Estate Taxation with Altruism Heterogeneity

American Economic Review: Papers & Proceedings 2013, 103(3): 489–495
http://dx.doi.org/10.1257/aer.103.3.489

OPTIMAL TAXATION AND SOCIAL INSURANCE‡

Estate Taxation with Altruism Heterogeneity†

By EMMANUEL FARHI AND IVÁN WERNING*

We develop a theory of optimal estate taxation in a model where bequest inequality is driven by differences in parental altruism. We show that a wide range of results are possible, from positive taxes to subsidies, depending on redistributive objectives implicit in the cardinal specification of utility and social welfare functions. We propose a normalization that is helpful in classifying these different possibilities. We isolate cases where the optimal policy bans negative bequests and taxes positive bequests, features present in most advanced countries.

I. Introduction

Many people's ideas about estate taxes take the perspective of children and build on the intuition that inheritances are pure luck—after all, children do nothing to deserve their parents—to conclude that bequests should be redistributed away to help level the playing field.

However, taking the perspective of parents, one can make a powerful argument against estate taxation on the grounds of fairness. This case is eloquently articulated in the form of a parable by Mankiw (2006):

> *Consider the story of twin brothers Spendthrift Sam and Frugal Frank. Each starts a dot-com after college and sells the business a few years later, accumulating a $10 million nest egg. Sam then lives the high life, enjoying expensive vacations and throwing lavish parties. Frank, meanwhile, lives more modestly. He keeps his fortune invested in the economy, where it finances capital accumulation, new technologies, and economic growth. He wants to leave most of his money to his children, grandchildren, nephews, and nieces.*
>
> *Now ask yourself: Which millionaire should pay higher taxes?…What principle of social justice says that Frank should be penalized for his frugality? None that I know of.*

In this paper, we offer a theory of estate taxation that reconciles these two philosophies. We analyze a model where parents with different degrees of altruism consume and leave bequests to their offspring. Altruism is private information, giving rise to a trade-off between equality of opportunity for newborns and incentives for altruistic parents. We consider a wide class of social welfare functions and characterize both optimal nonlinear and linear estate tax systems.

In Farhi and Werning (2010) we formulated a similar optimal tax problem by taking a canonical Mirrleesian tax model—where skill differences are the only source of heterogeneity—and adding a bequest decision. In the model of that paper, more productive parents earn more, consume more, and bequeath more.

Instead, in this paper we depart from the canonical optimal tax model, abstracting from parental earnings inequality to focus instead on differences in the degree of altruism.[1] Our main

‡ *Discussant:* Ali Shourideh, Wharton.

* Farhi: Harvard University, Littauer Center, 1805 Cambridge Street, Cambridge, MA 02138 (e-mail: efarhi@harvard.edu); Werning: MIT Economics Department, Building E52, 50 Memorial Drive, Cambridge, MA 02142 (e-mail: iwerning@mit.edu). We thank Ali Shourideh and Stefanie Stantcheva for useful comments and suggestions. Lucas Manuelli provided valuable research assistance.

† To view additional materials, and author disclosure statement(s),visit the article page at http://dx.doi.org/10.1257/aer.103.3.489.

[1] Piketty and Saez (2012) present a model with both altruism and productivity differences and study optimal linear taxes on savings/bequests. Our results suggest that altruism heterogeneity coupled with their choice of welfare functions may be key to understanding the simulations with positive and negative marginal tax rates that they report.

goal is to isolate what this different source for bequest inequality implies for estate taxation.

We find that optimal estate taxes depend crucially on redistributive objectives. Different welfare criteria lead to results ranging from taxes to subsidies. We identify a few useful benchmarks. First, optimal estate taxes are zero when no direct weight is placed on children and when parents' welfare is summarized by a Utilitarian criterion using a normalization of utility (Proposition 1). This formalizes Mankiw's intuition. Second, when the Utilitarian criterion is augmented with a positive weight on children's welfare, subsidies on estates emerge (Proposition 2). Finally, a clear cut case for positive taxes on estates is possible when one adopts a more extreme preference for equality of opportunity of children. With a Rawlsian maximin criterion optimal policy taxes positive bequests and bans negative ones (Proposition 4). These two properties are consistent with most actual tax codes, providing one possible justification for their use. We provide both results for nonlinear taxes (Propositions 1–4) and results for linear taxes (Propositions 5–9).

