Progressive Estate Taxation[*]

Emmanuel Farhi
MIT
efarhi@mit.edu

Iván Werning
MIT
iwerning@mit.edu

May 22, 2008 (4:50pm)

Abstract

We study efficient allocations in a Mirrleesian model with altruistic parents and focus on the implications for estate taxation. We show that optimal estate taxes have two important features. First, taxation should be progressive, so that more productive parents face a lower net return on bequests. Second, marginal taxes should be negative, so that parents face a marginal subsidy on bequests. We show that these features can be implemented using a simple nonlinear estate tax schedule, independent of income taxation. These prescriptions are shown to apply more broadly to other intergenerational transfers, such as education and human capital investments. Our results can be seen as generalizing the notion of non-inheritable debt, which obtains as a special case when the welfare criterion is Rawlsian. In extensions with heterogeneous family size, we show that inheritance taxation has an advantage over estate taxation.

1 Introduction

Arguably, the biggest risk in life is the family one is born into. Newborns partly inherit the luck, good or bad, of their parents and ancestors through the wealth accumulated within their dynasty. This makes them concerned not only with their own uncertain skills and earning potential, but also with that of their progenitors. They value insurance, from behind the veil of ignorance, against these risks. On the other hand, altruistic parents are partly motivated by the impact their efforts can have on their children's well-being through bequests. The intergenerational transmission of welfare determines the balance between insuring newborns and providing parental incentives.

One instrument that regulates this intergenerational transmission of welfare is estate taxation. This paper examines the optimal design of estate taxation by characterizing Pareto efficient allocations in an economy that captures the aforementioned tradeoff, between insuring children and creating incentives for parents. In our model, the estate tax affects the degree of inheritability of welfare.

We begin with a simple Mirrleesian economy with two generations linked by parental altruism (later we extend the analysis to an infinite horizon setting). In this economy, a continuum of parents live during the first period. In the second period each is replaced by a single child. Parents are altruistic towards their child, they work, consume and bequeath; children simply consume. Following Mirrlees (1971), parents first observe a random productivity draw and then exert work effort. Both productivity and work effort are private information; only output, the product of the two, is publicly observable. Our first objective is to study the entire set of constrained Pareto efficient allocations and derive their implications for marginal tax rates.

For this economy, if one assumes that the social welfare criterion coincides with the parent's expected utility, then Atkinson and Stiglitz's (1976) celebrated uniform-taxation result applies and implies that the intertemporal consumption choice made by parents is best left undistorted. That is, when no direct weight is placed on the welfare of children, parental labor income should be taxed non-linearly but bequests should go untaxed.

While this describes one efficient allocation, the picture is incomplete. In this economy, parent and child are distinct individuals, albeit linked by altruism. In positive analyses it is common to subsume both in a single fictitious 'dynastic agent'. However, a complete normative analysis must distinguish the welfare of both parents and children (Phelan, 2006; Farhi and Werning, 2007). Figure 1 depicts our economy's Pareto frontier, plotting the ex-ante expected utility for the child on the horizontal axis, and that of the parent on the vertical axis. The arrangement discussed in the previous paragraph corresponds to the peak
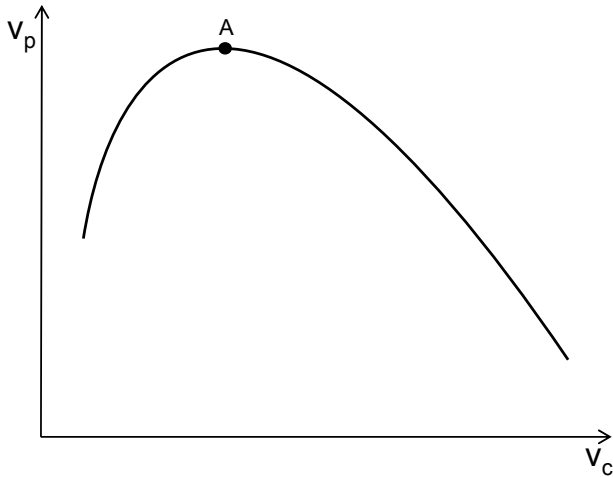
Figure 1: Pareto frontier between ex-ante utility for parent, v_p, and child, v_c.

marked as point A—which occurs at an interior point since parents are altruistic.

This paper explores other efficient points, those on the downward sloping section of the Pareto frontier. Indeed, to the right of point A, a role for estate taxation emerges. Our main result is that efficient estate taxation has two crucial features.

