National Tax Journal, March 2016, 69 (1), 11–40 http://dx.doi.org/10.17310/ntj.2016.1.01

OPTIMAL TAXATION WHEN CHILDREN'S ABILITIES DEPEND ON PARENTS' RESOURCES

Alexander Gelber and Matthew Weinzierl

Empirical research suggests that parents, and therefore the tax policy that affects them, can have a significant effect on their children's future earnings abilities. We take a first step toward characterizing how this intergenerational link matters for tax policy design. We find that the utilitarian welfare-maximizing policy in this context would be more redistributive toward low-income parents than under current U.S. tax policy. The additional income under such a policy would increase the probability that low-income children move up the economic ladder, and we estimate that it would generate an aggregate welfare gain equivalent to 1.75 percent of lifetime consumption.

Keywords: optimal taxation, family economics, intergenerational transmission of ability

JEL Code: H21

I. INTRODUCTION

Economists have long recognized that parents' resources may be a key determinant of their children's outcomes. Recent evidence shows, in particular, that increasing the disposable incomes of poor parents raises the performance of their children on tests of cognitive ability (Dahl and Lochner, 2012; Milligan and Stabile, 2011; Paxson and Schady, 2007; Akee et al., 2010; Løken, Mogstad, and Wiswall, 2012; Macours, Schady, and Vakis, 2012).[1] Moreover, this literature generally finds larger positive impacts of parents' resources on children's ability among lower-income families. For example, as our paper will explore in detail, Dahl and Lochner (2012) find that a $1,000 increase in family income raises math and reading test scores by about 6 percent of a standard

[1] Early work includes Leibowitz (1974), Becker and Tomes (1976), Becker (1981) and Becker and Tomes (1986).

Alexander Gelber: UC Berkeley Goldman School of Public Policy and National Bureau of Economic Research, Berkeley, CA, USA (agelber@nber.org)

Matthew Weinzierl: Harvard Business School and National Bureau of Economic Research, Boston, MA, USA (mweinzierl@hbs.edu)

deviation, with larger impacts in more disadvantaged families. How should tax policy respond to these findings? In this paper, we take a first step toward answering this question both qualitatively and quantitatively.[2]

We start with a formal model of the implications of this intergenerational link for tax policy design. We modify the standard modern optimal tax model to allow a child's income-earning ability to depend on three components: parental ability, parental disposable income, and a random shock. This mix of factors helps us to capture the complex process by which society's ability distribution is determined. Using that model, we derive analytical expressions for tax policy both within and across generations that would maximize utilitarian social welfare, that is, an unweighted aggregation of individual utility levels. These expressions make clear the new forces acting on tax design in this context. Within a generation, lower marginal tax rates at low incomes will yield gains to the extent that the marginal effect of resources on children's ability is relatively large for families with low incomes or low parental earnings ability. However, they will also yield losses to the extent that they reduce the incentives for individuals to act to increase the likelihood that their descendants will have high ability. Across generations, resources should be allocated to equalize the cost of raising welfare, taking into account not only the marginal utilities of individuals in each generation (as in a conventional setting) but also the effects of resources on prior generations' incentives and future generations' tax payments and utility levels.

We then take the model to data and use it to quantify the benefits society can gain by having a tax policy that takes advantage of this intergenerational link. We choose values for the model's key parameters by having the model's output (under existing U.S. tax policy) approximate new estimates of the transmission of ability across generations in the United States that we generate using data on parents and their children from the National Longitudinal Survey of Youth (NLSY) and Children of the National Longitudinal Survey of Youth (CNLSY). Next, we compute the utilitarian welfare-maximizing policy and find that it redistributes substantially more toward low-ability parents and earlier generations than does current policy. We find that these changes cause average ability to rise across generations, yielding a welfare gain that would be equivalent to a 1.75 percent permanent increase in aggregate disposable income in our baseline case. The gains are larger for lower-income families than for higher-income families.

The paper proceeds as follows. Section II lays out the model of the tax design problem, while Section III derives analytical conditions on utilitarian welfare-maximizing policy both within and across generations. Section IV generates new empirical estimates of the transmission of ability across generations and uses them to choose parameter values for the model. Section V uses that model to simulate and describe the effects of the

---

[2] Farhi and Werning (2010, p. 635) characterize optimal taxation across generations, noting in their opening sentences that "[o]ne of the biggest risks in life is the family one is born into. We partly inherit the luck, good or bad, of our parents through the wealth they accumulate." Their important analysis, however, largely assumed that children's skills are independent of their parents' abilities and economic resources. We take up the complementary analysis.

utilitarian welfare-maximizing tax policy on the evolution of the ability distribution, and it calculates the welfare implications of those effects. Section VI concludes. A lengthy online appendix, which follows the working paper version of this paper (Gelber and Weinzierl, 2012), contains details of the results, the robustness checks, the empirical strategy, and the relationship to prior literature.

II. MODEL

Our formal analysis starts with the standard dynamic optimal tax model based on Mirrlees (1971) and developed in Golosov, Kocherlakota, and Tsyvinski (2003), Kocherlakota (2005), Golosov and Tsyvinski (2006), Albanesi and Sleet (2006), Golosov, Tsyvinski, and Werning (2007), Farhi and Werning (2010, 2013), and Weinzierl (2011).[3] Individuals obtain utility U from consumption c — equivalently, disposable income — and disutility from exerting work effort l. Each individual has an unobservable "type" w, which measures the ability to earn pre-tax income y and is taken from a fixed set of possible values indexed by $i \in \{1,2,...,I\}$. The product of individual ability and work effort, both of which are unobservable, is the observable pre-tax income. The tax authority levies taxes as a function of pre-tax income to maximize a utilitarian social welfare function, which we define as the total present-value utility of the population of families starting from the first generation and ending in generation T. The tax authority is constrained in two ways: disposable income must be funded by output (feasibility), and individuals maximize their own utility taking the tax system as given (incentive compatibility). We also impose the constraint that taxes may depend on only the current generation's characteristics and choices, a restriction we believe is appropriate given the strong aversion in real-world policy to having children's taxes depend on their parents' outcomes.

The intergenerational focus of this paper requires some additional structure. Individuals are linked in families, with one individual per generation in each family. Generations are indexed by $t \in \{1,2,...,T\}$, where the choice of T is immaterial to the results below. The distribution of individuals across types is taken as given in the first generation, but in subsequent generations it is a function of the distribution of disposable income in the previous generation as well as of the inheritance of type. Formally, denote with $p^j\left(w_t^i, c_t^{i'}\right)$ the probability that an individual of generation $t+1$ is of type j, given that her parent (in generation t) was of type i and had disposable income of type i'. We assume no intergenerational transfers, so all disposable income is consumed within each generation. While bequests have interesting implications for the optimal taxation of testators and inheritors (Piketty and Saez, 2013; Kopczuk, 2013), they are highly concentrated among the very wealthy. Thus, individuals from most of the ability distribution will be unaffected by our assumption that there is no bequest motive, and even those who would be affected are already wealthy enough that marginal changes to parental resources are

[3] For work incorporating human capital accumulation, see Kapicka (2006a, 2006b), Grochulski and Piskorski (2010), Anderberg (2009), and Stantcheva (2015).

immaterial. As is common in optimal tax analyses, we translate the tax system set up by the tax authority into a menu of allocations of pre-tax and disposable income. Formally, the tax authority's problem is

Problem 1. Tax Authority's Problem

The tax authority chooses c_t^i, y_t^i to

$$\max_{\{c_t^i, y_t^i\}_{t=1,i=1}^{T,I}} \sum_i p^i U_1^i \tag{1}$$

subject to feasibility

$$\sum_i p^i R_1^i \geq \tilde{R} \tag{2}$$

and incentive compatibility

$$U_t^i \geq U_t^{i'|i} \text{ for all } t \text{ and } i, i', \text{ where} \tag{3}$$

$$U_t^{i'|i} = u\left(c_t^{i'}\right) - v\left(\frac{y_t^{i'}}{w_t^i}\right) + \beta \sum\nolimits_{j=1}^{I} p^j \left(w_t^i, c_t^{i'}\right) U_{t+1}^j,\ U_t^i = U_t^{i|i},$$

$$R_t^i = \left(y_t^i - c_t^i\right) + \beta \sum\nolimits_j p^j \left(w_t^i, c_t^i\right) R_{t+1}^j,$$

u represents the dependence of utility U on consumption c, v represents the dependence of utility U on hours worked l (equal to income y divided by ability w), β is a discount factor, and $\tilde{R}$ is an exogenous revenue requirement.

We make a number of simplifying assumptions. First, only tax policy is modeled in this paper, but that does not imply that other policies play no role in reality. Our empirical estimates take as given the existing set of non-tax policies and institutions, such as public education or related policies, including education grants, vouchers, loans, and subsidies, that have effects on children's abilities (including effects that may interact with the tax system). Our model implicitly assumes that these policies and institutions are held constant as taxes vary.

We restrict our model's instruments in this way for two reasons. First, it matches the empirical evidence we use to calibrate the model, which also is based on variation in tax policy holding other policies fixed. Second, the question of how tax policy can affect the ability distribution is of independent interest, both because in practice partial reforms to policy (e.g., to taxes only) are common and because a household's after-tax and transfer resources may have a direct effect on children.

At the least, we should view the tax policy component as a complement to the education policy component. Of course, broader policy reforms could achieve even greater gains, so our estimates of the welfare gains from changes to taxes can be considered a lower bound on the potential for a holistic optimal policy. Such an optimal policy would also take into account the potentially complex interactions between optimal tax policy, optimal educational policy, and other programs, and we hope the results of this paper encourage future work on these important questions.

The importance of the tax policy component relative to the education policy component — or other complementary components such as commodity taxes — may depend on the channels that link parental resources and children's abilities. For example, if parental resources matter for children's abilities primarily because the income elasticity of parental investments in schooling is positive, then targeted policies such as direct government expenditures on schooling could be more effective than the more indirect route of tax policy, as some of the decreased tax burden could be spent on items other than schooling investment. However, if greater parental income leads to better outcomes for children through a route that is less similar to feasible government policies — e.g., parental income leads to better outcomes for the children because it reduces stress in the household — then tax policy could be a preferable approach for shaping children's outcomes. Future empirical work that can observe such channels could help clarify these issues.

A second simplifying assumption is that the formulation above uses the same measure of parental resources as the quantity of consumption in the parent's utility function and the input to the child's ability production function. However, we are not asserting that the way in which parental disposable income is used is irrelevant to their child's ability. Rather, we are guided not only by tractability but also by the data. Our empirical evidence concerns the effect on a child's ability of transfers received by her parents; we have no data on how those transfers were allocated. In order to calibrate to this evidence, our model must also leave the allocation of these transfers unspecified. We use the term disposable income, rather than consumption, to make this aspect of our analysis clear.

In principle, one could attempt to use the limited available data on the division of parental expenditure into consumption and investment in children's abilities in order to model more subtle optimal policies. Identifying the separate effects of these categories of expenditure on children's ability would not be possible using our data and identification strategy, as studying the causal effect of the EITC on different categories of consumption is severely limited by empirical power issues (Gelber and Mitchell, 2012). Moreover, the appeal of more subtle policy that distinguishes between these categories would be diminished by incentives for (largely unobservable) misreporting of spending across categories. Similarly, in a dynamic framework one could model how consumption and savings over different periods are affected by taxes across periods. The empirical implementation of this would be difficult, however, as the intertemporal substitutability of consumption (i.e., the intertemporal elasticity of substitution) is the topic of much literature that is not considered settled, with elasticity estimates ranging from close to zero (Hall, 1988; Dynan, 1993) to over two (Gruber, 2006). In general, it is possible to argue that the welfare gains we calculate reflect a lower bound on the gains if separate taxation of different types of consumption were modeled.

A third assumption in the problem above is that the allocation of parental time has no effect on children's abilities. Later in the paper, we extensively explore the case in which children's ability depends on both parental income and parents' hours worked. If parents work more, they are likely to spend less time with their children. This in principle could either worsen children's outcomes (if, say, parents teach children skills in their non-work time) or could improve children's outcomes (if, say, parents' increased work serves as a role model for children's effort in school). While our empirical strategy strains the available data in this case, such that the results are best seen as suggestive, we find that on net allowing for a role of parental time allocation has only a small effect on our baseline results.

Finally, we do not constrain parent and child distributions of ability to be the same, as they might be in some steady state, in order to acknowledge that wage distributions and test scores have shown secular time trends in the data (Goldin and Katz, 2007; Flynn, 1987).

III. ANALYSIS OF UTILITARIAN OPTIMAL POLICY

Our analysis of the tax authority's problem in (1) through (3) generates two analytical results on utilitarian welfare-maximizing policy. First, we characterize the distortion to an individual's choice of how much to earn, that is, the marginal tax rate on income. Second, we derive a necessary condition on optimal allocations across generations that shows the effects of this intergenerational link in an intuitive but powerful way.

A. Marginal Tax Rates

In the absence of taxes, parent i in generation t would choose how much income to earn to maximize her own utility subject to a personal feasibility constraint and given the expectation that her descendants, whose abilities are determined by the production function $p^j\left(w_t^i, c_t^i\right)$, would also choose optimally for themselves. That parent's optimal private choice would satisfy

$$(4) \qquad \frac{v'\left(y_t^i / w_t^i\right)}{w_t^i u'\left(c_t^i\right)} = 1 + \beta \sum_j \frac{\partial p^j\left(w_t^i, c_t^i\right)}{\partial c_t^i} \frac{U_{t+1}^j}{u'\left(c_t^i\right)}.$$

To interpret (4), note that $v'\left(y_t^i / w_t^i\right) / \left[w_t^i u'\left(c_t^i\right)\right]$ is the ratio of the marginal disutility of labor to the marginal utility of consuming the income that labor earns. In a setting without endogenous child ability, this ratio equals one in the absence of taxes, and is less than one if the individual faces a positive marginal tax rate. In contrast, (4) implies that in this no-tax setting parents set this ratio *greater* than one. That is, parents take into account the effect of their disposable income on their child's ability, so they will appear to choose labor supply as though there were a marginal subsidy equal to the second term on the right-hand side of (4), relative to a model in which they took only their own disposable income into account.

Lemma 1 shows that the tax authority distorts a parent's private choice.

Lemma 1. Intratemporal Distortion

Let $\mu_t^{i'|i}$ denote the multiplier on (3). The solution to the tax authority's problem satisfies, for all $t \in [1,2,...,T]$ and all $j \in [1,2,...,I]$,

$$\text{(5)} \quad \frac{v'\left(y_t^j / w_t^j\right)}{w_t^j u'\left(c_t^j\right)} = A_t^j \frac{\left(B_t^j + C_t^j\right)}{\left(B_t^j + D_t^j\right)} \left(1 + \beta \sum_k \frac{\partial p^k\left(w_t^j, c_t^j\right)}{\partial c_t^j} \frac{U_{t+1}^k}{u'\left(c_t^j\right)}\right),$$

where the expressions for A_t^j, B_t^j, C_t^j, and D_t^j are as follows:

$$\text{(6)} \quad A_t^j = \frac{1}{1 - \beta \sum_k \frac{\partial p^k\left(w_t^j, c_t^j\right)}{\partial c_t^j} R_{t+1}^k}, \quad B_t^j = \beta^t \pi_t^j + \sum_{\tau=1}^{t-1} \beta^{t-\tau} \sum_i \sum_{i'} \mu_\tau^{i'|i} \left(\pi_t^j |_{c_\tau^i} - \pi_t^j |_{c_\tau^{i'}}\right)$$

$$\text{(7)} \quad C_t^j = \sum_{j'} \mu_t^{j'|j} - \sum_{j'} \mu_t^{j|j'} \frac{1 + \beta \sum_k \frac{\partial p^k\left(w_t^{j'}, c_t^j\right)}{\partial c_t^j} \frac{U_{t+1}^k}{u'\left(c_t^j\right)}}{1 + \beta \sum_k \frac{\partial p^k\left(w_t^j, c_t^j\right)}{\partial c_t^j} \frac{U_{t+1}^k}{u'\left(c_t^j\right)}},$$

$$D_t^j = \sum_{j'} \mu_t^{j'|j} - \sum_{j'} \mu_t^{j|j'} \frac{\frac{1}{w_t^{j'}} v'\left(\frac{y_t^j}{w_t^{j'}}\right)}{\frac{1}{w_t^j} v'\left(\frac{y_t^j}{w_t^j}\right)},$$

where $\pi_t^j |_{c_\tau^i}$ is the probability that a generation t descendant of parent type i from generation τ is of type j and $\sum_i \pi_t^j |_{c_\tau^i}$ is denoted by the unconditional probability π_t^j.

Lemma 1 shows that the product $A_t^j \left(B_t^j + C_t^j\right) / \left(B_t^j + D_t^j\right)$ is the optimal wedge distorting the parent's choice of earned income. To understand the determinants of this optimal wedge, we decompose it into three effects.

The first is the "revenue effect." Suppose the tax authority had full information on individuals' (endogenous) abilities. In that case, $B_t^j = \beta^t \pi_t^j$ and $C_t^j = D_t^j = 0$. Then, the wedge would be simply A_t^j, the value of which depends on the present value of the expected net revenue gain from raising the disposable income of parent *j*. The tax authority values that revenue gain, while the parent does not, so the revenue effect implies that a smaller downward distortion (or larger upward distortion) to parent *j*'s effort is optimal. To the extent that the marginal effect of additional parental disposable income on children's abilities is larger for lower-ability parents, this revenue effect will be greater at low incomes, and smaller marginal taxes at low incomes will be optimal.

The second is the "relative return effect." Suppose the tax authority cannot observe ability, but parental resources have no effect on children's abilities. In that case, $A_t^j = 1$, and $B_t^j = \beta^t \pi_t^j$, and $C_t^j = \sum_{j'} \mu_t^{j'|j} - \sum_{j'} \mu_t^{j|j'}$, the optimal distortion is driven by binding incentive constraints in the current generation. Introducing *endogenous* ability changes C_t^j because its value depends on how the inheritance of ability and parental resources interact in the production of child ability. In particular, including endogenous ability will increase C_t^j if the marginal effect of additional parental disposable income on children's abilities is larger for lower-ability parents. Meanwhile, D_t^j is unchanged, so C_t^j and D_t^j have more similar values and the optimal marginal taxes on low-skilled parents are smaller (note that $C_t^j < D_t^j$ when higher-skilled types are tempted to mimic lower-skilled types). Intuitively, if the extra resources redistributed to low-income families are of less value to high-ability parents because the intergenerational effect on abilities is smaller, then high-income parents will be less tempted in this model to claim low ability, and smaller marginal tax rates at low incomes will be optimal.

The third is the "ancestor incentive effect." Introducing endogenous ability also changes B_t^j, which measures how an increase in c_t^j affects the incentives of earlier generations to increase the likelihood that their descendants have the type *j*. All else equal, a smaller marginal tax rate for a low-skilled type raises the temptation for previous generations to work less and accept a higher probability of low-skilled descendants. In this way, the ancestor incentive effect acts to offset the revenue and relative return effects, providing a rationale for larger marginal taxes at low incomes.

In the end, the net effect on utilitarian welfare-maximizing marginal tax rates depends on the relative strength of these three effects.

B. Allocations across Generations

In the absence of the intergenerational link studied in this paper, the standard condition on utilitarian welfare-maximizing allocations across generations is

$$(8) \quad \sum_j \frac{\pi_t^j}{u'\left(c_t^j\right)} = \sum_k \frac{\pi_{t+1}^k}{u'\left(c_{t+1}^k\right)}.$$

This condition, parallel to the "Symmetric Inverse Euler Equation" in Weinzierl (2011), shows that the optimal allocation equalizes the cost, in disposable income units, of raising social welfare across generations. A version of it also applies to optimal tagging (Mankiw and Weinzierl, 2010).

With endogenous ability, a modified version of (8) applies, which we state in the following proposition.

Proposition 1. Solution to the Tax Authority's Problem

The solution to the tax authority's problem satisfies

$$\text{(9)}\quad \frac{1}{\Lambda_t}\sum_j \frac{\pi_t^j}{u'\left(c_t^j\right)} \frac{1-\beta\sum_k \frac{\partial p^k\left(w_t^j, c_t^j\right)}{\partial c_t^j} R_{t+1}^k}{1+\beta\sum_k \frac{\partial p^k\left(w_t^j, c_t^j\right)}{\partial c_t^j}\frac{U_{t+1}^k}{u'\left(c_t^j\right)}}$$

$$= \frac{1}{\Lambda_{t+1}}\sum_k \frac{\pi_{t+1}^k}{u'\left(c_{t+1}^k\right)} \frac{1-\beta\sum_l \frac{\partial p^l\left(w_{t+1}^k, c_{t+1}^k\right)}{\partial c_{t+1}^k} R_{t+2}^l}{1+\beta\sum_l \frac{\partial p^l\left(w_{t+1}^k, c_{t+1}^k\right)}{\partial c_{t+1}^k}\frac{U_{t+2}^l}{u'\left(c_{t+1}^k\right)}},$$

where Λ_t is defined as

$$\Lambda_t = 1 + \sum_{\tau=1}^{t-1}\frac{\beta^{t-\tau}}{\beta^t}\sum_i\sum_{i'}\mu_\tau^{i'|i}\sum_k\left(\pi_t^k\mid_{c_\tau^i} - \pi_t^k\mid_{c_\tau^{i'}}\right)$$

$$+\frac{\beta}{\beta^t}\sum_k\sum_{k'}\mu_t^{k|k'} \frac{\sum_j \frac{\partial p^j\left(w_t^k, c_t^k\right)}{\partial c_t^k}\frac{U_{t+1}^j}{u'\left(c_t^k\right)} - \sum_j \frac{\partial p^j\left(w_t^{k'}, c_t^k\right)}{\partial c_t^k}\frac{U_{t+1}^j}{u'\left(c_t^k\right)}}{1+\beta\sum_j \frac{\partial p^j\left(w_t^k, c_t^k\right)}{\partial c_t^k}\frac{U_{t+1}^j}{u'\left(c_t^k\right)}}.$$

To understand Proposition 1, recall that (8) implies that the cost of raising social welfare through transfers to one generation must be the same for all generations. Proposition 1 is the same condition, but in the more complicated context of this model. To help with its interpretation, we decompose it into three effects of making a transfer to a given generation *t*.

First, if the transfer to generation *t* raises individual *j*'s investment in her children's abilities resulting in increased tax revenue from future generations, these revenue gains act to offset the costs of the transfer. Formally, this factor is captured in the expression $\beta\sum_k\left[\partial p^k\left(w_t^j,c_t^j\right)/\partial c_t^j\right]R_{t+1}^k$, which is closely related to the revenue effect identified in the discussion of Lemma 1. This expression is the present value of the net change in future taxes paid by individual *j*'s children when c_t^j increases.

Second, under the assumption that the transfer to individual *j* raises her investment in her children's abilities resulting in increased utilities for future generations, these welfare gains augment any direct changes in utility from the transfer. Formally, this factor is captured in the expression $\beta\sum_k\left[\partial p^k\left(w_t^j,c_t^j\right)/\partial c_t^j\right]\left[U_{t+1}^k/u'\left(c_t^j\right)\right]$. This expression is the present value, per additional unit of utility for individual *j*, of the increase in utility enjoyed by individual *j*'s children when c_t^j increases.

Third, both the relative return and ancestor incentive effects from the discussion of Lemma 1 also matter for intertemporal allocations through their effects on Λ_t. The parameter $\Lambda_t = 1$ when two conditions hold: (1) the marginal effect of parental financial resources on child ability is independent of parental ability (i.e., $\partial p^k\left(w_t^j,c_t^j\right)/\partial c_t^j=\partial p^k\left(w_t^{j'},c_t^j\right)/\partial c_t^j$ for all j, j'); and (2) incentive constraints do not bind in preceding generations (i.e., $\mu_\tau^{i'|i}=0$ for all τ, i, i'). If, instead, transfers to a generation relax incentive constraints and increase the investment of low-income parents in their children, the expression Λ_t is less than one. Similarly, if transfers to a generation relax incentive constraints that bind on ancestors whose offspring are relatively common in the recipient generation, the expression Λ_t is less than one. The smaller is Λ_t, the larger is the optimal transfer to generation *t*.

To build intuition for how these effects combine, suppose that average ability is stable over time and the effects of parental resources on a child's ability are largest at lower skill levels (formally, $\sum_{k=1}^K\pi_t^k=\sum_{k=1}^K\pi_{t+1}^k$ for all $K \leq 1$ and $\left|\partial p^j\left(w_t^{k+1},c_t^{k+1}\right)/\partial c_t^\kappa\right|\leq\left|\partial p^j\left(w_t^k,c_t^k\right)/\partial c_t^k\right|$ for $\kappa > k$). A policy neglecting endogenous child ability would satisfy (8) and treat generations symmetrically. In contrast, an alternative policy that would satisfy (9) would transfer resources from generation $t+1$ to low-ability workers in generation *t*. Such a policy would raise the average inverse marginal utility of consumption in generation *t* but also increase the population proportion of higher-ability workers in generation $t + 1$, putting greater weight on workers with larger inverse marginal utilities of disposable income and smaller gains in future revenue and utility for their descendants from marginal resources. As a result, the cost of raising social welfare in generations *t* and $t + 1$ would rise together. In other words, transfers from future to earlier generations may generate gains for all generations: early generations from having higher disposable incomes and future generations from having improved ability distributions. While (9) does not prove that a policy improvement such as in this example exists, our data and simulations below show that this is a realistic scenario for the United States.

C. Labor as an Input to Children's Ability

As we discuss above, it is also possible that children's ability could depend on parents' hours worked. In this case, the planner's problem includes the dependence of $p(\cdot)$ on labor effort (or, equivalently, time not devoted to labor effort).

Problem 2. Planner's Problem with Parental Labor as an Input to Child Ability

The planner chooses c_t^i, y_t^i to

$$(10)\quad \max_{\{c_t^i, y_t^i\}_{t=1,i=1}^{T,I}} \sum_i p^i U_1^i,$$

where

$$(11)\quad U_t^i = u\left(c_t^i\right) - v\left(\frac{y_t^i}{w_t^i}\right) + \beta \sum_{j=1}^{I} p^j\left(w_t^i, c_t^i, \frac{y_t^i}{w_t^i}\right) U_{t+1}^j.$$

This is maximized subject to feasibility

$$(12)\quad \sum_i p^i R_1^i \geq \tilde{R},$$

where $\tilde{R}$ is an exogenous revenue requirement, and

$$(13)\quad \sum_i p^i \left[\left(y_t^i - c_t^i\right) + \beta \sum_j p^j\left(w_t^i, c_t^i, \frac{y_t^i}{w_t^i}\right) R_{t+1}^j\right] \geq \tilde{R},$$

and incentive compatibility for each generation

$$(14)\quad U_t^i \geq U_t^{i'|i} \text{ for all generations } t \text{ and types } i, i',$$

where $U_t^{i'|i}$ denotes the utility obtained by an individual of type i when claiming to be type i'

$$(15)\quad U_t^{i'|i} = u\left(c_t^{i'}\right) - v\left(\frac{y_t^{i'}}{w_t^i}\right) + \beta \sum_{j=1}^{I} p^j\left(w_t^i, c_t^{i'}, \frac{y_t^{i'}}{w_t^i}\right) U_{t+1}^j.$$

Simplifying as in the model without labor effort in the ability production function, we obtain the following lemma, which is analogous to Lemma 1.

Lemma 2. Intratemporal Distortion with Parental Labor as an Input to Child Ability

Let $U_t^{i'|i}$ denote the multiplier on (14). The solution to the Planner's Problem with Parental Labor as an Input to Child Ability satisfies, for all $t \in [1,2,...,T]$ and all $j \in [1,2,...,I]$,

$$\frac{v'\left(\frac{y_t^k}{w_t^k}\right)}{w_t^k u'(c_t^k)} = A_t^j \frac{\left(B_t^j + C_t^j\right)}{\left(B_t^j + D_t^j\right)} \frac{\left(1+\beta\sum_j \frac{\partial p^j\left(w_t^k, c_t^k, \frac{y_t^k}{w_t^k}\right)}{\partial c_t^k} \frac{U_{t+1}^j}{u'\left(c_t^k\right)}\right)}{\left(1+\beta\sum_j \frac{\frac{1}{w_t^k}\partial p^j\left(w_t^k, c_t^k, \frac{y_t^k}{w_t^k}\right)}{\partial y_t^k} \frac{U_{t+1}^j}{\frac{1}{w_t^k} v'\left(\frac{y_t^k}{w_t^k}\right)}\right)},$$

where

$$A_t^j = \frac{1+\beta\sum_j \frac{\frac{1}{w_t^k}\partial p^j\left(w_t^k, c_t^k, \frac{y_t^k}{w_t^k}\right)}{\partial y_t^k} R_{t+1}^j}{1-\beta\sum_k \frac{\partial p^k\left(w_t^j, c_t^j, \frac{y_t^j}{w_t^j}\right)}{\partial c_t^j} R_{t+1}^k},$$

$$B_t^j = \beta^t \pi_t^j + \sum_{\tau=1}^{t-1} \beta^{t-\tau} \sum_i \sum_{i'} \mu_\tau^{i'|i}\left(\pi_t^j\,|_{c_\tau^i} - \pi_t^j\,|_{c_\tau^{i'}}\right),$$

$$C_t^j = \sum_{j'} \mu_t^{j'|j} - \frac{1+\beta\sum_k \frac{\partial p^k\left(w_t^{j'}, c_t^j, \frac{y_t^j}{w_t^j}\right)}{\partial c_t^j} \frac{U_{t+1}^k}{u'\left(c_t^j\right)}}{1+\beta\sum_k \frac{\partial p^k\left(w_t^j, c_t^j, \frac{y_t^j}{w_t^j}\right)}{\partial c_t^j} \frac{U_{t+1}^k}{u'\left(c_t^j\right)}} \sum_{j'} \mu_t^{j|j'}, \text{ and}$$

$$D_t^j = \sum_{j'} \mu_t^{j'|j} - \frac{\frac{1}{w_t^{j'}} v'\left(\frac{y_t^j}{w_t^{j'}}\right)}{\frac{1}{w_t^j} v'\left(\frac{y_t^j}{w_t^j}\right)} \frac{\left(1+\beta\sum_k \frac{\frac{1}{w_t^{j'}}\partial p^k\left(w_t^{j'}, c_t^j, \frac{y_t^j}{w_t^{j'}}\right)}{\partial y_t^j} \frac{U_{t+1}^k}{\frac{1}{w_t^{j'}} v'\left(\frac{y_t^j}{w_t^{j'}}\right)}\right)}{\left(1+\beta\sum_k \frac{\frac{1}{w_t^j}\partial p^k\left(w_t^j, c_t^j, \frac{y_t^j}{w_t^j}\right)}{\partial y_t^j} \frac{U_{t+1}^k}{\frac{1}{w_t^j} v'\left(\frac{y_t^j}{w_t^j}\right)}\right)} \sum_{j'} \mu_t^{j|j'}.$$

Note that the division of the intratemporal result into a wedge and an expression equal to what the parent would choose continues to hold in this setting. The parent's optimum would now be

$$\frac{\frac{1}{w_t^i} v'\left(\frac{y_t^i}{w_t^i}\right)}{u'\left(c_t^i\right)} = \frac{1+\beta \sum_{j=1}^{I} \frac{\partial p^j\left(w_t^i, c_t^i, \frac{y_t^j}{w_t^j}\right)}{\partial c_t^i} \frac{U_{t+1}^j}{u'\left(c_t^i\right)}}{1+\beta \sum_{j=1}^{I} \frac{1}{w_t^j} \frac{\partial p^j\left(w_t^i, c_t^i, \frac{y_t^j}{w_t^j}\right)}{\partial y_t^j} \frac{U_{t+1}^j}{\frac{1}{w_t^i} v'\left(\frac{y_t^i}{w_t^i}\right)}}.$$

The main differences once parental labor effort enters the child ability production function are as follows (recall that parents internalize the direct effect of their time allocation on their children's abilities, as in the expression for the parental optimum above). Assume extra parental time at work is detrimental to child ability so that $\frac{1}{w_t^{j'}} \partial p^k\left(w_t^{j'}, c_t^j, \frac{y_t^j}{w_t^{j'}}\right) / \partial y_t^j < 0$. First, the term A_t^j now captures that extra parental effort will lower future revenues, so the optimal distortion discouraging labor effort is larger (i.e., the term A_t^j is smaller). Second, if extra parental time at work is more detrimental to child ability for low-ability parents, then D_t^j will be smaller. This reduces the optimal distortion of labor supply for lower skill types. Intuitively, incentive constraints are less binding with this effect, because high-skilled parents gain relatively less from the lower labor effort requirements they would enjoy if they claimed a low income.

IV. MODEL CALIBRATION UNDER EXISTING U.S. TAX POLICY

In this section, we generate new estimates of the effect of parental resources on children's ability under existing U.S. tax policy and then use those estimates to parameterize the model described in Section II.

A. Empirical Estimates of the Target Statistics

We adapt to our framework the empirical work from a recent major study of parents' taxes and children's outcomes. Dahl and Lochner (2012) study the effect of expansions of the EITC in the 1990s and other variation over this period in tax and transfer programs on children's test score outcomes.[4] Rather than calibrating our model using a cross-section of data, we use a modified version of the Dahl and Lochner empirical strategy in order to generate more credible estimates of the causal effect of parental income on child ability. We briefly describe their empirical strategy here, often borrowing from their description, but we refer readers to their paper for a full description of their

[4] See Hotz and Scholz (2003) or Eissa and Hoynes (2006) for detailed descriptions of the EITC program and a summary of related research.

empirical strategy and its motivation.[5] We note the caveat that not all papers have found important effects of transfer programs on children; for example, Jacob, Kapustin, and Ludwig (2014) find that randomly receiving a housing voucher has mostly statistically insignificant effects on children's outcomes, which are always smaller than those in recent studies of cash transfers.

Dahl and Lochner (2012) investigate how increases in parental income due to the EITC and other tax and transfer programs affect the cognitive achievement of disadvantaged children. Their estimation strategy is based on the observation that low-to-middle-income families received large increases in payments from expansions of the EITC in the late-1980s and mid-1990s, but higher-income families did not. If parental disposable income affects child ability, this disparity in the changes to disposable income should have caused an increase over time in the test scores of children from low-to-middle-income families relative to those from higher-income families. Dahl and Lochner use the CNLSY, which contains data on several thousand children matched to their mothers (from the main NLSY sample). Income and demographic measures are included in the data, in addition to as many as four repeated measures of cognitive test scores per child taken every other year. They use measures of child ability based on standardized scores on the Peabody Individual Achievement Test (PIAT), which measures oral reading ability, mathematics ability, word recognition ability, and reading comprehension. Dahl and Lochner's instrumental variables estimates suggest that a $1,000 increase in family income raises math and reading test scores by about 6 percent of a standard deviation.

We estimate a model similar to Dahl and Lochner's, designed to attain a slightly different goal. Motivated by our model above and the simulation below, we estimate the effect that income has on the probability that a parent of given ability type produces a child of a given ability type (controlling for the child's lagged ability type); details are provided in the online appendix.

One potential limitation of this approach is that EITC variation affects only lower- and lower middle-income families. However, it is important to note that our measure of tax liability is based on the NBER TAXSIM calculator, which incorporates all tax liabilities, not just those due to the EITC. At the same time, the biggest variation in income taxes over this period related to the EITC, as well as to increases in marginal tax rates for top incomes from the 1990 and 1993 reforms.[6]

In our main specification, we divide parents into four wage (ability) categories $\left\{P_i\right\}_{i=1}^{4}$ and divide children into four test score categories $\left\{C_i\right\}_{i=1}^{4}$. Each category comprises one

[5] As noted above, several papers estimate the effect of parents' income on their children's achievement levels. The marginal effect of parental resources on child ability is more elusive, in part because it is difficult to find exogenous variation in parents' disposable income and linked data on parents' and their children's outcomes. Our paper suggests that the effect of parents' disposable income on children's wages is an important topic for future research.

[6] To address this issue further, we examine the effects of analyzing higher income levels in the appendix, where we show the robustness of the results to including redistribution from still higher incomes in the calibration.

quartile of the sample distribution of wages or test scores, respectively, with subscript i indicating the quartile of the distribution, where $i = 1$ is the lowest quartile. Because there are four parent types, we estimate four separate regressions, in each of which the dependent variable is a dummy that equals one when the child has ability in the i-th category. We classify parents into wage types by ranking them according to their average wage over the full sample period.

In choosing the number of categories, we take into account competing technical and conceptual considerations. More categories will give the calibration more targets, as well as better describe the true heterogeneity of the population and therefore the potential gains from optimal policy. However, too many categories will prevent the regressions in the empirical estimation from having enough positive values of the dependent variable to yield meaningful results. Using too few categories fails to provide enough empirical targets for the calibration exercise to converge on a best set of parameter values; this factor requires us to use at least four categories.[7]

We construct the simulated model to match (i.e., minimize the sum of squared deviations from) three sets of statistics obtained from this analysis: (1) the marginal effects of parental resources on their children's abilities, (2) the ability transition matrix between generations, and (3) the expected log wage within generations. The signs of these estimates generally conform to expectations: higher parental income predominantly increases the probability that a child is high-ability (i.e., in the third or fourth quartile) and decreases the probability that a child is low-ability (i.e., in the first or second quartile), with the "correct" sign of the relationship occurring in 13 out of 16 regressions. Using the same dataset and definition of types, we next generate this transition matrix by calculating the fraction of the sample from each parent wage category that began the sample period with the child test score in each category. Finally, the expected log wage within a generation is readily calculated as the average of the log of the four ability levels. Details of these calculations are provided in the online appendix.

B. The Model's Parameter Values

We need to specify some features of the general model from Section II to generate output that can be matched to the targeted statistics.

First, we assume that the ability production function is

$$(16) \qquad E\left[\ln w_{t+1} \mid w_t^j, c_t^{j'}\right] = \alpha_a\left(\rho \ln w_t^j + (1-\rho)\ln \tilde{w}\right) + \alpha_c^j \ln c_t^{j'},$$

which implies that the child's expected ability is a function of the parent's ability, a fixed "reference" ability level, and the parent's disposable income. The child's expected ability is influenced by the parent's ability w_t^j relative to the fixed ability level $\tilde{w}$, indi-

[7] In the appendix, we also show results with five and 10 categories. Those results show heterogeneity in tax rates at a finer level of disaggregation at the cost of a substantial loss of power in the empirical estimates. Our analyses with five and 10 types yield similar results to those from our benchmark four-type model.

cating mean reversion in characteristics transmitted across generations, consistent with Haider and Solon (2006).

This log-linear functional form concisely captures the basic forces at work in this model. It allows us to adjust the relative importance of parental ability and mean ability (through ρ) and the relative importance of parental ability and parental resources (through α_a and α_c^j). Note that our measure of parental ability (i.e., the wage) implicitly includes not only innate talent but also the education and other factors that contribute to the parent's wage. A further step in the analysis would be to separate out the determinants of parental ability, including parents' own education decisions, and independently assess their importance for determining children's ability. Note that the importance of parental resources depends on parental ability through the dependence of α_c^j on j, establishing a direct connection between the exogenous and endogenous components of the ability production function. Importantly, we impose no restrictions on the form this connection takes — the marginal value of parental resources may increase, be constant, decrease, or exhibit complex nonlinearities as innate ability increases.[8]

Second, we translate the expected ability in (16) into an ability distribution for the population of children of parents of type j with disposable income $c_t^{j\prime}$ by assuming that ability is distributed lognormally with variance σ^2

$$(17) \quad \ln w_{t+1} \sim N\left(E\left[\ln w_{t+1} \mid w_t^j, c_t^{j\prime}\right], \sigma^2\right).$$

The ability distribution over the income range relevant to this paper is commonly calibrated as lognormal (Tuomala, 1990). The variance σ^2 represents an exogenous, stochastic shock to child ability common across parent types.

The simulations of this model will use a discrete distribution of abilities, consistent with the model described in Section II, whereas (16) and (17) appear to produce continuous ability distributions. To classify individuals into I discrete types, we define fixed ranges of w that correspond to each type $i \in I$. By "fixed," we mean that the boundary values of w that determine whether an individual is assigned wage w^i or w^{i+1} are exogenously given. With these fixed ranges, we can translate the distribution of ability for a given child implied by (17) into transition probabilities among types across generations.[9]

Finally, before proceeding with the calibration, we specify the tax system facing individuals, the utility function those individuals maximize, and the set of ability

[8] We do not estimate this production function directly using our empirical approach because our empirical approach relies on a fixed effects specification, which would difference out parent ability. Our regression specification estimates a coefficient on parental income that is comparable to the coefficient on parental income in (16).

[9] An example may help clarify the procedure. Suppose $I = 2$, so that there are two ability types. Denote the fixed wage level that separates types 1 and 2 as w^*. A mother of type j expects her child, on average, to have the ability $E\left[\ln w_2 \mid w_1^j, c_1^j\right]$ as defined by (16) . In reality, her child's ability is a random variable distributed according to $N\left(E\left[\ln w_2 \mid w_1^j, c_1^j\right], \sigma^2\right)$. The probability that her child's ability ends up in the lower half of the full distribution of wages across all children is, therefore, the value at w^* of the cumulative density function implied by this normal distribution.

types.[10] For the status quo tax system, we assume that the Kotlikoff and Rapson (2007) calculations of marginal effective tax rates on income for 30-year-old couples in the United States in 2005 are a good approximation of the status quo tax policy facing parents of young children.[11] These authors' detailed calculations go well beyond statutory personal income tax schedules and include a wide array of transfer programs, as well as corporate, payroll, and state and local income and sales taxes. Our computations require a smooth tax function, so we take a fifth-order polynomial approximation of the Kotlikoff-Rapson schedule up to $75,000, well above the EITC phase-out. The government's tax system also includes a grant to all individuals, which we determined by enforcing the government's budget constraint in each generation. The individual's current-generation utility takes a separable, isoelastic form,

$$u\left(c_t^i\right)-v\left(\frac{y_t^i}{w_t^i}\right)=\frac{\left(c_t^i\right)^{1-\gamma}-1}{1-\gamma}-\frac{\theta}{\sigma}\left(\frac{y_t^i}{w_t^i}\right)^{\sigma},$$

and we set $\gamma = 2$ and $\sigma = 3$ to be consistent with mainstream estimates of these parameters (Barsky et al., 1997; Chetty, 2006; Chetty, 2012). We choose $\theta = 2.5$ so that hours worked in the simulation approximately match the average labor supply in the population. We initially set $\beta = 1$, reflecting no discounting of utility across generations, the assumption preferred by Ramsey (1928). Guided by the empirical analysis discussed above, we assume ability comes in $I = 4$ fixed types (roughly interpretable as the hourly wage): $w_t^i \in \{3.44, 6.30, 9.42, 19.57\}$ for all $t = \{1,2,...,T\}$. The probability distribution across those types is uniform in the first generation but is endogenously determined in the model for subsequent generations.

C. Calibration Results

Expressions (16) and (17) indicate that the model calibration will search over values of the following seven parameters: $\{\rho, \alpha_a, \{\alpha_c^j\}_{j=1}^I, \sigma\}$. We impose values for two parameters based on prior research. We set $\rho = 0.5$ for the parameter controlling the transmission of ability across generations. This assumption is based on the voluminous evidence surveyed in Feldman, Otto, and Christiansen (2000), which has produced a range of plausible estimates for ρ — generally in the range of 0.3 to 0.7. This evidence enables us to impose a value of ρ (rather than estimate it), thereby increasing the model's ability to estimate the parameters on which we have more limited evidence.[12] We also impose the value of $\sigma = 0.76$, which we calculate using data on wages from the NLSY sample. This leaves five parameters to be chosen.

---

10 We demonstrate the robustness of our results to alternative specifications in the appendix.

11 The biggest subsequent changes in tax policy not captured in these calculations are the increases in marginal tax rates among the highest income earners that have occurred through the Patient Protection and Affordable Care Act (ACA) of 2010 and the American Taxpayer Relief Act of 2012, as well as the phase-outs of subsidies for health insurance for low-to-middle income earners under the ACA.

12 In the appendix, we show that our qualitative conclusions are robust to ρ ranging from 0.3 to 0.7.

To calibrate the five remaining parameters, we minimize a weighted sum of squared errors between our model's output and the three sets of target statistics: the marginal effects, the transition matrix, and the mean log wage. For reference, we calculate the marginal effects of parental disposable income from the model output as the increase in the probability of a given child type caused by an increase of 1 percent in a given parent type's disposable income, while the model output directly implies transition probabilities for all parent and child types and mean log wages. We weight the squared errors by the inverse of the targets' standard errors, which has the effect of putting much greater weight on the more-precisely-estimated transition matrix elements and the mean log wage. We use ten generations ($T = 10$) in the simulations, using the middle (fifth-to-sixth) generation as the target for the calibration exercise.[13]

Table 1 shows the parameter values chosen by the simulation.

Recall that α_a and α_c^j are the weights on the two channels, ability and economic resources, through which parents affect their child's ability. The product of ρ and α_a gives the weight on parental ability in expected child ability, while α_c^j gives the (parental type-specific) weight on parental resources. The monotonically declining values of α_c^j in Table 1 suggest that parental resources play a greater role among lower-ability parents, consistent with the empirical evidence. Key moments determining the estimates of the α_c^j are the coefficients on parent income in determining child ability. Key moments determining both the estimates of the α_c^j and the estimate of α_a are the elements of the transition matrix of parent ability to child ability, as these determine the combined role that parent ability and parent resources play in determining child ability.

The simulation performs well in matching the empirical targets for which the data are most informative, the transition matrix and mean log wage. The simulation yields marginal effects of parental resources that differ substantially from the data.[14] The calibrated status quo marginal effects exhibit a pattern much closer to what the empirical literature described previously would suggest — negative for lower child types and positive for higher child types — than do the estimated effects in the data. This divergence is not surprising, however, given the statistical insignificance of the empirical estimates and their often-unexpected signs.

Table 2 shows that the simulation closely matches the data for the parent-child type transition matrix between generations (we show the transition between the fifth and sixth generations of the simulation as an illustrative example).

Finally, the simulation matches the mean log wage, 2.07.

Table 1

Parameter Values Chosen in Calibration

	α_a	α_c^1	α_c^2	α_c^3	α_c^4
Value under status quo policy	0.88	0.57	0.23	0.17	0.00

[13] We show robustness to this choice in the appendix.

[14] Appendix Table 7 shows this for the fifth (middle) generation of the simulation.

Table 2
Parent-Child Type Transition Matrix

	Data					Calibrated Status Quo Policy			
	Child Type					Child Type			
Parent Type	1	2	3	4	Parent Type	1	2	3	4
1	0.31	0.24	0.22	0.22	1	0.30	0.26	0.26	0.19
2	0.29	0.25	0.23	0.23	2	0.28	0.25	0.26	0.20
3	0.20	0.27	0.29	0.25	3	0.22	0.24	0.28	0.26
4	0.20	0.23	0.26	0.31	4	0.19	0.22	0.29	0.30

V. EFFECTS OF OPTIMAL POLICY ON ABILITY DISTRIBUTION AND WELFARE

In this section, we simulate a many-period version of the tax authority's problem using the parameter values and the model from the previous sections. We characterize utilitarian welfare-maximizing policy by comparing it to the status quo policy used in that calibration, and we calculate the welfare gains from the optimal policy's effects on the evolution of the ability distribution.

A. Utilitarian Optimal Policy Compared to the Status Quo

Table 3 shows average and marginal tax rates for each type under the utilitarian welfare-maximizing and status quo policies. Average tax rates are calculated as the ratio $(y - c)/y$. For marginal tax rates, we compare the marginal tax rates imposed by the status quo policy to the marginal tax rates that would implement the utilitarian welfare-maximizing allocation.

The utilitarian welfare-maximizing policy is substantially more redistributive than the status quo, generating large transfers to low-skilled parents. This result is due entirely to the redistributive preferences in the assumed social welfare function. Nevertheless, these redistributive transfers generate an improved ability distribution by taking advantage of

Table 3
Marginal and Average Tax Rates

	Marginal Tax Rate (Percent)		Average Tax Rate (Percent)	
Type	W-max	Status Quo	W-max	Status Quo
1 (lowest)	32	21	–192	–42
2	44	31	–60	–14
3	46	35	–13	0
4 (highest)	0	32	37	16

the gap between the impact of increased disposable income on the children of low-ability parents and high-ability parents. The larger marginal distortions (other than on $I = 4$) under the utilitarian welfare-maximizing policy make the allocations enjoyed by lower types less attractive to those with higher ability, as is common in optimal tax analyses.[15]

One might be concerned that these results could be driven in important ways by having only four types. For example, in a Mirrlees framework, having a finite number of types yields a zero marginal tax rate at the top of the distribution, as shown above. However, the differences in welfare between the status quo and the welfare-maximizing policy stems largely from the much lower average tax rates at the bottom of the distribution under the welfare-maximizing policy, rather than from differences in marginal distortions at the top.[16]

The utilitarian welfare-maximizing policy adjusts intertemporal allocations to capitalize on the endogeneity of ability, as was suggested in the discussion of Proposition 1. Table 4 reports the difference between the tax authority's "budget balance" as a share of aggregate income in each generation under the welfare-maximizing policy and under the status quo. In other words, it is the additional average tax rate assessed on each generation relative to a balanced budget (assumed for the status quo).

The utilitarian welfare-maximizing policy borrows from future generations to fund greater investment in the skills of the current generation relative to the status quo. Of course, our model abstracts from many features of the economy, notably capital as a factor of production, some of which may make deficit-financed investment in children less appealing. However, the key point illustrated by Table 4 is that society can benefit by having later generations contribute, through higher taxes, to improving the ability distribution generated by earlier generations.

Table 4

Intertemporal Allocations

	Difference in Government Budget Balance (as Percent of Output)									
	1	2	3	4	5	6	7	8	9	10
W-max – Status Quo	–5.1	–0.7	–0.9	–0.9	–0.9	–0.9	–0.9	–0.8	–1.0	11.3

[15] If, as in a standard Mirrlees-type analysis, the planner believed the starting ability distribution was fixed, the optimal policy would generate average tax rates of [–253%, –73%, –28%, –2%, and 42%]. Therefore, the presence of endogenous ability reduces the extent of redistribution that is optimal in a standard utilitarian model. The intuition for this result is that low earners in the conventionally optimal policy face high marginal tax rates, with much of their welfare coming through increased leisure time. With endogenous ability, it can be optimal to have these workers exert more effort, earn more income, and, therefore, enjoy more disposable income than in a static model. Note that the status quo policy includes much less redistribution, so that the increased redistribution under the utilitarian-optimal policy generates large welfare gains through its effects on ability alone.

[16] Consistent with this logic, we show in the appendix that our results are similar with other numbers of types (i.e., five or 10 types).

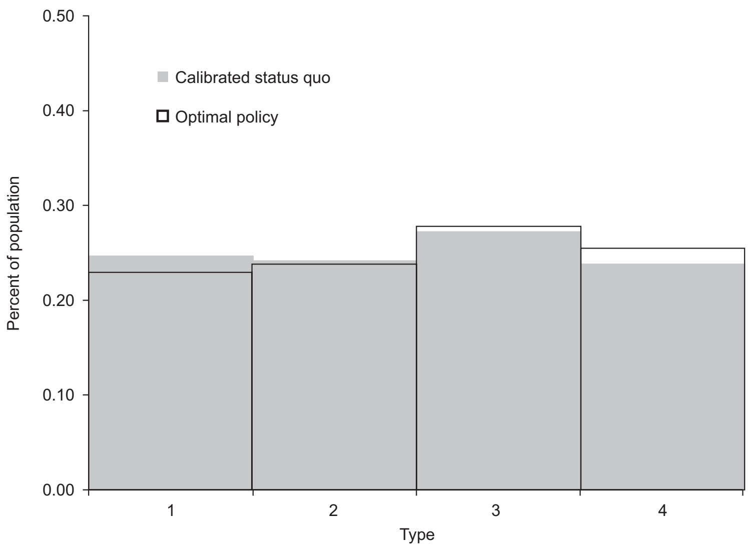
These differences in tax policy affect the evolution of the ability distribution. The ability transition matrices across generations are shown in Table 5.

The utilitarian welfare-maximizing policy enables a substantially greater share of the children of lower-skilled parents to move up the skill ladder than does the status quo policy, while it has much smaller — though non-negative — effects on the prospects of the children of higher-skilled parents. Figure 1 shows the result: the welfare-maximizing

Table 5
Transition Matrix

	W-max					Status Quo			
	Child Type					Child Type			
Parent Type	1	2	3	4	Parent Type	1	2	3	4
1	0.24	0.24	0.28	0.24	1	0.30	0.26	0.26	0.19
2	0.27	0.25	0.27	0.21	2	0.28	0.25	0.26	0.20
3	0.22	0.24	0.28	0.26	3	0.22	0.24	0.28	0.26
4	0.19	0.22	0.29	0.30	4	0.19	0.22	0.29	0.30

Figure 1
Ability Distribution under Two Policies



policy leads to 1.8 percent fewer individuals of the lowest type and 1.6 percent more of the highest type than does the status quo policy.

B. Welfare Gains from Improved Ability Distribution

Finally, we turn to welfare impacts. While utilitarian welfare is much higher under the welfare-maximizing policy than under the status quo, that difference is driven by the model's assumed social preference for income equality. For the purposes of this paper, we are more interested in the welfare gains due to endogenous ability by itself, so we consider the following thought experiment. We define a setting called the "adjusted status quo" in which the distribution of abilities generated by the utilitarian welfare-maximizing policy prevails but the within-period utility levels of individuals of each ability type are those from the status quo policy outcome. We then calculate the factor by which disposable income would have to rise in the status quo to reach the welfare of the adjusted status quo, as our measure of the welfare gain is solely due to the utilitarian welfare-maximizing policy's effects on the ability distribution over time. Similar factors can be calculated for each type of first-generation parent, indicating how the welfare gains through this channel are shared. Table 6 shows the results for the baseline case of ten generations.

As these results show, the utilitarian welfare-maximizing policy has the potential to generate a welfare gain equivalent to 1.75 percent of aggregate disposable income simply by shifting the ability distribution over time. The gains are somewhat larger among low-skilled parents, as would be expected. Nevertheless, high-skilled parents gain substantially, as the gains accruing to future generations raise the current generation's present-value welfare. Gains for future generations follow the same patterns.[17]

Table 6

Welfare

	Welfare Level (Utils)		Welfare Gain
	Status Quo Policy	Adjusted Status Quo	(Percent of Disposable Income)
Overall	6.44	6.49	1.75
Type-specific			
1 (lowest)	6.25	6.30	2.2
2	6.41	6.46	1.6
3	6.49	6.54	1.6
4 (highest)	6.61	6.66	1.6

[17] In the appendix, we explore the robustness of these baseline results to variation in time discounting β, the number of generations T, the assumed persistence of type across generations ρ, the number of types I, and other details of our specification. The qualitative and quantitative lessons of the baseline analysis prove to be robust, consistent with our message that a utilitarian welfare-maximizing policy that redistributes more than the status quo and shifts resources to earlier generations raises average ability over time and yields a sizeable welfare gain.

C. Effects of Parental Time Allocation

In this brief subsection, we incorporate parental time allocation into the quantitative analysis by estimating the welfare gain from the version of the model in which parental time working can affect child ability.[18] In particular, we target these marginal effects, as well as the transition matrix and mean log wage as in Section IV, calibrating the following reduced-form equation, which is analogous to (16),

$$(18) \qquad E\left[\ln w_2 \mid w_1^j, c_1^{j'}\right] = \alpha_a\left(\rho \ln w_1^j + (1-\rho)\ln \tilde{w}\right) + \alpha_c^j \ln c_1^{j'} + \alpha_l^j l_1^{j'},$$

where $l_1^{j\prime}$ is the labor effort of the parent of type j'. The calibration therefore searches for values of the parameters $\left\{\rho, \alpha_a, \left\{\alpha_c^j\right\}_{j=1}^{I}, \left\{\alpha_l^j\right\}_{j=1}^{I}, \sigma\right\}$ to match the empirical targets; with $I = 4$ there are 11 values to estimate. As before, we assume baseline values for $\rho = 0.50$ and $\sigma = 0.76$.

The results of the calibration exercise indicate that the values for α_a and α_c^j (the weights on the parental ability and parental economic resources) are only slightly changed by the inclusion of parental effort. Most of the α_l parameters are near zero, though the substantially larger negative value for the highest ability type suggests that creating greater incentives for effort among the higher earners in this group may have some costs in terms of child ability; see the online appendix for further details.

As in the baseline case, the simulation does well in matching the empirical targets for which the data are most informative, the transition matrix and mean log wage. The calibrated marginal effects of parental resources and effort imply that child ability is increasing in parental resources and leisure (i.e., decreasing in parental labor effort), patterns not significantly apparent in the data. For brevity, we omit these results, which are very similar to the relevant analogs shown previously.

The optimal policy given the calibrated model also strongly resembles that in the baseline case (where parental effort did not affect children's abilities). In particular, the optimal policy is substantially more redistributive across incomes than the status quo policy and borrows from future generations. The ability distribution improves over time under the optimal policy, so that 1.3 percent more of the population is of the highest type and 1.5 percent less is of the lowest type in the fifth (middle) generation than under the status quo policy. The welfare gain from this improvement in the ability distribution is 1.51 percent (compared to 1.75 percent in the baseline case). The slightly smaller values for the shifts in the ability distribution and welfare gain are due to the negative effects of increased parental work effort on child ability (the α_l coefficients), which offset in part the gains from having greater output and thus parental disposable income.

VI. CONCLUSION

This paper starts with the recent empirical finding that the cognitive performance of children in poor families improves when their parents have more disposable income, and asks how tax policy might respond to this relationship.

[18] In appendix Table 5, we estimate the marginal effects of both parental disposable income and parental work effort on child ability, which we use as additional targets in a calibration exercise similar to that described in the main text of the paper.

We take an initial step toward answering this question, providing both qualitative and quantitative results on tax policy in this context. First, we characterize conditions describing welfare-maximizing tax policy both within and across generations when children's abilities are a potentially complex function of inherited characteristics and parental (financial) resources. We use these conditions to better understand the ways in which this dependence changes the proper design of tax policy. Second, we use the model to quantify the potential benefits of well-designed policy. We specify the model's parameters to match new estimates of the intergenerational transmission of skills that we obtain by analyzing panel data from the NLSY. We then simulate the effects of the utilitarian welfare-maximizing policy in this model and show that its substantially greater progressivity shifts the ability distribution up over time. This shift generates an aggregate welfare gain equivalent to 1.75 percent of total disposable income in perpetuity with larger gains for the poor. Even higher-skilled members of the current generation benefit substantially, however, as the welfare gains experienced by future generations increase the present value of the welfare of the current generation.

Our paper abstracts from an important issue: some individuals in the population do not have children. Thus, we must be careful in interpreting our results; for example, the welfare gains from the policy change analyzed apply to *families with children*, not to the full population. Aside from this interpretive point, however, three factors suggest that omitting individuals without children will have at most minor effects on our results. First, Child Trends (2002) reports that 84 percent of men and 86 percent of women age 45 and older have ever had a biological child. In a benchmark lifecycle model, the lifetime tax burden is relevant for expenditure decisions, so our analysis of mothers with children captures the effects of our policy for the great majority of the population. Second, our calculations of the welfare gains from tax reform are restricted to its effects on the ability transition matrix. Those effects are relevant only for families with children, by definition. Finally, to the extent that tax policy can be (and often is, as with the Child Tax Credit, the EITC, or Temporary Assistance for Needy Families) made to depend on the presence of children, implementation of the policy change we consider may apply only to adults with children.

Nonetheless, it may be valuable to consider how including adults without children would affect our analysis. If the policy changes were not restricted to households with children, the additional revenue would be raised from and distributed to individuals with and without children. Because childless households have no effect on the ability transition matrix, the benefits and costs of redistribution — in terms of its effects on the ability transition matrix — would be reduced, and the welfare gains from reform may change.[19] In future work, it could be possible to model taxing or spending differentially on those with and without children, though in principle this could create distortions to the choice of whether to have children.

[19] This discussion also raises important issues about policy's design toward the number of children in a household, which we leave to future work.

We take a "welfarist" approach to defining optimal policy by using the standard utilitarian social welfare function, which equals the sum of utilities across the population. Philosophers and social choice theorists have suggested alternative goals for policy, including one of particular relevance to this paper, namely, equality of opportunity.[20] Roemer's (2000) leading treatment of that principle emphasizes that it requires a distinction between "before" and "after," in that opportunities should be equalized "before" competition among individuals begins but outcomes "after" competition begins should not. A classic tension in following this principle arises when parents' resources affect children's disutility of effort, so that what is "before" for children is "after" for their parents (Fishkin, 1987). A related point is that our use of a utilitarian social welfare function means that we do not fully insure children against "before" factors (those over which they have no control, such as their parents' financial resources) because the benefits from doing so must be weighed against the costs of dampening parents' incentives to earn more income. Future research could explore such issues.

Of course, future research may be able to improve our understanding of the tax policy studied in this paper. For example, when a panel dataset of sufficient duration allows us to link data on parents' and children's wages, this will allow estimates of the intergenerational effect of parental income on parent-child wage transitions. Exploring the potential for interactions between optimal tax policy and optimal government policy in other domains, including education, also seems to be a promising area for future research. Incorporating other dimensions of parental influence is another natural next step. We have shown that parental leisure versus work time does not seem to exert an important influence in this case, but one might study how the composition of parents' available resources (i.e., as disposable income or in-kind, such as education) affects the results. Such analyses may have implications for a broader class of policies that, like the taxes in this paper, could be used to affect — rather than merely respond to — the dynamics of the ability distribution.

ACKNOWLEDGMENTS AND DISCLAIMERS

We thank participants at the Cowles Summer Conference on Advances in Dynamic Taxation, the National Tax Association annual meeting, and seminars at IFS/LSE, Harvard University, Hebrew University, the University of Pennsylvania, and Stanford University. We thank Jillian Popadak, Jerry Yeh, and Roger Ou for research assistance and Benjamin B. Lockwood for exceptional work on the computational simulations. We thank the Entrepreneurship and Family Business Research Centre, Center for Human Resources, Mack Center, Zicklin Center, Risk and Decision Processes Center, and Global Initiatives Center, all at Wharton, for generous research support. We would especially like to thank Doug Bernheim, Caroline Hoxby, Marek Kapicka,

[20] Economists have begun including unconventional objectives in optimal tax research (Weinzierl, 2014a, 2014b; Saez and Stantcheva, forthcoming; Lockwood, Nathanson, and Weyl, 2015).

Robert E. Lucas, Luigi Pistaferri, Monica Singhal, Michele Tertilt, Aleh Tsyvinski, and several anonymous referees for comments and suggestions. We thank Gordon Dahl, Lance Lochner, Dave Rapson, and Larry Kotlikoff for sharing data.

DISCLOSURES

Gelber served as Deputy Assistant Secretary for Economic Policy at the U.S. Department of the Treasury from 2012 to 2013, although he had a minimal role in the development of tax policy. We have no financial arrangements that might give rise to conflicts of interest with respect to the research reported in this paper.

REFERENCES

Akee, Randall K. Q., William E. Copeland, Gordon Keeler, Adrian Angold, and E. Jane Costello, 2010. "Parents' Income and Children's Outcomes: A Quasi-Experiment Using Transfer Payments from Casino Profits." *American Economic Journal: Applied Economics* 2 (1), 86–115.

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

Anderberg, Dan, 2009. "Optimal Policy and the Risk Properties of Human Capital Reconsidered." *Journal of Public Economics* 93 (9–10), 1017–1026.

Barsky, Robert B., Thomas F. Juster, Miles S. Kimball, and Matthew D. Shapiro, 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics* 112 (2), 537–580.

Becker, Gary S., 1981. "Altruism in the Family and Selfishness in the Market Place." *Economica* 48 (189), 1–15.

Becker, Gary S., and Nigel Tomes, 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy* 84 (4, part 2), S143–162.

Becker, Gary S., and Nigel Tomes, 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4 (3), S1–39.

Chetty, Raj, 2006. "A New Method of Estimating Risk Aversion." *American Economic Review* 96 (5), 1821–1834.

Chetty, Raj, 2012. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica* 80 (3), 969–1018.

Child Trends, 2002. *Charting Parenthood: A Statistical Portrait of Fathers and Mothers in America*. Child Trends, Washington, DC.

Dahl, Gordon B., and Lance Lochner, 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *American Economic Review* 102 (5), 1927–1956.

Dynan, Karen E., 1993. "How Prudent Are Consumers?" *Journal of Political Economy* 101 (6), 1104–1113.

Eissa, Nada, and Hilary W. Hoynes, 2006. "Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply." In Poterba, James M. (ed.), *Tax Policy and the Economy, Volume 20*, 73–110. MIT Press, Cambridge, MA.

Farhi, Emmanuel, and Iván Werning, 2010. "Progressive Estate Taxation." *Quarterly Journal of Economics* 125 (2), 635–673.

Farhi, Emmanuel, and Iván Werning, 2013. "Insurance and Taxation over the Life Cycle." *Review of Economic Studies* 80 (2), 596–635.

Feldman, Marcus W., Sarah R. Otto, and Freddy B. Christiansen, 2000. "Genes, Culture, and Inequality." In Arrow, Kenneth, Samuel Bowles, and Steven Durlauf (eds.), *Meritocracy and Economic Inequality*, 61–85. Princeton University Press, Princeton, NJ.

Fishkin, James S., 1987. "Liberty Versus Equal Opportunity." *Social Philosophy and Policy* 5 (1), 32–48.

Flynn, James R., 1987. "Massive IQ Gains in 14 Nations: What IQ Tests Really Measure." *Psychological Bulletin* 101 (2), 171–191.

Gelber, Alexander M., and Joshua W. Mitchell, 2012. "Taxes and Time Allocation: Evidence from Single Women and Men." *Review of Economic Studies* 79 (3), 863–897.

Gelber, Alexander M., and Matthew C. Weinzierl, 2012. "Equalizing Outcomes and Equalizing Opportunities: Optimal Taxation when Children's Abilities Depend on Parents' Resources." NBER Working Paper No. 18332. National Bureau of Economic Research, Cambridge, MA.

Goldin, Claudia, and Lawrence F. Katz, 2007. "Long-Run Changes in the U.S. Wage Structure: Narrowing, Widening, Polarizing." *Brookings Papers on Economic Activity* 38 (2), 135–167.

Golosov, Mikhail, Aleh Tsyvinski, and Ivan Werning, 2007. "New Dynamic Public Finance: A User's Guide." In Acemoglu, Daron, Kenneth Rogoff, and Michael Woodford (eds.), *NBER Macroeconomics Annual 2006, Volume 21*, 317–388. MIT Press, Cambridge, MA.

Golosov, Mikhail, and Aleh Tsyvinski, 2006. "Designing Optimal Disability Insurance: A Case for Asset Testing." *Journal of Political Economy* 114 (2), 257–279.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski, 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies* 70 (3), 569–587.

Grochulski, Borys, and Tomasz Piskorski, 2010. "Risky Human Capital and Deferred Capital Income Taxation." *Journal of Economic Theory* 145 (3), 908–943.

Gruber, Jonathan, 2006. "A Tax-Based Estimate of the Elasticity of Intertemporal Substitution." NBER Working Paper No. 11945. National Bureau of Economic Research, Cambridge, MA.

Haider, Steven, and Gary Solon, 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96 (4), 1308–1320.

Hall, Robert E., 1988. "Intertemporal Substitution in Consumption." *Journal of Political Economy* 96 (2), 339–357.

Hotz, V. Joseph, and John Karl Scholz, 2003. "The Earned Income Tax Credit." In Moffitt, Robert A. (ed.), *Means-Tested Transfer Programs in the United States*, 141–197. University of Chicago Press, Chicago, IL.

Jacob, Brian, Max Kapustin, and Jens Ludwig, 2014. "Human Capital Effects of Anti-Poverty Programs: Evidence from a Randomized Housing Voucher Lottery." NBER Working Paper No. 20164. National Bureau of Economic Research, Cambridge, MA.

Kapicka, Marek, 2006a. "The Dynamics of Optimal Taxation when Human Capital is Endogenous." UCSB Working Paper. Department of Economics, University of California, Santa Barbara, Santa Barbara, CA.

Kapicka, Marek, 2006b. "Optimal Income Taxation with Human Capital Accumulation and Limited Record Keeping." *Review of Economic Dynamics* 9 (4), 612–639.

Kocherlakota, Narayana R., 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation." *Econometrica* 73 (5), 1587–1621.

Kopczuk, Wojciech, 2013. "Incentive Effects of Inheritances and Optimal Estate Taxation." *American Economic Review* 103 (3), 472–477.

Kotlikoff, Laurence, and David Rapson, 2007. "Does It Pay, at the Margin, to Work and Save? Measuring Effective Marginal Taxes on Americans' Labor Supply and Savings." In Poterba, James M. (ed.), *Tax Policy and the Economy, Volume 21,* 83–144. MIT Press, Cambridge, MA.

Leibowitz, Arleen, 1974. "Home Investments in Children." *Journal of Political Economy* 82 (2, part 2), S111–131.

Lockwood, Benjamin, Charles Nathanson, and E. Glen Weyl, 2015. "Taxation and the Allocation of Talent." Manuscript. Department of Economics, Harvard University, Cambridge, MA.

Løken, Katrine V., Magne Mogstad, and Matthew Wiswall, 2012. "What Linear Estimators Miss: The Effects of Family Income on Child Outcomes." *American Economic Journal: Applied Economics* 4 (2), 1–35.

Macours, Karen, Norbert Schady, and Renos Vakis, 2012. "Cash Transfers, Behavioral Changes, and Cognitive Development in Early Childhood: Evidence from a Randomized Experiment." *American Economic Journal: Applied Economics* 4 (2), 247–273.

Mankiw, N. Gregory, and Matthew Weinzierl, 2010. "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution." *American Economic Journal: Economic Policy* 2 (1), 155–176.

Milligan, Kevin, and Mark Stabile, 2011. "Do Child Tax Benefits Affect the Well-Being of Children? Evidence from Canadian Child Benefit Expansions." *American Economic Journal: Economic Policy* 3 (3), 175–205.

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

Paxson, Christina, and Norbert Schady, 2007. "Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health and Parenting." *Journal of Human Resources* 42 (1), 49–84.

Piketty, Thomas, and Emmanuel Saez, 2013. "A Theory of Optimal Inheritance Taxation." *Econometrica* 81 (5), 1851–1886.

Ramsey, Frank P., 1928. "A Mathematical Theory of Saving." *Economic Journal* 38 (152), 543–559.

Roemer, John E., 2000. *Equality of Opportunity*. Harvard University Press, Cambridge, MA.

Saez, Emmanuel, and Stefanie Stantcheva, forthcoming. "Generalized Social Marginal Welfare Weights for Optimal Tax Theory." *American Economic Review*.

Stantcheva, Stefanie, 2015. "Optimal Taxation and Human Capital Policies over the Life Cycle." NBER Working Paper No. 21207, National Bureau of Economic Research, Cambridge, MA.

Tuomala, Matti, 1990. *Optimal Income Tax and Redistribution*. Clarendon Press, Oxford, UK.

Weinzierl, Matthew, 2011. "The Surprising Power of Age-Dependent Taxes." *Review of Economic Studies* 78 (4), 1490–1518.

Weinzierl, Matthew, 2014a. "The Promise of Positive Optimal Taxation: Normative Diversity and a Role for Equal Sacrifice." *Journal of Public Economics* 118 (Oct.), 128–142.

Weinzierl, Matthew, 2014b. "Revisiting the Classical View of Benefit-Based Taxation." Harvard Business School Working Paper No. 14-401. Harvard Business School, Boston, MA.