II. The Model

There are two generations, parents born at $t = 0$ and children born at $t = 1$; each living for one period. Parents are altruistic and each has exactly one offspring. There is a storage technology between periods with constant return R. Parents are heterogenous. A parent of type θ has strictly quasi-concave preferences represented by the utility function

$$U^p(c_0, c_1; \theta),$$

where U^p is increasing, strictly concave and twice differentiable in $(c_0, c_1; \theta)$.[2] The type θ is distributed in the population according to a continuous density $f(\theta)$ on the interval $[\underline{\theta}, \bar{\theta}]$. We make the following standard single-crossing condition assumption.

[2] With a few additional assumptions, any strictly quasi-concave function can be monotonically transformed into a strictly concave utility function; see, e.g., Connell and Rasmusen (2012).

ASSUMPTION 1: *The parent's utility function U^p satisfies*

$$\frac{\partial}{\partial \theta}\left(\frac{U^p_{c_1}(c_0, c_1; \theta)}{U^p_{c_0}(c_0, c_1; \theta)}\right) > 0.$$

Higher types are more altruistic; lower types more selfish. Single crossing is an assumption about ordinal preferences, not cardinal utility. It will be useful to make a normalization regarding cardinal utility. Define the indirect utility function

$$V^p(I, R; \theta) \equiv \max_{c_0, c_1} U^p(c_0, c_1; \theta)$$

$$\text{such that} \quad c_0 + \frac{1}{R}c_1 = I.$$

ASSUMPTION 2: *The parent's utility function U^p is such that marginal utility is constant without redistribution*

$$V^p_I(I, R; \theta) = V^p_I(I, R; \theta') \quad \textit{for all } \theta, \theta', \textit{ and } I.$$

Assumption 2 amounts to a renormalization of cardinal utility that does not change ordinal preferences (see the online Appendix for details). Nevertheless, it will prove useful to categorize different cases and results.

We maintain Assumptions 1 and 2 throughout the paper. For a few results we need the following additional assumption.

ASSUMPTION 3: *The parent's utility function U^p satisfies*

$$U^p_{c_0,\theta}(c_0, c_1; \theta) \le 0 \quad \textit{and} \quad U^p_{c_1,\theta}(c_0, c_1; \theta) \ge 0.$$

Assumption 3 implies the single crossing condition in Assumption 1. A simple example satisfying all three assumptions is $U^p(c_0, c_1; \theta) = (1 - \theta)\log(c_0) + \theta\log(c_1)$.

We will employ a weighted Utilitarian criterion

$$\int \left(\lambda_\theta U^p(c_0(\theta), c_1(\theta); \theta) + \alpha_\theta U^c(c_1(\theta))\right) f(\theta)\, d\theta,$$

where λ_θ is the weight on a parent of type θ, α_θ is the weight on a child with parent of type θ, and U^c is increasing, concave, and differentiable. There are two interpretations of these weights. First, by varying the weights across types and generations one traces out the Pareto frontier.

Under this interpretation we adopt the ordinal preferences of parents and children and simply place flexible weights on different members of society; cardinal utility is irrelevant. A second interpretation, especially for λ_θ, is possible if we imagine evaluating expected utility behind the veil of uncertainty, before θ is realized. In this case, we interpret cardinal utility for parents to be $\lambda_\theta U^p(c_0, c_1; \theta)$. Observed consumption-savings behavior only identifies ordinal, not cardinal, utility.[3] Thus, flexible weights $\lambda_\theta, \alpha_\theta$ are required to consider a wide range of different tastes for redistribution or specifications of cardinal utility.

With α_θ constant, the curvature of U^c captures a preference for equality of children's consumption. We also want to consider a welfare function with extreme egalitarian preferences for children. To this end, we combine a weighted utilitarian criterion for parents' welfare, $\int \lambda_\theta U^p(c_0(\theta), c_1(\theta); \theta) f(\theta)\, d\theta$, with a Rawlsian maximin criterion for children's welfare,

$$\min_\theta U^c(c_1(\theta)).$$

This delivers the same implications as the weighted-Utilitarian criterion for some appropriate endogenous weights α_θ.