The first feature concerns the shape of marginal tax rates: we show that estate taxation should be progressive. That is, more fortunate parents, leaving larger bequests, should face a higher marginal tax on their bequests. Since more fortunate parents face lower returns on their bequests than the less fortunate ones, their bequests are more similar than would be otherwise. This mechanism generates mean reversion in consumption across generations, which helps lowers consumption inequality for newborns. This arrangement still provides incentives to parents, since the child's consumption varies with that of the parent, but it now varies less than one-for-one. In this way, progressive estate taxation mitigates the inheritability of luck within a dynasty.

Note that our result characterizes the optimal shape of the marginal tax rates on bequests, but does not imply that the overall tax system should be progressive. In fact, our analysis has nothing to say about the shape of labor income taxes. Moreover, our results on the estate tax apply regardless of the amount of redistribution across parents with different productivities.[1] Our stark conclusion regarding the progressivity of estate taxation contrasts with the well-known lack of sharp results regarding the shape of the optimal income tax schedule (Mirrlees, 1971; Seade, 1982; Tuomala, 1990; Ebert, 1992; Diamond, 1998; Saez, 2001).[2]

---

[1]We make this point rigorously in Section 4.1 where we show that our results about marginal tax rates on bequests hold for a large class of welfare functions for the generations of parents.

[2]Mirrlees's (1971) seminal paper established that for bounded distributions of skills the optimal marginal

3

The second feature concerns the level of marginal tax rates: estate taxation should be negative, taking the form of a subsidy. That is, the estate tax should impose a marginal subsidy (i.e. a negative tax) on bequests that decreases with bequests (i.e. a progressive tax). A subsidy encourages bequests to compensate newborns for the uncertainty introduced in their consumption. In other words, it would be inefficient to improve the welfare of children by simply reducing their consumption inequality, since this reduction comes at the expense of parental incentives. It is best to combine such a reduction with an increase in bequests. Or put differently, the first generation buys inequality from the second in exchange for a higher average consumption level, creating a Pareto improvement for both generations.

The second main objective of the paper is to derive an explicit tax system that implements these efficient allocations. We prove that a simple system, that confronts parents with separate nonlinear schedules for income and estate taxes, works. In our implementation the optimal estate tax schedule is decreasing and convex, reflecting our results regarding the sign and progressivity of the marginal tax. Thus, our results are not simply about implicit marginal tax rates or wedges, but also about marginal tax rates in an explicit tax system.

We illustrate the flexibility of our basic model by extending it in a number of directions that allow us to address some key issues related to estate taxation. These extensions also help interpret our results for optimal policies and compare them to actual policies.

Our first extension shows that the optimal estate tax schedule is independent of the welfare criterion used to evaluate the parents' welfare and that the progressive estate tax result holds for all welfare functions that value equality of outcomes for the child's generation. In the Rawlsian case, where the planner is concerned with the lowest welfare within the children's generation, a particularly simple implementation that makes debt non-inheritable is possible. The optimum can be implemented by a combination of a non-linear labor income tax levied on the parent and a borrowing constraint that prevents parents from leaving negative estates to their child. In equilibrium, poor parents will be up against this constraint because they know that their child will enjoy higher consumption and they would like to borrow against this. The borrowing constraint prevents them from doing this.

Importantly, this implementation should not be seen as an exception to our estate tax results, but, rather, as a special limiting case. To see this, note that the corresponding implicit tax rates are actually negative and progressive: when the borrowing constraint binds it is as if the parent faces a higher shadow interest rate that discourages borrowing, the more so the more binding the borrowing constraint. Thus, the implicit tax is negative and higher

---

income tax rates are regressive at the top (see also Seade, 1982; Tuomala, 1990; Ebert, 1992). More recently, Diamond (1998) has shown that the opposite—progressivity at the top—is possible if the skill distribution is unbounded (see also Saez, 2001). In contrast, our results on the progressivity of the estate tax do not depend on any assumptions regarding the distribution of skills.

for poorer parents that find themselves more constrained by the lack of borrowing. This sheds light on a feature of estate policy that is taken for granted and not often emphasized in academic and policy debates. In most countries, such as the US, the law stipulates that descendants are not liable for the debt of their parents. This in itself contributes to an implicit progressive negative marginal tax in estate policy, regardless of the explicit estate tax schedule. This also puts our general result in context. The progressive tax schedule obtained for smoother, non-Rawlsian welfare functions can be seen as generalizing the feature provided by a borrowing constraint, by creating a smoother incentive for parents to accumulate that is more intense for poorer parents.

We then incorporate human capital so that parents can make transfers to their children in two ways: through a better education, or through a larger bequest. In our model this margin should not be distorted. Thus, the marginal tax on financial bequests and on human capital investment should be equated. Since the optimal estate tax is a progressive subsidy, it immediately implies the same for the optimal human capital tax.

Finally, we allow for heterogeneous family sizes and altruism in order to compare the merits of estate and inheritance taxation. We first show that efficient allocations can be implemented with a tax on estates paid by the parent, but that this schedule must depend on family size and altruism. We then show that a simpler implementation is possible where children pay taxes on their own inheritances using a schedule that is independent of family size and altruism. These extensions highlight an important advantage of inheritance taxes over estate taxes.

We then turn to an infinite-horizon version of our economy. This is important for at least two reasons. First, it allows us to make contact with a growing literature on dynamic Mirrleesian models (for references, see Golosov, Tsyvinski and Werning, 2006). Second, it provides an important motivation for weighing the welfare of future generations. The reason is that allocations that maximize the expected utility for the very first generation may be disastrous for the average welfare of distant generations. Atkeson and Lucas (1992) proved an immiseration result of this kind for a taste shock economy, showing that inequality rises steadily over time, with everyone converging to zero consumption. As a result, no steady state for the cross section of welfare and consumption exists; we provide an example where a similar result holds in our Mirrleesian economy. This extreme outcome motivates placing a positive weight on future generations. As we show here, a steady state then exists where inequality is constant, everyone enjoys positive consumption and there is social mobility.

Tax implementations are necessarily more involved in an infinite horizon setting, but our main results extend. In a lifecycle context, Kocherlakota (2005) proposed an implementation with linear taxes on wealth. The tax rate is a function of the history of labor income and has

the property that at any point the average tax in the next period is zero. In our intergenerational context, when future generations are not valued, Kocherlakota's characterization of wealth taxes can be immediately applied to estate taxes. This then provides a simple benchmark, similar to the one offered by Atkinson and Stiglitz (1976) in the two period economy. When future generations have positive weight, we show that efficient allocations can be obtained by augmenting Kocherlakota's implementation with a nonlinear estate tax schedule. The nonlinear schedule is history independent and similar to the one used in the two period economy: it is decreasing and convex. Thus, relative to the case where future generations are not valued, optimal estate taxation is negative and progressive.

Progressivity creates mean reversion across generations, playing a key role in ensuring that inequality remains bounded and that a steady state exists. In this infinite horizon setting our results regarding the sign of estate taxation should be interpreted with care. When future generations are not valued the "zero average tax" obtained by Kocherlakota (2005) actually reflects a strictly positive intertemporal wedge (Golosov et al., 2003) where savings are discouraged.[3] In this implementation, savings are discouraged by making the rate of return on bequests risky, instead of offering a lower but riskless return. Relative to the case where no weight is given to future generations, placing a positive weight does provide a force for subsidies that encourages bequests. However, the starting point, with no weight on future generations, already discourages bequests and features a positive intergenerational wedge. Thus, when future generations are valued, the sign of the intergenerational wedge may still be positive. This should be kept in mind when interpreting our results on the sign of estate taxes.

Although our approach is normative, it is interesting to compare the tax prescriptions we obtain with actual policies. Taken together, our results share similarities and differences with policy used in most countries. In the two period model, the two features we found are that marginal tax rates on bequests should be progressive and negative. Progressivity of marginal tax rates is, broadly speaking, a feature of actual tax policy in developed economies. For example, in the United States bequests are exempt up to a certain level, then taxed linearly at a positive rate. Our paper provides the first theoretical justification, to the best of our knowledge, for this common feature of policy.

As for the sign of marginal tax rates, one generally does not think of current policy as subsidizing estates. One interpretation is that our normative model stresses a connection between progressive and negative marginal tax rates which may be overlooked in current thinking on estate tax policy. However, it turns out that the comparison with actual policies is more nuanced. On the one hand, most explicit taxes impose positive marginal rates on

---

[3] The intertemporal or intergenerational wedge is explicitly defined in equation (25).

bequests, although a large fraction of bequests may lie below the exemption level and, thus, face a zero marginal tax rate. On the other hand, as we argued above, restrictions on debt inheritability constitute an implicit marginal subsidy on bequests, while educational policies provide an explicit subsidy for another important intergenerational transfer.

Thus, perhaps a better stylized summary of actual policy is that marginal tax rates on intergenerational transfers are negative at the bottom and possibly positive near the very top. Indeed, the infinite horizon version of our model suggests that a similar pattern is possible in terms of implicit intergenerational wedges. Obviously, because actual policies use tax systems that differ from the history-dependent implementation needed in the infinite horizon setting, it may also be sensible to compare the implicit intergenerational wedges, rather than just tax rates. As explained above, even when the welfare of future generations is not considered by the planner, optimal allocations feature a positive intertemporal wedge. More generally, in the infinite horizon setting, unlike the simple two period model, optimal policies may feature a positive intergenerational wedge, especially at the top.

Related Literature. Cremer and Pestieau (2001) also study optimal estate taxation in a two-period economy, but their results are quite different from ours. In particular, they find that marginal tax rates may be regressive and positive over some regions. These results are driven by their implicit assumption that parental consumption and work are complements, departing from the Atkinson-Stiglitz benchmark of separability, which is our starting point.[4] Kaplow (1995) and Kaplow (2000) discuss estate and gift taxation in an optimal taxation framework with altruistic donors or parents. These papers make the point that gifts or estates should be subsidized, but assume away unobserved heterogeneity and are therefore silent on the issue of progressivity.

Our work also relates to a number of recent papers that have explored the implications of including future generations in the social welfare criterion. Phelan (2006) considered a social planning problem that weighted all generations equally, which is equivalent to not discounting the future at all. Farhi and Werning (2007) considered intermediate cases, where future generations receive a geometrically declining weight. This is equivalent to a social discount factor that is less than one and higher than the private one. Sleet and Yeltekin (2005) have studied how such a higher social discount factor may arise from a utilitarian planner without commitment. However, none of these papers consider implications for estate taxation.

---

[4]In the main body of the paper, Cremer and Pestieau (2001) studies a model without work effort, with an exogenous wealth shock that is privately observed by parents. However, in their appendix they develop a more standard Mirrlees model with the assumption that parental consumption and work are complements.

2   Parent and Child: A Two Period Economy

Consider the following two-period economy. A continuum of parents live during period $t = 0$, with each succeeded by a single descendant, or child, in period $t = 1$. To keep things simple, in this two period version, parents work and consume, while children just consume.

At the beginning of period $t = 0$, parents learns their productivity θ_0, which is drawn from a distribution $F(\theta_0)$. They then produce n_0 efficiency units of labor, requiring work effort n_0/θ_0. An allocation is a triplet of functions $(c_0(\theta_0), c_1(\theta_0), n_0(\theta_0))$, where c_0 and n_0 represents the parent's consumption and output, and c_1 represents the child's consumption.

Each parent is altruistic towards her child. The utility of a parent with productivity θ_0 is given by

$$v_0(\theta_0) = u(c_0(\theta_0)) - h\left(\frac{n_0(\theta_0)}{\theta_0}\right) + \beta v_1(\theta_0), \tag{1}$$

with $\beta < 1$. The child's utility is simply

$$v_1(\theta_0) = u(c_1(\theta_0)). \tag{2}$$

The utility function $u(c)$ is increasing, concave and differentiable; the dis-utility function $h(n)$ is assumed increasing, convex and differentiable. Combining equations (1) and (2) gives $v_0(\theta_0) = u(c_0) + \beta u(c_1) - h(n_0/\theta_0)$, a standard expression showing that the parent's utility can be reinterpreted as that of a fictitious dynastic agent that lives in both periods with a discount factor β.

An allocation is *resource feasible* if

$$K_1 + \int_0^\infty c_0(\theta_0)\, dF(\theta_0) \le e_0 + \int_0^\infty n_0(\theta_0)\, dF(\theta_0),$$
$$\int_0^\infty c_1(\theta_0)\, dF(\theta_0) \le e_1 + RK_1.$$

Combining these two inequalities yields the present-value resource constraint

$$\int_0^\infty c_0(\theta_0)\, dF(\theta_0) + \frac{1}{R}\int_0^\infty c_1(\theta_0)\, dF(\theta_0) \le e_0 + \frac{1}{R}e_1 + \int_0^\infty n_0(\theta_0)\, dF(\theta_0). \tag{3}$$

If productivity were observable by the planner, first-best efficient allocations would equalize consumption within each generation and require parents with higher productivity to produce more. Instead, we assume productivity is privately observed by the parent. As a result, consumption varies with output to provide incentives. By the revelation principle, we can restrict attention to direct mechanisms. Agents announce reports of their productivity and

receive an allocation as a function of these reports. We say that an allocation is *incentive compatible* if the parent finds it optimal to reveal her shock truthfully:

$$u(c_0(\theta_0)) + \beta u(c_1(\theta_0)) - h\left(\frac{n_0(\theta_0)}{\theta_0}\right) \geq u(c_0(\theta_0')) + \beta u(c_1(\theta_0')) - h\left(\frac{n_0(\theta_0')}{\theta_0}\right) \quad \forall \theta_0, \theta_0'. \quad (4)$$

An allocation is *feasible* if it satisfies the resource constraint (3) and the incentive constraints (4).

Next, we define two utilitarian welfare measures

$$V_0 \equiv \int_0^\infty v_0(\theta_0)\, dF(\theta_0) \qquad \text{and} \qquad V_1 \equiv \int_0^\infty v_1(\theta_0)\, dF(\theta_0).$$

Note that

$$V_0 = \int_0^\infty \left(u(c_0(\theta_0)) - h(y(\theta_0)/\theta_0)\right) dF(\theta_0) + \beta V_1,$$

so that the utilitarian welfare of the second generation, V_1, enters that of the first generation, V_0, through the altruism of parents. In addition to this indirect channel, we allow the welfare of the second generation, V_1, to enter our planning problem directly.

We say that an allocation $(c_0(\theta_0),\ c_1(\theta_0),\ n_0(\theta_0))$ is *efficient* if it is feasible, delivers welfare (V_0, V_1) and there is no other feasible allocation $(\tilde{c}_0(\theta_0),\ \tilde{c}_1(\theta_0),\ \tilde{n}_0(\theta_0))$ that delivers welfare $(\tilde{V}_0, \tilde{V}_1)$ with $V_0 \leq \tilde{V}_0$ and $V_1 \leq \tilde{V}_1$ and at least one of these two inequalities holding strictly. Efficient allocations solve the *social planning problem*

$$\max V_0$$

subject to the resource constraint (3), the incentive-compatibility constraints (4) and

$$V_1 \geq \underline{V}_1. \quad (5)$$

The social planning problem is indexed by $\underline{V}_1$. For low enough values of $\underline{V}_1$ the constraint (5) is not binding and the social planning problem then maximizes parental welfare V_0 subject to feasibility. Let V_1^* be the corresponding level of welfare obtained by the second generation in the social planning problem when constraint (5) is not imposed. Then, constraint (5) is not binding for all $\underline{V}_1 \leq V_1^*$. The solution corresponds to the peak on the Pareto frontier illustrated in Figure 1. The second generation obtains a finite level of welfare V_1^* because they are valued indirectly, through the altruism of the first generation. For values of $\underline{V}_1 > V_1^*$, constraint (5) binds and the solution corresponds to the downward sloping section in the figure.

3   The Main Result: Progressive Estate Taxation

In this section we derive two main results for the two-period economy laid out in the previous section. For any allocation, define the implicit estate tax or wedge $\tau(\theta_0)$ by

$$1 = \beta R(1 - \tau(\theta_0)) \frac{u'(c_1(\theta_0))}{u'(c_0(\theta_0))}. \tag{6}$$

This equation represents an intertemporal-Euler equation with a distortion equal to $\tau(\theta_0)$. Our results focus on properties of this implicit tax or wedge. Our main result shows that, at an efficient allocation, the implicit estate tax is progressive. We also construct an explicit tax system that implements efficient allocations.

To derive an intertemporal-optimality condition, let ν be the multiplier on constraint (5), μ be the multiplier on the resource constraint (3) and form the corresponding lagrangian:

$$L \equiv \int_0^\infty [v_0(\theta_0) + (\beta + \nu)v_1(\theta_0)] \, dF(\theta_0) - \mu \int_0^\infty [c_0(\theta_0) + c_1(\theta_0)/R - n_0(\theta_0)] \, dF(\theta_0)$$

so that the social planning problem is equivalent to maximizing L subject to incentive constraints (4). Suppose an allocation is optimal and consider the following perturbation, at a particular point θ_0. Let $c_0^\varepsilon(\theta_0) = c_0(\theta_0) + \varepsilon$ and define $c_1^\varepsilon(\theta_0)$ as the solution to $u(c_0^\varepsilon(\theta_0)) + \beta u(c_1^\varepsilon(\theta_0)) = u(c_0(\theta_0)) + \beta u(c_1(\theta_0))$. This construction ensures that the incentive constraints are unaffected by ε. A first order necessary condition is that the derivative of L with respect to ε be equal to zero, which delivers:

$$1 = \beta R \left(1 + \frac{1}{\beta} \frac{\nu}{\mu} u'(c_0(\theta_0)) \right) \frac{u'(c_1(\theta_0))}{u'(c_0(\theta_0))}. \tag{7}$$

Our first result, derived from equation (7) setting $\nu = 0$, simply echoes the celebrated Atkinson-Stiglitz uniform-commodity taxation result for our economy.

Proposition 1. *The constrained-efficient allocation with $\underline{V}_1 \leq V_1^*$ has a zero implicit estate tax, $\tau(\theta_0) = 0$.*

Atkinson and Stiglitz (1976) showed that if preferences over a group of consumption goods is separable from work effort, then the tax rates on these goods can be set to zero. In our context, this result applies to the group (c_0, c_1) and implies a zero implicit estate tax, $\tau(\theta_0) = 0$.

For these efficient allocations there is perfect inheritability of welfare across generations, in the sense that the Euler equation $u'(c_0) = \beta R u'(c_1)$ implies that dynastic consumption is smoothed. For example, if the utility function is CRRA $u(c) = c^{1-\sigma}/(1-\sigma)$ then consumption

in both periods are proportional to each other, $c_1(\theta_0) = (\beta R)^{\frac{1}{\sigma}} c_0(\theta_0)$, or

$$\log c_1(\theta_0) - \log c_0(\theta_0) = \frac{1}{\sigma} \log(\beta R).$$

Thus, the consumption of the parent and child vary, across dynasties with different θ_0, one-for-one in logarithmic terms. Making the children's consumption depends on the parent's productivity θ_0 provides the parent with added incentives. The fact that consumption moves one-for-one is efficient because only the welfare of parents is considered. Children are used to provide parental incentives and are not insured against the risk of their parent's fortune because their expected welfare is of no direct concern.

In contrast, when $\underline{V}_1 > V_1^*$, so that $\nu > 0$, then equation (7) implies that the ratio of marginal utilities is not equalized across agents and the marginal estate tax must be nonzero. Indeed, since consumption increases with θ_0 estate taxation must be progressive.

Proposition 2. *For $\underline{V}_1 > V_1^*$ the constrained efficient allocation the implict estate tax is strictly negative, increasing in the parent's productivity θ_0 and given by*

$$\tau(\theta_0) = -\frac{1}{\beta}\frac{\nu}{\mu}u'(c_0(\theta_0)) \qquad or \qquad \frac{\tau(\theta_0)}{1 - \tau(\theta_0)} = -R\frac{\nu}{\mu}u'(c_1(\theta_0)). \tag{8}$$

The proposition provides two expressions for the implicit estate tax. The first one relates the tax to the parent's consumption, while the second relates it to the child's consumption. The progressivity of the estate tax is implied by both, since both $c_0(\theta_0)$ and $c_1(\theta_0)$ are increasing in θ_0.

To understand these results, consider the CRRA utility case again. The consumption of parent and child are related by

$$\log c_1(\theta_0) - \log c_0(\theta_0) = \frac{1}{\sigma}\log\left(1 + \frac{1}{\beta}\frac{\nu}{\mu}c_0(\theta_0)^{-\sigma}\right) + \frac{1}{\sigma}\log(\beta R). \tag{9}$$

The right hand side of equation (9) is strictly decreasing in $c_0(\theta_0)$. Thus, the child's consumption still varies with that of the parent's but less than one-for-one in logarithmic terms. In this way, the intergenerational transmission of welfare is imperfect, with consumption mean reverting across generations. Thus, when the expected welfare of the second generation taken into account, insurance is provided to the children's generation to improve their average welfare V_1.

The progressivity of the implicit estate tax reflects the mean reversion in the allocation. Fortunate parents, with high productivity, must face a lower net-of-tax return on bequests so that their dynastic consumption slopes downward. Likewise, poor parents, with low

productivity, require higher net-of-tax returns on bequests so that their consumption slopes upward.

Another intuition is based on interpreting our economy with altruism as an economy with an externality. In the presence of externalities, corrective Pigouvian taxes are always desirable. In the simplest case, the externality is assumed to affect utility through the average consumption of the externality producing good (e.g. pollution from gasoline consumption).

In our model, we can think of c_1 as a consumption good chosen by the parent that has a positive externality on the child. Since the externality is positive, a Pigouvian subsidy (negative tax) is generally optimal. However, the externality measure is average welfare $V_1 = \int u(c_1(\theta_0)) \, dF(\theta_0)$, instead of aggregate consumption $\int c_1(\theta_0) \, dF(\theta_0)$. Thus, in contrast to the standard externality case, a constant corrective subsidy is not optimal. In particular, since the utility function $u(c_1)$ is concave, the externality from V_1 is strongest for children that have relatively low consumption, implying that the Pigouvian subsidy should be highest for poorer parents. This explains the progressivity of the implicit tax $\tau(\theta_0)$.

Private information is not crucial for these arguments. What is needed is something that gives rise to differences in parental consumption. In our setup, private information plays a role because if productivity or effort were publicly observable, then the first best would be achievable, which equates consumption across parents of different productivity. In this sense, our results are derived from an interaction between redistributive and corrective motives for taxation.

It is worth stressing that, while our main result characterizes marginal tax rates on bequests as progressive, it says nothing about the overall progressivity of the tax system, or the extent of redistribution across parents with different productivities. Our analysis does not characterize the shape of labor income taxes. Indeed, roughly speaking, any incidence that our estate tax may have affecting the overall redistribution within the first generation could be counterbalanced by adjusting the income tax schedule. Our result is not about the overall degree of inequality. Rather, it is about shifting inequality from the second generation to the first.

We next show that we can implement efficient allocations with a simple tax system. We say that an allocation is *implementable by non-linear income and estate taxation $T^y(n_0)$ and $T^b(b)$* if, for all θ_0, the allocation $(c_0(\theta_0), c_1(\theta_0), n_0(\theta_0))$ solves

$$\max_{c_0, c_1, n_0} \left(u(c_0) + \beta u(c_1) - h\left(\frac{n_0}{\theta_0}\right) \right)$$

subject to

$$c_0 + b = n_0 - T^b(b) - T^y(n_0),$$
$$c_1 = Rb.$$

We now establish that there exist tax schedules that implement efficient allocations. Define the tax on estates as

$$T^{b\prime}(b) \equiv -R\frac{\nu}{\mu}u'(Rb) \tag{10}$$

with the arbitrary normalization that $T^b(0) = 0$. The estate tax and the inheritance tax are thus given by simple first order ODEs. These definitions guarantee that the marginal rate of taxation on intergenerational transfers is equal to the optimal intergenerational wedge τ.

The proof then exploits the fact that marginal tax rates are progressive to ensure that the bequest problem is convex. The proof is contained in the appendix.

Proposition 3. *The optimal allocation is implementable with a non-linear income tax and an estate tax. The estate tax T^b is decreasing and convex.*

We also propose another natural implementation with an inheritance tax instead of an estate tax. The difference is that the tax is paid by the child on the bequest Rb left by the parent. In this simple context, the difference between the two implementations is tenuous. A starker contrast will appear in Section 4, we extend our basic setup to incorporate richer features relevant to intergenerational transfers.

Similarly, we say that an allocation is *implementable by non-linear income and inheritance tax $\hat{T}^y(n_0)$ and $\hat{T}^b(Rb)$* if, for all θ_0, the allocation $(c_0(\theta_0), c_1(\theta_0), n_0(\theta_0))$ solves

$$\max_{c_0,c_1,n_0} \left(u(c_0) + \beta u(c_1) - h\left(\frac{n_0}{\theta_0}\right) \right)$$

subject to

$$c_0 + b = n_0 - \hat{T}^y(n_0),$$
$$c_1 = Rb - \hat{T}^b(Rb).$$

We define the inheritance tax by the following first order ODE:

$$\frac{\hat{T}^{b\prime}(Rb)}{1 - \hat{T}^{b\prime}(Rb)} \equiv -R\frac{\nu}{\mu}u'(Rb - \hat{T}^b(Rb)). \tag{11}$$

The following result parallels Proposition 3 for the case of an inheritance tax.

Corollary. *The optimal allocation is implementable with a non-linear income tax and an inheritance tax. The inheritance tax $\hat{T}^b$ is decreasing and convex.*

The proof of this corollary follows exactly the same steps as that of Proposition 3 and is omitted for brevity.

We have stated our results in terms of the implicit marginal tax rates, as well as a particular tax implementation. Two other implementations are worth briefly mentioning. First the optimal allocation can be implemented by a non-linear income tax and a progressive consumption subsidy in the second period: $T^{c_1}(c_1)$. Second, the optimal allocation can also be implemented with a non-linear income tax and a regressive consumption tax in the first period: $T^{c_0}(c_0)$. In this two period version of the model all these implementations seem equally plausible. However, in the infinite horizon setting, implementations that rely on taxation of consumption require marginal have tax rates $T^{c_t\prime}(c_t(\theta_0))$ which grow without bound. Although this is certainly feasible in the model, it seems like an unappealing feature due to considerations outside the scope of the model, such as tax evasion.[5] In any case, all these implementations share that the intertemporal choice of consumption will be distorted, so that the implicit marginal tax rate on estates is progressive and given by $\tau(\theta_0)$.

4 Extensions

Our main result highlights two properties of optimal policy. First, the marginal taxes on intergenerational transfers should be progressive. Second, these marginal tax rates should be negative, so that transfers are subsidized. In this section we discuss a few extensions of the basic model that provide insight into these two results.

We first consider a planning problem with more general welfare functions. We show that the way the welfare of the first generation is considered is of no importance for the main result. Instead, the crucial assumption underlying our results is a preference for equality in the child's generation. In the Rawlsian limit we show that optimal policy collapses to the imposition that parents cannot bequeath debt. We argue that this policy implements progressive and negative implicit marginal tax rates. This provides a new perspective for our results: the progressive tax schedules we obtain, are a smoother manifestation of the same principle underlying the prohibition of negative bequests.

We then extend the model to incorporate other parental transfers. In particular, we show that our results are relevant for human capital investments and imply that education should

---

[5] Moreover, in a multi-period extension where each agent lives for more than one period a consumption tax on annual consumption would not work, because the progressive intertemporal distortions should only be introduced across generations, not across a lifetime (Farhi et al., 2005).

be subsidized in a progressive manner. The rationale is that it is optimal not to distort the relative prices of the different forms of transfers: all forms of parental transfers should face the same marginal tax rates, and this tax rate should be negative and decreasing in parental wealth. Interestingly, education is subsidized in most countries. Indeed, basic education is typically more heavily subsidized than higher education.

In the third extension, we investigate the relative merits of estate and inheritance taxation. In our baseline model, each parent has a single child. Estate and inheritance taxes are then completely equivalent. This equivalence breaks down when family size is allowed to vary. We show that in this context, inheritance taxes allow for a simpler implementation with a single decreasing and convex inheritance tax schedule that does not depend on family size.

Finally, we extend the model to allow for heterogeneity in parental altruism. Perhaps surprisingly, we show that the bequests of low-altruism parents should not be more heavily subsidized: the estate tax can be made independent of the parent's altruism.

4.1 Non-inheritable Debt

An important feature of intergenerational policies is that children are not liable for their parent's debt. Here we argue that the borrowing constraint this imposes on parents creates implicit marginal subsidies that are progressive since poorer parents find the constraint more binding. Indeed, we show that when the welfare criterion for children is Rawlsian instead of utilitarian, the optimal allocation can be implemented with a borrowing constraint preventing parents from leaving debt to their child. In this way, the Rawlsian limit collapses the smoother progressive subsidies, that we show hold for more general welfare functions, around a single point at the borrowing constraint. Thus, our main results can be seen as generalizing the accepted principle that parents should not be allowed to borrow against their children.

Define two *social welfare functions* W_0 and W_1 for parents and children, respectively, by

$$W_0 = \int_0^\infty \hat{W}_0(v_0(\theta_0), \theta_0) \, dF(\theta_0)$$

$$W_1 = \int_0^\infty \hat{W}_1(v_1(\theta_0)) \, dF(\theta_0)$$

where $v_0(\theta_0) = u(c_0(\theta_0)) + \beta u(c_1(\theta_0)) - h(n_0(\theta_0)/\theta_0)$ and $v_1(\theta_0) = u(c_1(\theta_0))$. We assume that $\hat{W}_1$ is increasing, concave and differentiable, and that $\hat{W}_0(\cdot, \theta_0)$ is increasing and differentiable for all θ_0. The utilitarian case corresponds to the identity functions, $\hat{W}_0(v) = v$ and $\hat{W}_1(v) = v$, while the Rawlsian case obtains as the limit when $\hat{W}_1$ is infinitely concave.

Note that we allow the function $\hat{W}_0$ to depend on θ_0. For example, this permits a weighted utilitarian criterion $\hat{W}_0 = \pi(\theta_0)v_0(\theta_0)$ with Pareto weights $\pi(\theta_0)$. Essentially, our analysis only requires a welfare criterion that ensures Pareto efficiency for the first generation. In contrast, our results require the welfare criterion for the second generation to capture a preference for equality.

As in the utilitarian case, efficient allocations solve the social planning problem of maximizing W_0 subject to the resource constraint (3), the incentive-compatibility constraints (4) and $W_1 \geq \underline{W}_1$, for some $\underline{W}_1$. Using the same perturbation argument developed for the utilitarian case, we find that the implicit estate tax is given by

$$\tau(\theta_0) = -\frac{1}{\beta}\frac{\nu}{\mu}\hat{W}_1'(u(c_1(\theta_0))u'(c_0(\theta_0)) \quad \text{or} \quad \frac{\tau(\theta_0)}{1-\tau(\theta_0)} = -R\frac{\nu}{\mu}\hat{W}_1'(u(c_1(\theta_0))u'(c_1(\theta_0)). \quad (12)$$

Since $\hat{W}_1$ is concave and consumption for both parent and child, $c_0(\theta_0)$ and $c_1(\theta_0)$, are increasing in θ_0 it follows that $\tau^b(\theta_0)$ is increasing in θ_0. In other words, the estate tax is progressive. Note that τ does not depend directly on $\hat{W}_0$ but only indirectly through μ, ν, $c_0(\theta_0)$ and $c_1(\theta_0)$. Equation (12) illustrates that the progressivity of estate taxes is mainly determined by the welfare function for children $\hat{W}_1$ and its embedded concern for consumption equality. The welfare function for parents $\hat{W}_0$ on the other hand plays an important role in determining the shape of the income tax schedule and thereby the overall progressivity of the tax system.

Exactly as in Proposition 3, we can show that the optimal allocation is implementable with a non-linear income tax T^y and an estate tax T^b, as well as with a non-linear income tax $\hat{T}^y$ and an inheritance tax $\hat{T}^b$. