\boldsymbol{\omega}_{t,t+\tau}^i, \omega_t^i, \boldsymbol{\omega}_{t-1}\right\}\right)}{w\left(\omega_{t+\tau}^i, c_{t+\tau-1}\left(\left\{\boldsymbol{\omega}_{t,t+\tau-1}^i, \omega_t^i, \boldsymbol{\omega}_{t-1}\right\}\right)\right)}\right)\right) \end{pmatrix}$$
$$\geq \begin{pmatrix} u\left(c_t\left(\left\{\omega_t^{i'}, \boldsymbol{\omega}_{t-1}\right\}\right)\right) - v\left(\frac{y_t\left(\left\{\omega_t^{i'}, \boldsymbol{\omega}_{t-1}\right\}\right)}{w\left(\omega_t^i, c_{t-1}(\boldsymbol{\omega}_{t-1})\right)}\right) + \\ \sum_{\tau=1}^{T-t} \beta^t \sum_{\boldsymbol{\omega}_{t+\tau}} \pi\left(\boldsymbol{\omega}_{t+\tau} | (\boldsymbol{\omega}_{t+\tau-1})\right) \left(u\left(c_{t+\tau}\left(\left\{\boldsymbol{\omega}_{t,t+\tau}^i, \omega_t^{i'}, \boldsymbol{\omega}_{t-1}\right\}\right)\right) - v\left(\frac{y_{t+\tau}\left(\left\{\boldsymbol{\omega}_{t,t+\tau}^i, \omega_t^{i'}, \boldsymbol{\omega}_{t-1}\right\}\right)}{w\left(\omega_{t+\tau}^i, c_{t+\tau-1}\left(\left\{\boldsymbol{\omega}_{t,t+\tau-1}^i, \omega_t^{i'}, \boldsymbol{\omega}_{t-1}\right\}\right)\right)}\right)\right) \end{pmatrix}$$

which must hold for all i', $\boldsymbol{\omega}_{t-1}$, and $\boldsymbol{\omega}'_{t-1}$. The object $\boldsymbol{\omega}_{t,t+\tau}^i$ is the sequence of natural abilities running from generation t to generation $t+\tau$ ($\boldsymbol{\omega}_{t+\tau}^i$ indicates the series starting at $t=0$ and ending at $t+\tau$). Note that each generation expects the subsequent generations to choose their intended allocations, no matter this generation's report. That truth-telling is guaranteed by backward induction from the generation T incentive constraint.

1.1 Two-type, two-generation model

We focus on a smaller version of the model economy to highlight the main forces at work. First, natural ability comes in two levels: ω^h and ω^l, representing "high" and "low" abilities. Second, only two generations,

[3] Incorporating overall intergenerational transfers would be straightforward but would introduce a set of issues around intergenerational justice that this paper is not designed to address.

1 and 2, exist.[4]

Before specifying the planner's problem, we characterize the choices of a parent of type j.

1.2 Private optimum

For reference, consider a parent of type j solving her own private optimization, cognizant of the dependence of her child's ability on her own disposable income. Consider the results if we make the following assumption:

Assumption 1: *Parents observe their children's natural abilities only after choosing their own optimal effort levels..*

Given this assumption, parents choose their effort levels knowing only the conditional distributions of children's ability given parent's ability. The parent's choices will satisfy the following optimality condition (see the Appendix for a derivation):

$$\frac{v'\left(\frac{y_1\left(\omega_1^j\right)}{w_1^j}\right)\frac{1}{w_1^j}}{u'\left(c_1^j\right)} = 1 + \sum_i \left(p^i|\omega_1^j\right)\Omega^{j,i} \tag{1}$$

where

$$\Omega^{j,i} = \beta v'\left(\frac{y_2^{j,i}}{w_2^{i,j}}\right)\frac{y_2^{j,i}}{w_2^{i,j}}\varepsilon^{j,i}$$

and

$$\varepsilon^{j,i} = \frac{\partial w\left(\omega_2^i, c_1\left(\omega_1^j\right)\right)}{\partial c_1\left(\omega_1^j\right)}\frac{c_1\left(\omega_1^j\right)}{\left(w\left(\omega_2^i, c_1\left(\omega_1^j\right)\right)\right)}$$

is the elasticity of a child's ability with respect to parent's after-tax income.

The left hand side in (1) is the ratio of the parent's marginal disutility of income to marginal utility of consumption, and it is set to one in a standard private optimization that neglects the dependence of children's ability on parental resources. The second term on the right hand side is positive, causing a parent to exert more effort and/or obtain more disposable income because her children's disutility of earning income will fall when their wages are thereby increased.[5] The term $\Omega^{j,i}$ measures this effect for a child of type i.

Importantly, $\Omega^{j,i}$ is not constant across parent types. First, in these models, income and labor effort tend to rise with natural ability. If $v''(l) > 0$, this implies that $\Omega^{j,i}$ is larger for higher-skilled children. Second, $\frac{\partial\Omega^{j,i}}{\partial\varepsilon^{j,i}} > 0$, so $\Omega^{j,i}$ raises the optimal disposable income level $c_1\left(\omega_1^j\right)$ more for those parents whose children's wages are particularly elastic with respect to parental disposable income. These two factors most likely work in opposite directions. If natural ability is positively correlated across generations within families, the weighting of $\Omega^{j,i}$ by the conditional probabilities $\left(p^i|\omega_1^j\right)$ will have parents with higher natural abilities

[4] If more than two generations existed, the optimal policy in the first generation would take into account not only the implications of redistribution for the second generation's ability distribution but also what distributions that policy makes possible for the third generation and beyond. In other words, policy will gradually aim to generate the optimal distribution of abilities, subject to incentive constraints and the distribution of natural abilities. The two-generation model provides a view of the first step of that process.

[5] Note that the utility gain from a marginal increase in the child's wage is:

$$v\left(\frac{y_2^{j,i}}{w_2^{i,j}}\right)\frac{-\partial v\left(\frac{y_2^{j,i}}{w_2^{i,j}}\right)}{\partial w_2^{i,j}}\frac{w_2^{i,j}}{v\left(\frac{y_2^{j,i}}{w_2^{i,j}}\right)} = v'\left(\frac{y_2^{j,i}}{w_2^{i,j}}\right)\frac{y_2^{j,i}}{w_2^{i,j}}$$

exerting relatively more effort. On the other hand, if the elasticity of the wage is higher for children of lower-income parents, parents with lower natural abilities will work relatively hard. Which of these factors dominates is an empirical question.

For future reference, let

$$\left(\tau^j\right)^{Private} = 1 - \frac{v'\left(\frac{y_1^j}{w_1^j}\right)\frac{1}{w_1^j}}{u'\left(c_1^j\right)}$$

denote the distortion to the conventional consumption-leisure margin of the parent of type j in the private optimum. Using (1), this is:

$$\left(\tau^j\right)^{Private} = -\sum_i \left(p^i|\omega_1^j\right)\beta v'\left(\frac{y_2^{j,i}}{w_2^{i,j}}\right)\frac{y_2^{j,i}}{w_2^{i,j}}\varepsilon^{j,i}$$

so that the parent of type j acts as if there were a marginal subsidy to work relative to a model without endogenous ability.

1.3 Planner's problem

The planner's problem with two natural ability types and two generations is as follows.

Problem 1 (Two-generation planner's problem)

$$\max_{c,y}\sum_{j=h,l}\pi^j\left\{\begin{array}{c} u\left(c_1\left(\omega_1^j\right)\right) - v\left(\frac{y_1\left(\omega_1^j\right)}{w_1^j}\right) \\ +\beta\sum_{i=h,l}\left(p^i|\omega_1^j\right)\left(u\left(c_2\left(\left\{\omega_1^j,\omega_2^i\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^i\right\}\right)}{w\left(\omega_2^i,c_1\left(\omega_1^j\right)\right)}\right)\right)\end{array}\right\},$$

where π^j is the population proportion of natural ability j and $p^i|\omega_1^j$ is the probability of a child having natural ability i given that the parent's ability is j. Optimization is subject to period-by-period feasibility:

$$\sum_{j=h,l}\pi^j\left[\left(y_1\left(\omega_1^j\right) - c_1\left(\omega_1^j\right)\right)\right] + X_1 \geq 0, \qquad (\phi_1)$$

and

$$\sum_{j=h,l}\pi^j\sum_{i=h,l}\left(p^i|\omega_1^j\right)\left[\left(y_2\left(\left\{\omega_1^j,\omega_2^i\right\}\right) - c_2\left(\left\{\omega_1^j,\omega_2^i\right\}\right)\right)\right] + X_2 \geq 0, \qquad (\phi_2)$$

where X_1 and X_2 are defined below; and subject to incentive compatibility constraints:

$$\begin{aligned} & u\left(c_1\left(\omega_1^h\right)\right) - v\left(\frac{y_1\left(\omega_1^h\right)}{w_1^h}\right) + \beta\sum_{i=h,l}\left(p^i|\omega_1^h\right)\left(u\left(c_2\left(\left\{\omega_1^h,\omega_2^i\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^h,\omega_2^i\right\}\right)}{w\left(\omega_2^i,c_1\left(\omega_1^h\right)\right)}\right)\right) \quad (\mu^{l|h}) \\ \geq\ & u\left(c_1\left(\omega_1^l\right)\right) - v\left(\frac{y_1\left(\omega_1^l\right)}{w_1^h}\right) + \beta\sum_{i=h,l}\left(p^i|\omega_1^h\right)\left(u\left(c_2\left(\left\{\omega_1^l,\omega_2^i\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^l,\omega_2^i\right\}\right)}{w\left(\omega_2^i,c_1\left(\omega_1^l\right)\right)}\right)\right)\end{aligned}$$

$$
\begin{aligned}
& u\left(c_1\left(\omega_1^l\right)\right)-v\left(\frac{y_1\left(\omega_1^l\right)}{w_1^l}\right)+\beta\sum_{i=h,l}\left(p^i|\omega_1^l\right)\left(u\left(c_2\left(\left\{\omega_1^l,\omega_2^i\right\}\right)\right)-v\left(\frac{y_2\left(\left\{\omega_1^l,\omega_2^i\right\}\right)}{w\left(\omega_2^i,c_1\left(\omega_1^l\right)\right)}\right)\right);\ \left(\mu^{h|l}\right) \\
\geq\ & u\left(c_1\left(\omega_1^h\right)\right)-v\left(\frac{y_1\left(\omega_1^h\right)}{w_1^l}\right)+\beta\sum_{i=h,l}\left(p^i|\omega_1^l\right)\left(u\left(c_2\left(\left\{\omega_1^h,\omega_2^i\right\}\right)\right)-v\left(\frac{y_2\left(\left\{\omega_1^h,\omega_2^i\right\}\right)}{w\left(\omega_2^i,c_1\left(\omega_1^h\right)\right)}\right)\right)
\end{aligned}
$$

for all i*,and:*

$$
u\left(c_2\left(\left\{\omega_1^j,\omega_2^h\right\}\right)\right)-v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^h\right\}\right)}{w\left(\omega_2^h,c_1\left(\omega_1^j\right)\right)}\right)\geq u\left(c_2\left(\left\{\omega_1^j,\omega_2^l\right\}\right)\right)-v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^l\right\}\right)}{w\left(\omega_2^l,c_1\left(\omega_1^j\right)\right)}\right);\qquad\left(\lambda^{j,l|j,h}\right)
$$

$$
u\left(c_2\left(\left\{\omega_1^j,\omega_2^l\right\}\right)\right)-v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^l\right\}\right)}{w\left(\omega_2^l,c_1\left(\omega_1^j\right)\right)}\right)\geq u\left(c_2\left(\left\{\omega_1^j,\omega_2^h\right\}\right)\right)-v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^h\right\}\right)}{w\left(\omega_2^h,c_1\left(\omega_1^j\right)\right)}\right)\qquad\left(\lambda^{j,h|j,l}\right)
$$

for all j*.*

The last two incentive constraints guarantee that a child will reveal its natural ability regardless of the parent's report. We denote the multipliers on the constraints above ϕ_1 and ϕ_2 for the feasibility constraints, $\mu_1^{j'|j}$ for the incentive constraint on the parent of natural ability level j, and $\lambda_2^{j,i'|j,i}$ for the incentive constraint on the child of natural ability level i whose parent had natural ability level j.

Some notation will simplify our results. Recall that $\varepsilon^{j,i}$ denotes the elasticity of a child's ability with respect to its parent's after-tax income. Denote $y_2\left(\left\{\omega_1^j,\omega_2^{i'}\right\}\right)$ as $y_2^{j,i'}$, and similarly for other variables. We also make the following simplifying functional form assumption:

Assumption 2: *The components of the individual utility function take the forms:*

$$
\begin{aligned}
u\left(c\right) &= \ln c \\
v\left(l\right) &= \frac{1}{\sigma}l^{\sigma}.
\end{aligned}
$$

For the rest of this example, Assumption 2 is maintained unless stated otherwise.

1.3.1 Optimal policy

The first order conditions for y_1^h and c_1^h, income and consumption of parents of type h, yield

$$
1-\tau^j=\mathbf{A}^j\left(1+\sum_i\mathbf{B}^{j,i}\Omega^{j,i}\right)\tag{2}
$$

where

$$
\mathbf{A}^j=\frac{\pi^j+\mu_1^{j'|j}-\mu_1^{j|j'}}{\pi^j+\mu_1^{j'|j}-\mu_1^{j|j'}\frac{v'\left(\frac{y_1\left(\omega_1^j\right)}{w_1^{j'}}\right)\frac{1}{w_1^{j'}}}{v'\left(\frac{y_1\left(\omega_1^j\right)}{w_1^j}\right)\frac{1}{w_1^j}}}
$$

$$\mathbf{B}^{j,i} = \frac{\pi^j \left(p^i|\omega_1^j\right) + \mu_1^{j'|j} \left(p^i|\omega_1^j\right) - \mu_1^{j|j'} \left(p^i|\omega_1^{j'}\right) + \sum_{i'} \lambda_2^{j,i'|j,i} \left(1 - \frac{v'\left(\frac{y_2^{j,i'}}{w_2^{i,j}}\right)}{v'\left(\frac{y_2^{j,i}}{w_2^{i,j}}\right)} \frac{y_2^{j,i'}}{y_2^{j,i}}\right)}{\pi^j + \mu_1^{j'|j} - \mu_1^{j|j'}}$$

$$\Omega^{j,i} = \beta v' \left(\frac{y_2^{j,i}}{w_2^{i,j}}\right) \frac{y_2^{j,i}}{w_2^{i,j}} \varepsilon^{j,i} \tag{3}$$

Before analyzing expression (2) in general, consider the following special case.

First-best When the planner can observe natural ability, allocations satisfy:

$$\left(\tau^j\right)^{First-Best} = \left(\tau^j\right)^{Private}. \tag{4}$$

The first-best policy does not distort private decisions. This equivalence does not imply that parents are indifferent between the first-best and private allocations, however. The first-best planner will redistribute across parent types with lumpsum taxation not reflected in this marginal condition.

Note that the equivalence in (4) would hold in a standard optimal tax model without the dependence of children's ability on parental resources.

Constrained Optimal Policy While the first-best scenario is a useful benchmark, we now turn to analyzing the optimal policy.

With two natural ability types, only one incentive constraint multiplier can be nonzero in each generation. Assume that $\frac{w_1^h}{w_1^l} > 1$ and $\frac{w\left(\omega_2^h, c_1\left(\omega_1^j\right)\right)}{w\left(\omega_2^l, c_1\left(\omega_1^j\right)\right)} > 1$ such that $\mu_1^{h|l} = 0$ and $\lambda_2^{j,h|j,l} = 0$ for all j: that is, only the incentive constraints on the high-skilled parents and children bind.

High-skilled in the Constrained Optimal Policy The components of condition (2) for high-skilled parents are

$$\mathbf{A}^h = 1$$

$$\mathbf{B}^{h,h} = \left(p^h|\omega_1^h\right) + \frac{\lambda_2^{h,l|h,h}}{\pi^h + \mu_1^{l|h}} \left(1 - \frac{v'\left(\frac{y_2^{h,l}}{w_2^{h,h}}\right)}{v'\left(\frac{y_2^{h,h}}{w_2^{h,h}}\right)} \frac{y_2^{h,l}}{y_2^{h,h}}\right)$$

$$\mathbf{B}^{h,l} = \left(p^l|\omega_1^h\right)$$

$$\Omega^{h,i} = \beta v' \left(\frac{y_2^{h,i}}{w_2^{i,h}}\right) \frac{y_2^{h,i}}{w_2^{i,h}} \varepsilon^{h,i}.$$

To compare the policy implied by these results to the first-best and private optimum allocations, the key calculation is the difference $\sum_i \mathbf{B}^{h,i}\Omega^{h,i} - \sum_i \left(p^i|\omega_1^h\right)\Omega^{h,i}$. Given our assumption on the form of $v(l)$ and the results above, this difference is:

$$\sum_i \mathbf{B}^{h,i}\Omega^{h,i} - \sum_i \left(p^i|\omega_1^h\right)\Omega^{h,i} = \beta \frac{\lambda_2^{h,l|h,h}}{\pi^h + \mu_1^{l|h}} \frac{\left(y_2^{h,h}\right)^\sigma - \left(y_2^{h,l}\right)^\sigma}{\left(w_2^{h,h}\right)^\sigma} \varepsilon^{h,h} > 0 \tag{5}$$

The inequality in result (5), combined with (2), implies that the constrained optimal policy subsidizes the work of parents with high natural ability relative to the first-best or privately optimal allocations. That inequality holds only if $\varepsilon^{h,h} > 0$: that is, only if high-skilled children's wages are positively affected by their parents' resources. Intuitively, raising high-skilled children's wages makes it easier to satisfy their incentives not to mimic the low-skilled, relaxing the constraints on the planner. The more elastic are high-skilled children's wages to parental resources, the more powerful is this factor.

In contrast, result (5) also implies that $\mathbf{B}^{h,h}$ is negatively related to $\mu_1^{l|h}$. This partially offsets the factors favoring subsidies to the effort of high-ability parents. When the high-skilled parents are tempted to mimic the low-skilled, distorting the high-skilled parents toward more effort makes this temptation stronger. Therefore, the planner does less of this distortion than it would if parents' types were known.

In sum, the optimal condition for a high-skilled parent is:

$$1 - \left(\tau^h\right)^{Optimal} = 1 + \sum_i \left(p^i|\omega_1^h\right)\Omega^{h,i} + \frac{\lambda_2^{h,l|h,h}}{\pi^h + \mu_1^{l|h}}\left(1 - \left(\frac{y_2^{h,l}}{y_2^{h,h}}\right)^{\sigma}\right)\Omega^{h,h}$$

so that $\left(\tau^h\right)^{Optimal}$ provides a larger spur to effort than in the first-best or privately optimal allocations.

Low-skilled in the Constrained Optimal Policy The components of condition (2) for low-skilled parents are

$$\mathbf{A}^l = \frac{\pi^l - \mu_1^{l|h}}{\pi^l - \mu_1^{l|h}\frac{v'\left(\frac{y_1\left(\omega_1^l\right)}{w_1^h}\right)\frac{1}{w_1^h}}{v'\left(\frac{y_1\left(\omega_1^l\right)}{w_1^l}\right)\frac{1}{w_1^l}}}$$

$$\mathbf{B}^{l,h} = \left(p^h|\omega_1^l\right)\frac{\pi^l - \mu_1^{l|h}\frac{\left(p^h|\omega_1^h\right)}{\left(p^h|\omega_1^l\right)}}{\pi^l - \mu_1^{l|h}} + \frac{\lambda_2^{l,l|l,h}\left(1 - \frac{v'\left(\frac{y_2^{l,l}}{w_2^{h,l}}\right)}{v'\left(\frac{y_2^{l,h}}{w_2^{h,l}}\right)}\frac{y_2^{l,l}}{y_2^{l,h}}\right)}{\pi^l - \mu_1^{l|h}}$$

$$\mathbf{B}^{l,l} = \left(p^l|\omega_1^l\right)\frac{\pi^l - \mu_1^{l|h}\frac{\left(p^l|\omega_1^h\right)}{\left(p^l|\omega_1^l\right)}}{\pi^l - \mu_1^{l|h}}$$

$$\Omega^{l,i} = \beta v'\left(\frac{y_2^{l,i}}{w_2^{i,l}}\right)\frac{y_2^{l,i}}{w_2^{i,l}}\varepsilon^{l,i}$$

We can simplify these if we assume the following:

$$\pi^l > \mu_1^{l|h},$$

which is difficult to check analytically but which numerical simulations confirm.

We also assume that the children of low-ability parents are more likely to be low-ability than the children of high-ability parents:

$$\left(p^l|\omega_1^l\right) > \left(p^l|\omega_1^h\right),$$

a relationship we will quantify later in the paper.

Along with the functional forms specified above, these assumptions imply:

$$\mathbf{A}^l = \frac{\pi^l - \mu_1^{l|h}}{\pi^l - \mu_1^{l|h}\left(\frac{w_1^l}{w_1^h}\right)^\sigma} < 1.$$

The inequality for$\mathbf{A}^l$ reflects a standard second-best distortion that discourages effort by the low-skilled. It lowers the marginal disutility of income relative to the marginal utility of consumption for the lower type, discouraging higher types from mimicking it.

We can also derive the difference $\sum_i \mathbf{B}^{l,i}\Omega^{l,i} - \sum_i \left(p^i|\omega_1^l\right)\Omega^{l,i}$, which plays a key role here as it did for the high-skilled parents, indicating whether the planner encourages effort by the low-skilled parent more or less in the constrained optimal policy than in the first best:

$$\sum_i \mathbf{B}^{l,i}\Omega^{l,i} - \sum_i \left(p^i|\omega_1^l\right)\Omega^{l,i} = \beta\frac{\lambda_2^{l,l|l,h}}{\pi^l - \mu_1^{l|h}}\left(1 - \left(\frac{y_2^{l,l}}{y_2^{l,h}}\right)^\sigma\right)\Omega^{l,h} + \beta\frac{\mu_1^{l|h}}{\pi^l - \mu_1^{l|h}}\left[\left(\left(p^l|\omega_1^l\right) - \left(p^l|\omega_1^h\right)\right)\left(\Omega^{l,l} - \Omega^{l,h}\right)\right] \tag{6}$$

For the most plausible cases, expression (6) is positive, pushing the planner toward encouraging effort by low-skilled parents and working against the influence of $\mathbf{A}^l$. We discuss each term separately.

As with the high-skilled parent's expression (5), the first term on the right-hand side in (6) indicates that satisfying the incentives of the high-skilled children of low-skilled parents is made easier if they have higher skills. Note that the role of $\mu_1^{l|h}$ is reversed in this expression relative to (5): that is, it is subtracted in the denominator, so it increases the size of the optimal marginal distortion. Intuitively, the desire to subsidize the effort of low-ability parents and therefore increase the resources available to their high-ability children must be tempered so as to prevent the high-skilled parents from mimicking the low-skilled.

The second term in (6) captures the planner's gain from targeting the parents whose children benefit the most from extra parental resources. Formally, this term is positive if either of two situations exist: 1) if natural ability is positively correlated across generations within a family and parental resources are more valuable when invested in children of low natural ability than of high-natural ability (i.e., $\left(p^l|\omega_1^l\right) > \left(p^l|\omega_1^h\right)$ and $\Omega^{l,l} > \Omega^{l,h}$); or 2) if the opposite holds for both of these (i.e., $\left(p^l|\omega_1^l\right) < \left(p^l|\omega_1^h\right)$ and $\Omega^{l,l} < \Omega^{l,h}$). In either case, the planner has low-skilled parents earn more (and have higher disposable income) when low-skilled parents have a greater proportion than high-skilled parents of the children who benefit most from parental resources. If these conditions do not hold, so that the children who benefit most from parental resources are had in a greater proportion by high-skilled parents, the planner discourages work among the low-skilled.

Another way to interpret this second term starts with the fact that the income tax has pushed high-skilled parents to the point of indifference between revealing their type and mimicking the low-skilled. The planner can relax the incentive constraint on these parents by encouraging the low-skilled parents to work more, giving them a smaller transfer, and leaving them with a higher after tax income. This is valued by low-skilled parents because their higher disposable income raises their children's skills. It is less valuable to high-skilled parents to the extent that their children are less likely to have the low skills for which such investment is particularly valuable (this is the role of the second condition). To see this, note that this first term in brackets equals zero if $\left(p^l|\omega_1^l\right) = \left(p^l|\omega_1^h\right)$, that is if each parent type is equally likely to have low-skilled children. In that case, the benefits of higher disposable income are equal for each type, and the incentive constraint on the high-skilled is not loosened.

In sum, the optimality condition for the low-skilled worker can be written as:

$$1-\left(\tau^l\right)^{Optimal}=\frac{\pi^l-\mu_1^{l|h}}{\pi^l-\mu_1^{l|h}\left(\frac{w_1^l}{w_1^h}\right)^\sigma}\left(1+\sum_i\left(p^i|\omega_1^h\right)\Omega^{h,i}\frac{\pi^l-\mu_1^{l|h}\frac{\left(p^h|\omega_1^h\right)}{\left(p^h|\omega_1^l\right)}}{\pi^l-\mu_1^{l|h}}+\frac{\lambda_2^{l,l|l,h}}{\pi^l-\mu_1^{l|h}}\left(1-\left(\frac{y_2^{l,l}}{y_2^{l,h}}\right)^\sigma\right)\Omega^{l,h}\right). \tag{7}$$

This expression may be less than or greater than one, distorting the low-skilled parent either to work less than or more than she would choose on her own or than the planner would have her work in the first-best.

Examining two simplified cases helps with building intuition for (7). First, suppose parental type (but not each child's type) is observed by the planner. Then, $\mu_1^{l|h}=0$ and expression (7) becomes:

$$1-\left(\tau^l\right)^{Optimal}=1+\sum_i\left(\left(p^i|\omega_1^h\right)\Omega^{h,i}\right)+\frac{\lambda_2^{l,l|l,h}}{\pi^l-\mu_1^{l|h}}\left(1-\left(\frac{y_2^{l,l}}{y_2^{l,h}}\right)^\sigma\right)\Omega^{l,h}. \tag{8}$$

which is very similar to the expression for the distortion to the high-skilled parent, expression (5). By making the multiplicative factor in front of the right-hand side of (7) equal to one, and by eliminating the incentive problem created by investment in the children of the low-skilled, which was the ratio $\frac{\pi^l-\mu_1^{l|h}\frac{\left(p^h|\omega_1^h\right)}{\left(p^h|\omega_1^l\right)}}{\pi^l-\mu_1^{l|h}}$, observable parental type has reduced the need for distortions on the low-skilled parent. What remains is the final term in (8), which captures the loosening of incentive constraints across children types that the planner can achieve by providing more resources to low-ability parents. Therefore, expression (8) encourages effort by low-ability parents.

Next, suppose children's type (but not the parent's type) is observed by the planner. Then, $\lambda_2^{l,l|l,h}=0$ and expression (7) becomes:

$$1-\left(\tau^l\right)^{Optimal}=\frac{\pi^l-\mu_1^{l|h}}{\pi^l-\mu_1^{l|h}\left(\frac{w_1^l}{w_1^h}\right)^\sigma}\left(1+\sum_i\left(\left(p^i|\omega_1^h\right)\Omega^{h,i}\frac{\pi^l-\mu_1^{l|h}\frac{\left(p^h|\omega_1^h\right)}{\left(p^h|\omega_1^l\right)}}{\pi^l-\mu_1^{l|h}}\right)\right). \tag{9}$$

Now, the benefit from providing greater resources to low-ability parents is less than in the first-best, because doing so merely worsens the incentive problems among parents. Therefore, expression (9) discourages effort by low-ability parents.

1.4 Mistaken planner's problem

In this section, we introduce the formal statement of the problem solved by a social planner who is unaware of the effect of parental resources on children's skills. The problem perceived by the mistaken planner does not determine ex post allocations because of the planner's errors, which we discuss below, but the perceived problem is as follows:

Problem 2 (Two-generation mistaken planner's perceived problem)

$$\max_{c,y}\sum_{j=h,l}\pi^j\left\{\begin{array}{c}u\left(c_1\left(\omega_1^j\right)\right)-v\left(\frac{y_1\left(\omega_1^j\right)}{w_1^j}\right)\\+\beta\sum_{i=h,l}\left(p^i|\omega_1^j\right)\left(u\left(c_2\left(\left\{\omega_1^j,\omega_2^i\right\}\right)\right)-v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^i\right\}\right)}{w\left(\omega_2^i,c_1\left(\omega_1^j\right)\right)}\right)\right)\end{array}\right\},$$

where π^j is the population proportion of natural ability j and $p^i|\omega_1^j$ is the probability of a child having natural ability i given that the parent's ability is j. Optimization is subject to period-by-period feasibility:

$$\sum_{j=h,l} \pi^j \left[\left(y_1 \left(\omega_1^j \right) - c_1 \left(\omega_1^j \right) \right) \right] \geq 0, \qquad (\phi_1)$$

and

$$\sum_{j=h,l} \pi^j \sum_{i=h,l} \left(p^i|\omega_1^j \right) \left[\left(y_2 \left(\left\{ \omega_1^j, \omega_2^i \right\} \right) - c_2 \left(\left\{ \omega_1^j, \omega_2^i \right\} \right) \right) \right] \geq 0, \qquad (\phi_2)$$

and subject to perceived *incentive compatibility constraints:*

$$\begin{aligned} & u\left(c_1\left(\omega_1^h\right)\right) - v\left(\frac{y_1\left(\omega_1^h\right)}{w_1^h}\right) + \beta \sum_{i=h,l}\left(p^i|\omega_1^h\right)\left(u\left(c_2\left(\left\{\omega_1^h,\omega_2^i\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^h,\omega_2^i\right\}\right)}{w\left(\omega_2^i\right)}\right)\right) \quad (\mu^{l|h}) \\ \geq \quad & u\left(c_1\left(\omega_1^l\right)\right) - v\left(\frac{y_1\left(\omega_1^l\right)}{w_1^h}\right) + \beta \sum_{i=h,l}\left(p^i|\omega_1^h\right)\left(u\left(c_2\left(\left\{\omega_1^l,\omega_2^i\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^l,\omega_2^i\right\}\right)}{w\left(\omega_2^i\right)}\right)\right) \end{aligned}$$

$$\begin{aligned} & u\left(c_1\left(\omega_1^l\right)\right) - v\left(\frac{y_1\left(\omega_1^l\right)}{w_1^l}\right) + \beta \sum_{i=h,l}\left(p^i|\omega_1^l\right)\left(u\left(c_2\left(\left\{\omega_1^l,\omega_2^i\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^l,\omega_2^i\right\}\right)}{w\left(\omega_2^i\right)}\right)\right); \quad (\mu^{h|l}) \\ \geq \quad & u\left(c_1\left(\omega_1^h\right)\right) - v\left(\frac{y_1\left(\omega_1^h\right)}{w_1^l}\right) + \beta \sum_{i=h,l}\left(p^i|\omega_1^l\right)\left(u\left(c_2\left(\left\{\omega_1^h,\omega_2^i\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^h,\omega_2^i\right\}\right)}{w\left(\omega_2^i\right)}\right)\right) \end{aligned}$$

for all i,and:

$$u\left(c_2\left(\left\{\omega_1^j,\omega_2^h\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^h\right\}\right)}{w\left(\omega_2^h\right)}\right) \geq u\left(c_2\left(\left\{\omega_1^j,\omega_2^l\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^l\right\}\right)}{w\left(\omega_2^l\right)}\right); \qquad (\lambda^{j,l|j,h})$$

$$u\left(c_2\left(\left\{\omega_1^j,\omega_2^l\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^l\right\}\right)}{w\left(\omega_2^l\right)}\right) \geq u\left(c_2\left(\left\{\omega_1^j,\omega_2^h\right\}\right)\right) - v\left(\frac{y_2\left(\left\{\omega_1^j,\omega_2^h\right\}\right)}{w\left(\omega_2^h\right)}\right) \qquad (\lambda^{j,h|j,l})$$

for all j.

The difference between this planner's perceived problem and the optimal planner's problem is that this planner *perceives* the wages of the children (generation 2) as independent from the consumption levels of their parents (generation 1). That incorrect perception means that the planner solving this problem will offer allocations of (c, y) to parents that, when they choose optimally, are collectively either infeasible or leave a surplus. For instance, because the planner underappreciates the importance of parental resources, high-ability parents may be more tempted to cheat down and claim low-ability when facing a mistaken planner than when facing the optimal planner. We use this mistaken planner's problem in the next section to perform comparison simulations with the optimal policy.

The mistaken planner's problem is not completely specified, however, without accounting for how the planner fails to notice that its policy generates a deficit or surplus. If the mistaken planner observed the ex post infeasibility or excess resources generated by its policy, it would learn of its mistake. To address this, we allow the mistaken planner's policy to generate ex post deficits or surpluses. We label these as

Pop. Share	**Parent's natural ability**				
	60	80	100	120	140
	0.05	0.26	0.40	0.24	0.05

Probability of child's ability given parent's		**Child's natural ability**				
		60	80	100	120	140
Parent's natural ability	60	0.27	0.48	0.24	0.01	0.00
	80	0.08	0.41	0.42	0.09	0.00
	100	0.02	0.26	0.46	0.24	0.02
	120	0.00	0.09	0.41	0.41	0.09
	140	0.00	0.02	0.25	0.47	0.26

"windfalls" and assign them the following symbols:

$$X_1 = -\sum_{j=h,l} \pi^j \left[\left(y_1 \left(\omega_1^j \right) - c_1 \left(\omega_1^j \right) \right) \right], \tag{10}$$

and

$$X_2 = -\sum_{j=h,l} \pi^j \sum_{i=h,l} \left(p^i | \omega_1^j \right) \left[\left(y_2 \left(\left\{ \omega_1^j, \omega_2^i \right\} \right) - c_2 \left(\left\{ \omega_1^j, \omega_2^i \right\} \right) \right) \right], \tag{11}$$

where the quantities of income and consumption in (10) and (11) are those offered by the mistaken planner.

These windfalls enter the feasibility constraints of the optimal planner as specified above to make comparisons between policies possible. If the mistaken planner runs a deficit (e.g., if $X_1 < 0$), the optimal planner is granted an equivalent amount of resources as an endowment.

2 Calibration of the model

The model described requires three empirical inputs: 1) the distribution of natural abilities among parents and children; 2) the heritability of natural ability across generations and conditional probabilities for children's natural ability given parent's natural ability; 3) the function that turns natural ability and parental disposable income into a skill (wage) level for children.

We base our assumptions for natural ability and its heritability on the intelligence quotient (IQ), which has been used for decades to measure cognitive aptitude. As conventionally calculated, IQ is distributed normally with a mean of 100 and a standard deviation of 15. Following this distribution, we assign parents to one of five natural ability levels $\omega^l \in \{60, 80, 100, 120, 140\}$. For children, we rely on evidence that the correlation of IQ across generations is approximately 0.50 [citation]. This implies that the conditional distribution of a child's IQ given a parent's IQ of ω is normal with a mean of $\omega + 0.50\,(100 - \omega)$ and a standard deviation of $(225\,(1 - 0.50))^{\frac{1}{2}} = 10.6$. Using this conditional distribution, we generate 1000 children for each parental IQ level and assign them to the same five IQ levels. The results of these steps are the following probability distributions for parents and children's natural abilities:

The third component is the subject of recent study. In particular, Dahl and Lochner (2008) estimate that a \$1000 increase in family income, due to a change in the rules governing EITC eligibility in the United States, generates a change equal to 6% of a standard deviation in test scores among children. Milligan and Stabile (2009) give a similar figure when studying the effects of cash transfers to Canadian families that vary by province, though they also report some results suggesting smaller impacts. Paxson and Schady find that cash transfers in Ecuador yield improvements in cognitive performance among children, but that the effects are strong only for the poor.

To translate these findings into the model above, we must specify and calibrate a form for the wage as a function of natural ability and parental resources. The empirical literature provides virtually no guidance on the proper function form. For simplicity, we assume the following additive form:

$$w_2^{i,j} = \omega_2^j + \left(c_1^i - c^*\right)^\rho \tag{12}$$

so that natural ability is supplemented by a concave return to parental resources above some minimal level c^*.[6] Note that

$$\frac{\left(\partial w_2^{i,j}\right)^2}{\partial \omega_2^j \partial c_1^i} = 0$$

so that natural ability has no effect on the absolute change in the child's wage induced by an increase in parental resources. Parents will therefore find that investing in a child's skills will yield a smaller percentage increase in the child's wage the greater is the child's natural ability. Because parents with high incomes tend to have children with higher natural ability, increasing the disposable incomes of high-income parents will yield smaller percentage changes in children's wages, consistent with the findings of the papers mentioned above.[7]

From the evidence cited above, we know that a \$1000 increase in family income generates a change in the wage of 6% of the standard deviation of wages for the children of low-income parents. \$1000 is approximately 8.5% of the average income for the sample under study in Milligan and Stabile (2009). Assuming a wage of \$6 with a standard deviation of \$10 for these parents, and letting $c_1 = \$10$ (consumption per hour) and $c^* = \$0$, we can use this evidence to calculate a value for the only parameter in (12), $\rho = 0.88$.

2.1 Optimal policy simulation results

In this section, we compare the allocations under the optimal policy to the allocations chosen in equilibrium by individuals facing the mistaken planner's offered policy (which we call the mistaken planner equilibrium). This comparison provides the most direct evidence on the effects of interest in this paper. While real world policy does not follow the mistaken planner's policy, using an existing policy as the comparison to the optimal policy would conflate the effect of taking parental resources into account with the many other effects of using a Mirrleesian optimal tax policy.

We compare the policies along four dimensions: 1) parents' average tax rates; 2) children's equilibrium wages; 3) parents' and children's individual utilities; 4) social welfare.[8]

[6] The elasticity of the child's wage with respect to parent's c using this form is: $\varepsilon = \frac{\partial w_2^{i,j}}{\partial c_1^i} \frac{c_1^i}{w_2^{i,j}} = \rho \left(c_1^i - c^*\right)^{\rho - 1} \frac{c_1^i}{w_2^{i,j}}$

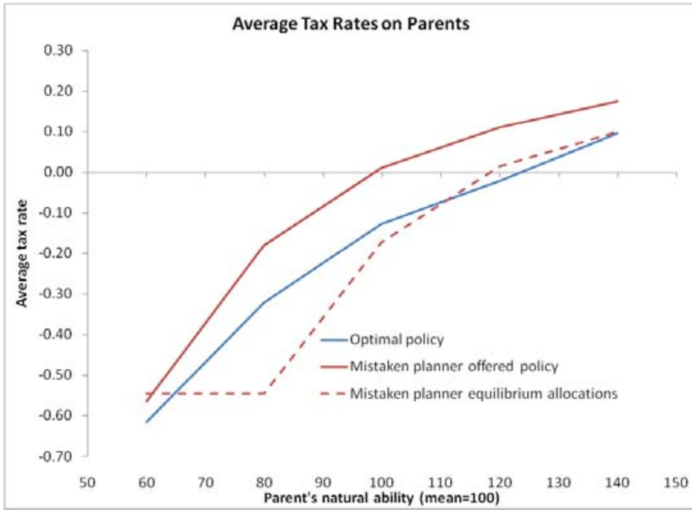
[7] A more general version that would allow for different assumptions about the complementarity or substitutability of ability and parental resources is:

$$w_2^{i,j} = A\left(\alpha\left(\omega_2^j\right)^\gamma + (1-\alpha)\left(c_1^j - c^*\right)^{\rho\gamma}\right)^{\frac{1}{\gamma}}$$

where the model in the paper assumes $A = 2$, $\alpha = \frac{1}{2}$, and $\gamma = 1$ but $\gamma \in [-\infty, 1]$ and $\frac{1}{1-\gamma}$ is the elasticity of substitution.

[8] TBD: use the numerical simulations to show the effects of relaxing parents' and children's constraints

Parents' average tax rates Optimal policy assesses lower average tax rates on parents across the ability distribution than does the mistaken planner, as shown in the following figure:



The mistaken planner assesses higher average tax rates because parents respond to its offered policy by mimicking lower-ability types. To illustrate this, the dashed line in the above figure shows the average taxes actually paid by parents facing the mistaken planner's offered policy. These parents' deviations result in less revenue raised by the planner, so taxes must be higher overall than in the optimal policy.

Children's equilibrium wages The optimal policy raises the wages of all types of children. The table below shows the ratio of children's wages in the optimal policy to their wages in the mistaken equilibrium.

Table 1: Ratio of children's wages in optimal policy to their wages in mistaken planner equilibrium

		Child's natural ability					
		60	80	100	120	140	*Weighted average*
Parent's natural ability	60	1.01	1.01	1.01	1.01	1.01	*1.01*
	80	1.06	1.05	1.04	1.03	1.03	*1.04*
	100	1.06	1.05	1.04	1.04	1.03	*1.04*
	120	1.07	1.05	1.05	1.04	1.03	*1.04*
	140	1.06	1.05	1.04	1.03	1.03	*1.03*

Memorandum: conditional probabilities

		Child's natural ability				
		60	80	100	120	140
Parent's natural ability	60	0.27	0.48	0.24	0.01	-
	80	0.08	0.41	0.42	0.09	-
	100	0.02	0.26	0.46	0.24	0.02
	120	-	0.09	0.41	0.41	0.09
	140	-	0.02	0.25	0.47	0.26

Note: darker shading indicates larger values

The largest gains are for the low-ability children of medium- or high-ability parents, who have wages more than five percent higher in the optimal policy than in the mistaken planner's equilibrium.

Parents' and Children's individual utilities Utilities are higher for all individuals in the optimal policy, even for those individuals who mimick and deviate from their intended allocations in the mistaken planner's problem. To measure the utility gains for children, we calculate the percent by which consumption would need to be increased for each child in the mistaken planner equilibrium to yield the child's utility in the optimal policy. To measure the gains for parents, we calculate the analogous percent by which the parent's consumption and the consumption of its potential offspring would have to be raised. The results are shown in the following table:

Table 2: Consumption-equivalent gain in individual utilities from optimal policy vs. mistaken planner equilibrium

		Parent's natural ability					
		60	80	100	120	140	
Parent's utility gain		8.9%	5.8%	4.7%	5.7%	4.5%	
		Child's natural ability					
Child's utility gain		60	80	100	120	140	*Weighted average*
Parent's natural ability	60	13%	12%	13%	14%	n/a	*13%*
	80	5%	6%	5%	4%	n/a	*5%*
	100	11%	4%	3%	3%	3%	*3%*
	120	n/a	3%	4%	4%	4%	*4%*
	140	n/a	24%	9%	7%	7%	*8%*

Notes: n/a indicates no individuals of this type; darker shading indicates larger values

The largest gain is for the child with natural ability 80 and parents with high natural ability (140). These children receive almost no redistribution in the mistaken planner's equilibrium, getting a net transfer worth 10 percent of income. In contrast, their net average tax rate is -248 percent in the optimal policy

Welfare Finally, we calculate the gain to overall social welfare of going from the mistaken planner's equilibrium to the optimal policy. Again, we calculate the percent increase in consumption (now of all individuals) that would generate the same welfare in the mistaken planner's equilibrium as in the optimal policy. The answer is 5.4% of total consumption.

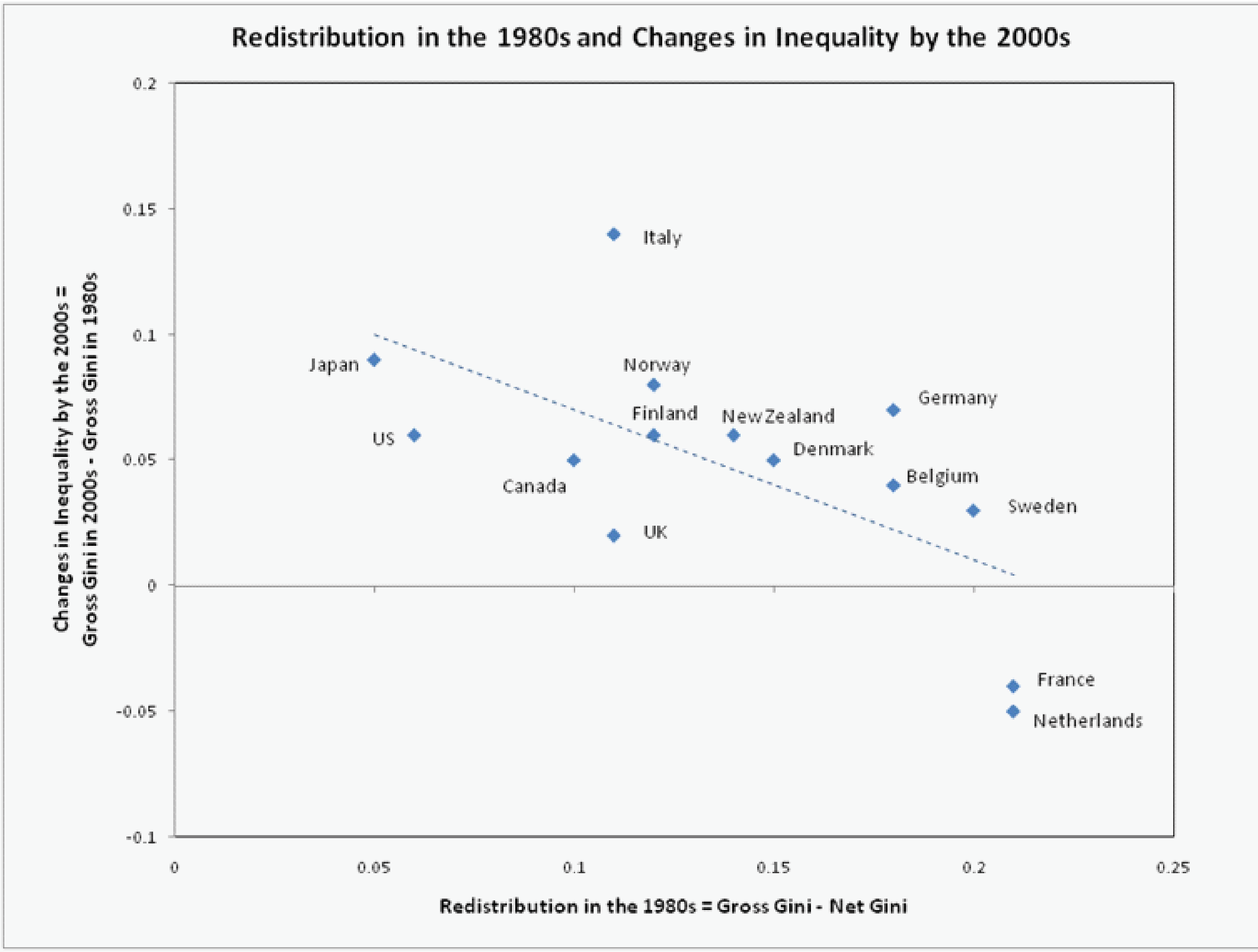
3 Empirical Test of the Model

This paper's model has a testable implication: all else the same, more redistributive policy ought to lead, about a generation later, to less inequality in gross earnings. Using data from the OECD covering the 1980s through today, this section looks for such a relationship.

A different but related empirical relationship has recently been documented by Guvenen, Kuruscu, and Ozkan (2009). These authors use a sample of eight OECD countries (U.S., U.K., France, Gemany, Netherlands, Denmark, Sweden, and Finland) and document that the more progressive a country's tax and transfer programs in 2003, the smaller the change in the log 90-50 and the change in the log 50-10 wage differentials from 1980 to 2003. The authors interpret this as evidence for their model, in which tax progressivity has a negative effect on the incentive to accumulate human capital, and progressivity thereby dulls the rise in wage dispersion in the presence of skill-biased technical change. Our empirical methods differ in the sample of countries, the measures of progressivity, the outcome variable, and the control variables. Most important, we focus on the effect of progressivity on subsequent inequality, while these authors' evidence speaks to the effect of future (expected) progressivity on the evolution of inequality.

To measure the extent of redistribution in a country, we use the difference between the pre-tax and transfer Gini coefficient and the after-tax and transfer Gini coefficient. We have data on these Gini coefficients from the mid-1980s to the mid-2000s for 14 of the 30 current OECD countries.[9] The model's prediction is that countries with greater redistribution early in the sample will have smaller increases (or greater decreases) in pre-tax and transfer inequality at the end of the sample because their low-skilled parents were given greater resources with which to raise their children. To measure the change in gross inequality over time, we use the change in the pre-tax and transfer Gini coefficient.

The figure below plots the change in the pre-tax and transfer Gini coefficient from the 1980s through the 2000s against the starting gap between the gross income and net income Ginis. The strong negative relationship is consistent with the theoretical model as well as with the micro evidence cited above.



The bivariate relationship shown in the figure is statistically significant at the 2 percent level.

Many other factors could be causing this relationship, however, so to gauge the economic significance of this relationship it is important to control for other natural explanatory variables. Foremost among these is the level of redistribution in the mid-2000s, as we might expect more redistributive policy to be persistent

[9] A fifteenth country, Portugal, has data for the 1970s and 1990s. Including it yields very similar results, as does including all countries' observations over the 1970s-1990s timeframe.

across time and to discourage effort among high earners and thereby lower gross inequality independent of the level of redistribution in the 1980s. We use the same measure as the independent variable above, but for the latter time period, to measure the level of redistribution in the 2000s. We also control for a direct measure of the marginal distortions to effort at the top of the income distribution: the top marginal income tax rate. Three other standard control variables for cross-country analyses, the log of PPP-adjusted GDP per capita, growth in that measure of output, and the extent of ethnic heterogeneity are also included.

As the following table shows, adding these controls strengthens both the statistical and economic significance of the relationship between redistribution and gross inequality approximately a generation later.

Redistribution in the 1980s and Changes in Inequality by the 2000s

	Case 1: no controls	Case 2: current redistribution controls	Case 3: all controls
Starting level of redistribution (1980s)	**-0.60**	**-1.06**	**-1.04**
	0.21	*0.20*	*0.24*
Ending level of redistribution (2000s)		**0.87**	**0.83**
		0.22	*0.27*
Top marginal income tax rate (2005)		-1.09E-03	-1.21E-03
		8.68E-04	*1.15E-03*
Growth in Log PPP real GDP per capita, 1985-2005			-1.13
			3.67
Log PPP-adjusted real GDP per capita, 2000			-9.61E-07
			2.87E-06
Ethnic fractionalization[1]			1.08E-03
			5.10E-02
Adjusted R-squared	*0.36*	*0.74*	*0.65*
Observations	*14*	*14*	*14*

Notes: Data from OECD unless noted. Standard errors shown in italics below coefficient estimates. The dependent variable is the change in the pre-tax and transfer Gini from the mid-1980s to the mid-2000s. ***Bold*** *indicates significance at the 5% level or lower.*

1: Source: Alesina et al. (2003), date varies by country

In particular, including measures of the extent of redistribution in the 2000s substantially increases the size of the (negative) relationship between redistribution in the 1980s and gross inequality in the 2000s. It also increases the statistical significance of our estimate of that relationship.

4 Conclusion

A key question in the design of taxation is how to take into account current policy's effects on the constraints facing future policy. In this paper, we explore one component of that question: the effect of parents' after-tax income on the income-earning abilities of their children. We focus on three contributions.

First, we analytically characterize how the endogenous ability distribution of a subsequent generation affects tax policy toward the current generation. We show the key forces pushing toward more current redistribution (i.e., it makes the future tax problem easier) and those pushing toward less. We also develop a formalization of a mistaken social planner that allows us to compare policy made by an optimal planner to that made by a planner who ignores the effects of parents' resources on children's ability. This methodological step may prove useful in other contexts.

Second, we quantitatively explore, using a specification of the general model guided by existing research, the optimal policy response to endogenous children's ability. We show that optimal policy uses lower average tax rates than the mistaken policy and boosts all children's wages. Moreover, the optimal policy generates a sizeable welfare gain of 5.4% of total consumption, producing gains for every type of individual.

Finally, we add to the micro evidence on the effects of parental resources suggestive evidence at a macro, cross-country level that redistribution in a current generation yields lower pre-tax inequality in the subsequent generation.

Together, these results suggest that the effects of current policies on the future distribution of productive ability ought to, and perhaps does, substantially influence the design of tax policy.

References

[1] Albanesi, Stefania and Christopher Sleet, (2006). "Dynamic Optimal Taxation with Private Information," *Review of Economic Studies* 73, pp. 1-30.

[2] Anderberg, D., (2009). "Optimal policy and the risk properties of human capital reconsidered." *Journal of Public Economics* 93.

[3] Dahl, Gordon and Lance Lochner, (2008). "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." Working paper.

[4] Farhi, Emmanuel and Ivan Werning (2007). "Capital Taxation: Quantitative Explorations of the Inverse Euler Equation." MIT working paper, April.

[5] Golosov, Mikhail, Aleh Tsyvinski, and Ivan Werning (2007). "New Dynamic Public Finance: A User's Guide," *NBER Macroannual* 2006.

[6] Golosov, Mikhail and Aleh Tsyvinski (2006). "Designing optimal disability insurance: A case for asset testing," *Journal of Political Economy.*

[7] Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski (2003). "Optimal Indirect and Capital Taxation," *Review of Economic Studies* 70, pp. 569-587.

[8] Grochulski, Borys and Tomasz Piskorski (2006). "Risky Human Capital and Deferred Capital Income Taxation." FRB of Richmond Working Paper 06-13.

[9] Kapicka, Marek, (2006). "Optimal Income Taxation with Human Capital Accumulation and Limited Recordkeeping," *Review of Economic Dynamics,* 9(4), October.

[10] Kocherlakota, Narayana (2005). "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica* 73(5), September.

[11] Milligan, Kevin and Mark Stabile, (2008). "Do child tax benefits affect the wellbeing of children? Evidence from Canadian Child Benefit expansions." NBER Working Paper 14624.

[12] Mirrlees, J.A., (1971). "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies* 38, 175-208.

[13] Paxson, Christina and Norbert Schady, (2007). "Does Money Matter? The Effects of Cash Transfers on Child Health and Development in Rural Ecuador." *Economic Development and Cultural Change*, forthcoming, working paper version 2007.

[14] Weinzierl, Matthew (2009). "The Surprising Power of Age-Dependent Taxes." Working paper.