We assume each parent's θ type is private information. This makes the first-best unavailable and creates a trade-off between redistribution and incentives. We follow both a Mirrleesian approach, with no exogenous restrictions on policy instruments beyond those implied by private information, and a Ramsey approach with restricted taxes.

III. Nonlinear Taxation

We begin with the Mirrleesian approach, without arbitrary restrictions on tax instruments, by studying the mechanism design problem that incorporates the incentive constraints. Similar to Mirrlees (1971), the optimum can be implemented with a nonlinear tax of bequests. Parents are subject to the budget constraints

$$c_0 + B + T(B) = I_0 \tag{1}$$

$$c_1 = I_1 + RB, \tag{2}$$

where T is a nonlinear tax on bequests. At points where T is differentiable, the marginal tax rate on bequests equals the implicit marginal tax rate on estates $T'\left(\frac{c_1(\theta) - I_1}{R}\right) = \tau(\theta)$, defined by $(1 + \tau(\theta)) U^p_{c0}(c_0(\theta), c_1(\theta); \theta) \equiv R U^p_{c1}(c_0(\theta), c_1(\theta); \theta)$. Next we characterize the optimal allocation and the associated implicit marginal tax rate.

A. *A Weighted Utilitarian Objective*

The dual planning problem we study is

$$\min_{c_0, c_1, v} \int \left(c_0(\theta) + \frac{1}{R} c_1(\theta) \right) f(\theta)\, d\theta, \tag{3}$$

subject to $c_1(\theta)$ monotone increasing and

$$v(\theta) = U^p(c_0(\theta), c_1(\theta); \theta), \tag{4}$$

$$\dot{v}(\theta) = U^p_\theta(c_0(\theta), c_1(\theta); \theta), \tag{5}$$

$$\int (\lambda_\theta\, U^p(c_0(\theta), c_1(\theta); \theta) + \alpha_\theta U^c(c_1(\theta))) f(\theta)\, d\theta \geq V. \tag{6}$$

This problem minimizes the resources required to achieve a certain level of welfare subject to incentive compatibility. The second constraint is the envelope condition which, together with the monotonicity condition, is necessary and sufficient for incentive compatibility (see, e.g., Milgrom and Segal 2002).

Our first results focus on cases with no weight on children's welfare.

PROPOSITION 1: *Suppose that Assumptions 1 and 2 hold, and that there is no weight on children* $\alpha_\theta = 0$. *Then* (*i*) *if* λ_θ *is constant the optimum coincides with the first-best, and estate taxes are* 0, $\tau(\theta) = 0$; (*ii*) *if* λ_θ *is decreasing and in addition Assumption 3 holds, then marginal estate taxes are positive* $\tau(\theta) \geq 0$; *and* (*iii*) *if* λ_θ *is increasing and in addition Assumption 3 holds, then marginal estate taxes are negative*, $\tau(\theta) \leq 0$.

When the weight on parents, λ_θ, is constant, the first-best allocation is incentive compatible and, hence, optimal. This sets up an important

[3] See Lockwood and Weinzierl (2012) for an application of this principle to the taxation of labor.

benchmark where bequests are not taxed. It formalizes the parable by Mankiw (2006) cited in the Introduction.

In contrast, when weights λ_θ are decreasing, favoring selfish parents, this creates a force for positive taxation of estates. The reverse is true when weights λ_θ are increasing, favoring altruistic parents, leading to a subsidy on estates. These results emphasize that ordinal preferences cannot settle the sign of estate taxes, which depends crucially on the weights λ_θ. The specification of cardinal utility or social welfare functions is crucial.

We now analyze the case where we allow for arbitrary weights on children. In the online Appendix we show that at points where the monotonicity constraint is not binding the implicit marginal tax rate on estates equals

$$(7)\quad \tau(\theta) = -\nu\alpha_\theta R U^c_{c_1}(\theta) - \nu\frac{\mu(\theta)}{f(\theta)} R U^c_{c_1}(\theta)\left(\frac{U^p_{\theta,c_1}(\theta)}{U^p_{c_1}(\theta)} - \frac{U^p_{\theta,c_0}(\theta)}{U^p_{c_0}(\theta)}\right),$$

where $\nu > 0$ is the multiplier on the promise keeping constraint (6) and $\mu(\theta)$ is the costate variable associated with (5), satisfying $\mu(\underline{\theta}) = \mu(\bar{\theta}) = 0$. This formula is equivalent to the one in Farhi and Werning (2010), except for the term involving $\mu(\theta)$.

PROPOSITION 2: *Suppose that Assumptions 1 and 2 hold. Suppose no bunching at the extremes. Then (i) marginal tax rates are negative at the extremes $\tau(\underline{\theta}) < 0$, $\tau(\bar{\theta}) < 0$; (ii) if α_θ is constant or decreasing then $\tau(\underline{\theta}) < \tau(\bar{\theta})$.*

This result indicates that, unless we place zero weight on children, a force for subsidies is always present. It also highlights a force for progressive taxation, in the sense of a rising marginal tax rate. Both features are in line with the main results in Farhi and Werning (2010). Indeed, there are parental weights that lead to exactly the same formula as in this canonical tax model. These parental weights are precisely those such that the first-best is incentive compatible so that $\mu(\theta) = 0$.

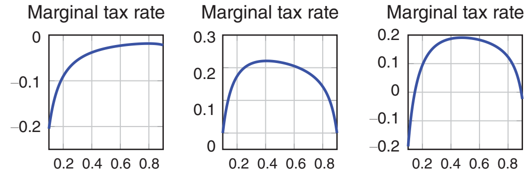

FIGURE 1. OPTIMAL IMPLICIT MARGINAL ESTATE TAX RATES $\tau(\theta)$ AS A FUNCTION OF θ FOR THE WEIGHTED UTILITARIAN CASE

Notes: The Pareto weights λ_θ and α_θ are as follows: $\lambda_\theta = 1$ and $\alpha_\theta = 0.02$ (first panel), $\lambda_\theta = \frac{e^{-\theta}}{E[e^{-\theta}]}$ and $\alpha_\theta = 0$ (second panel) and $\lambda_\theta = \frac{e^{-\theta}}{E[e^{-\theta}]}$ and $\alpha_\theta = 0.02$ (third panel).

PROPOSITION 3: *For constant weights on children $\alpha_\theta = \alpha \geq 0$, there exist parental weights λ_θ such that $\tau(\theta) = -\nu\alpha_\theta R U^c_{c_1}(\theta)$ for all θ.*

Numerical Illustration.—Figure 1 collects a few illustrative examples, using logarithmic utility $U^p(c_0, c_1; \theta) = (1-\theta)\log(c_0) + \theta \log(c_1)$, $U^c(c_1) = \log(c_1)$ and a uniform distribution for θ over $[0.1, 0.9]$.

The first panel in Figure 1 has constant positive weights on both parents and children. Proposition 2 leads us to expect negative tax rates near the extremes. In this example, tax rates remain negative throughout and are increasing in most of the range. This outcome is essentially as in Farhi and Werning (2010).

The second panel puts no weight on children, but assumes a decreasing weight on parents λ_θ. Tax rates are positive throughout, as expected from Proposition 1 part (ii). The third panel combines this decreasing λ_θ with a constant and positive weight α_θ ; tax rates are negative near the extremes, but positive over an interior interval.

B. *A Rawlsian Criterion for Children*

We now evaluate the welfare of children using a Rawlsian criterion. This amounts to studying the same planning problem in (3)–(6) with the additional constraint

$$U^c(c_1(\theta)) \geq \underline{u}.$$

Define θ^* to be the highest value of θ for which this constraint holds with equality. For high enough $\underline{u}$ we have $\theta^* > \underline{\theta}$. All types $\theta \in [\underline{\theta}, \theta^*]$ are bunched, so the implicit marginal tax $\tau(\theta)$ is increasing in θ by single crossing. Thus, $\tau(\theta) \leq \tau(\theta^*)$ for all $\theta \leq \theta^*$. Indeed, it is possible that $\tau(\theta) < 0$ for some $\theta < \theta^*$ even if $\tau(\theta) \geq 0$ for $\theta \geq \theta^*$. We now show that, indeed, tax rates are positive above θ^*.

PROPOSITION 4: *Suppose Assumptions 1, 2, and 3 hold. Suppose further that λ_θ is constant and that c_1 is a normal good. Then marginal taxes are positive $\tau(\theta) \geq 0$ for $\theta \geq \theta^*$ and strictly positive over a positive measure of θ.*

Even though the weight on parents is constant, the optimum involves positive taxation wherever the Rawlsian constraint is slack. Intuitively, children with bequests above the minimum do not contribute towards the maximin criterion, so they are taxed to redistribute towards the poorest children, as well as their selfish parents, who may otherwise be hurt by the imposition to improve their children's welfare. The implicit marginal tax rate at the bottom may or may not be negative, but it is positive for $\theta \geq \theta^*$. Given Proposition 1 part (i), positive taxes can be entirely attributed to placing a positive weight on children.[4] The second panel in Figure 2 illustrates this result. In this example, the implicit tax in the bunching region indeed becomes negative for low enough θ.

The optimal allocation has bunching below θ^*, so the tax schedule T must feature a kink, with marginal tax rates jumping upward. Indeed, it may require a marginal subsidy, coming from the left. A simple alternative implementation can avoid this by imposing the same budget constraints (1)–(2) but adding the constraint that $B \geq \underline{B}$.[5] By a suitable choice of lump-sum transfers, determining I_0 and I_1, we can normalize $\underline{B} = 0$. The tax code then imposes only positive marginal tax rates, but negative implicit taxes may be generated by the borrowing constraint, $B \geq 0$. Strictly positive taxes and the outlawing of negative bequests are common features of policy across developed countries.

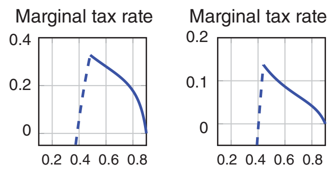

FIGURE 2. OPTIMAL IMPLICIT MARGINAL ESTATE TAX RATES $\tau(\theta)$ AS A FUNCTION OF θ WITH A RAWLSIAN CRITERION FOR GENERATION 1.

Notes: The Pareto weights λ_θ are as follows: $\lambda_\theta = \frac{e^{-\theta}}{E[e^{-\theta}]}$ (first panel), $\lambda_\theta = 1$ (second panel). The dashed portion coincides with values θ for which the borrowing constraint is binding ($\theta \leq \theta^*$). For these values, the implicit marginal tax rate $\tau(\theta)$ is lower than the explicit marginal tax rate $\tau(\theta^*)$ that agents face in the implementation with a nonlinear tax and a borrowing limit, reflecting the binding borrowing constraint.

IV. Linear Taxes and Limits to Borrowing

We now restrict estate taxes to be linear. The planner taxes bequests at a constant rate τ, balancing its budget with a lump-sum tax (or transfer). We also consider the imposition of constraints on borrowing that limit parents from passing on debt to their children. To keep things simple, we start by discussing the logarithmic utility case. We then provide tax formulas for general preferences.

A. *A Weighted Utilitarian Objective*

We first consider the case with the weighted Utilitarian criterion and no borrowing limits. The planning problem, stated in the online Appendix, is relatively straightforward and maximizes our welfare criterion subject to the resource constraint. The first-order conditions deliver a useful tax formula.

PROPOSITION 5: *Assume logarithmic utility $U^p(c_0, c_1; \theta) = (1-\theta)\log(c_0) + \theta\log(c_1)$, $U^c(c_1) = \log(c_1)$. The optimal linear estate tax is given by*

$$\frac{\tau}{1+\tau} = -\frac{\nu}{I}\,\frac{\frac{\mathrm{Cov}(\theta, \lambda_\theta + \alpha_\theta)}{\int\theta(1-\theta)f(\theta)\,d\theta} + \frac{\int\alpha_\theta(1-\theta)f(\theta)\,d\theta}{\int\theta(1-\theta)f(\theta)\,d\theta}}{1 + \frac{\mathrm{Var}(\theta)}{\int\theta(1-\theta)f(\theta)\,d\theta}}.$$

[4] Formula (7) can still be applied with endogenous positive weights on children α_θ that are decreasing in θ and are zero for all $\theta > \theta^*$; the costate $\mu(\theta)$ negative and zero at the extremes.

[5] This implementation is natural because it highlights that the optimal allocations will typically feature parents below θ^* bunched to satisfy the Rawlsian constraint $U^c(c_1(\theta)) \geq \underline{u}$. These same allocations could be obtained with an appropriate kink in the T function, typically requiring a sufficiently high subsidy rate to the left of the bunching point.

The numerator is the sum of a Ramsey covariance term and a Pigouvian average term. The term in the denominator is a Ramsey adjustment.[6] Roughly speaking, the Ramsey terms reflect the costs and benefits of redistribution across dynasties, while the Pigouvian term reflects the value of redistribution from parents to children. The Pigouvian term has a corrective nature because when social welfare places direct weight on children, parents necessarily undervalue bequests. When $\alpha_\theta = 0$, the Pigouvian term vanishes, leaving only the Ramsey terms, and the formula specializes to a version of the many-person Ramsey tax problem of Diamond (1975).

The Ramsey covariance term in the numerator may be positive or negative and neatly highlights the importance of the weights λ_θ and α_θ. The Pigouvian average term in the numerator is always negative or zero, providing a force for a subsidy as long as children have positive weight. The Ramsey adjustment term in the denominator only scales taxes proportionately but does not affect their sign.

If both weights are constant then the covariance term is 0 and the second term takes over. If we further assume that $\alpha_\theta = 0$ then the optimal tax is zero, $\tau = 0$, in line with Proposition 1 part (i); if, on the contrary, we place a positive and constant weight on children the optimal tax is a subsidy: $\tau < 0$. This linear tax result is consistent with the nonlinear results on negative marginal tax rates at the extremes in Proposition 2.

When α_θ and λ_θ are not constant the covariance term is not zero and decreasing weights provide a force for a tax. Whether the optimal tax is positive or negative depends on the net effect of the two terms in the numerator.

This formula can be generalized away from logarithmic utility. We define the after tax interest rate $\tilde{R} = \frac{R}{1+\tau}$, the uncompensated demand functions $c_0(I, \tilde{R}, \theta)$ and $c_1(I, \tilde{R}, \theta)$, the compensated elasticity $\varepsilon_{c_1,\tilde{R}}(I, \tilde{R}, \theta)$ of c_1 to the after tax interest rate $\tilde{R}$, the indirect utility function $V^p(I, \tilde{R}, \theta)$, and $W(I, \tilde{R}, \theta) = \lambda_\theta V^p(I, \tilde{R}, \theta) + \alpha_\theta\, U^c(c_1(I, \tilde{R}, \theta))$.

[6] This adjustment term encapsulates the impact on tax revenues of the income effect associated with a marginal tax change.

PROPOSITION 6: *For general preferences, the optimal linear estate tax is given by*

$$\frac{\tau}{1+\tau} = -\nu \frac{\dfrac{\mathrm{Cov}(c_1(\theta), W_I(\theta))}{\int \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta)\, d\theta} + \tilde{R}\, \dfrac{\int \alpha_\theta U^c_{c_1}(\theta)\ \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta)\, d\theta}{\int \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta)\, d\theta}}{1 + \dfrac{1}{\tilde{R}}\, \dfrac{\mathrm{Cov}(c_1(\theta), c_{1,I}(\theta))}{\int \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta)\, d\theta}}.$$

The formula takes the form of a ratio as in the logarithmic utility case, with similar terms in the numerator and denominator. The formula highlights the role of the interest rate elasticity of bequests. Basically, the Ramsey terms are hit by the inverse of the elasticity of bequests, while the Pigouvian term is not. More precisely, the Pigouvian term is a weighted average, and the interest rate elasticity of bequests affects only the corresponding weights. In this sense, the inverse-elasticity rule applies to the Ramsey terms as in Diamond (1975), but not to the Pigouvian term. Indeed, the average Pigouvian term is best thought as representing a Pigouvian motive for taxation. And to a large extent, Pigouvian taxes do not depend on elasticities.

B. *A Rawlsian Criterion for Children*

We now evaluate children's welfare according to a Rawlsian maximin criterion, exactly as in Section IIIB. In addition to a linear tax on bequests we provide the planner with one additional instrument: a minimum bequest requirement $\underline{B}$. As in Section IIIB, appropriate intergenerational transfers allow us to normalize and set $\underline{B} = 0$, so we can interpret this as a constraint that outlaws parents passing on debt to their children. We assume that the Rawlsian constraint is binding, which is the case for high enough $\underline{u}$.

PROPOSITION 7: *Assume logarithmic utility* $U^p(c_0, c_1; \theta) = (1-\theta)\log(c_0) + \theta\log(c_1)$ *and* $U^c(c_1) = \log(c_1)$. *Suppose* λ_θ *is constant and that children's welfare is evaluated by a Rawlsian maximin criterion. Then the optimum is such that the tax rate is strictly positive* $\tau > 0$ *and a borrowing constraint is strictly binding for some agents.*

The optimum features a tax coupled with a borrowing limit.[7] This result is a linear counterpart to the nonlinear conclusions in Proposition 4. The economic logic is similar: the revenue from a positive tax is used to improve the welfare of the poorest children, as well as the welfare of parents that are hurt by the imposition of the borrowing constraint.

Proposition 7 requires logarithmic utility. We now provide a more general related local result. Although it does not fully settle the sign, this result does suggest that positive estate taxes may be optimal for a wide class of preferences.

PROPOSITION 8: *Suppose* λ_θ *is constant and that children's welfare is evaluated by a Rawlsian maximin criterion. Suppose that Assumptions 1 and 2 hold. In addition, assume that* c_1 *is a normal good. There exists a positive tax* $\tau > 0$ *that improves on the no-intervention equilibrium with* $\tau = 0$.

We also provide an optimal tax formula for general preferences. We need to adapt the definitions of the demand functions, the indirect utility function and the interest rate elasticity of bequests to incorporate a borrowing constraint (see the online Appendix).

PROPOSITION 9: *For general preferences, the optimal tax rate is given by*

$$\frac{\tau}{1+\tau} = -\nu \frac{\dfrac{\mathrm{Cov}(c_1(\theta), \lambda_\theta V_I^p(\theta))}{\int \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta)\, d\theta}}{1 + \dfrac{1}{\tilde{R}} \dfrac{\mathrm{Cov}(c_1(\theta), c_{1,I}(\theta))}{\int \varepsilon_{c_1,\tilde{R}}(\theta) c_1(\theta) f(\theta)\, d\theta}}.$$

This optimal tax formula features only Ramsey terms and no Pigouvian term: the Pigouvian motive for taxation is addressed entirely by the borrowing constraint.

V. Conclusions

We have singled out one case where optimal policy takes a simple form: a ban on negative bequests and a positive tax on positive ones. These properties are features of tax codes in most developed economies. However, this result applies to a particular, albeit defensible, combination of welfare criteria (maximin for children) and cardinal normalizations. The conclusions are sensitive to the form of redistributive tastes, embedded in assumptions on the cardinality of utility and social welfare functions, as well as the source of the inequality in bequests, such as altruism heterogeneity versus parental earnings heterogeneity.

[7] If λ_θ is decreasing in θ, there is an additional force for a tax.

REFERENCES

Connell, Christopher, and Eric B. Rasmusen. 2012. "Concavifying the QuasiConcave." Unpublished.

Diamond, Peter A. 1975. "A Many-Person Ramsey Tax Rule." *Journal of Public Economics* 4 (4): 335–42.

Farhi, Emmanuel, and Ivan Werning. 2010. "Progressive Estate Taxation." *Quarterly Journal of Economics* 125 (2): 635–73.

Lockwood, Benjamin B., and Matthew C. Weinzierl. 2012. "De Gustibus non est Taxandum: Theory and Evidence on Preference Heterogeneity and Redistribution." National Bureau of Economic Research Working Paper 17784.

Mankiw, Gregory N. 2006. "The Estate Tax Debate." *Greg Mankiw's Blog*, June 5. (http://gregmankiw.blogspot.com/2006/06/estate-tax-debate.html)

Milgrom, Paul, and Ilya Segal. 2002. "Envelope Theorems for Arbitrary Choice Sets." *Econometrica* 70 (2): 583–601.

Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38 (114): 175–208.

Piketty, Thomas, and Emmanuel Saez. 2012. "A Theory of Optimal Capital Taxation." National Bureau of Economic Research Working Paper 17989.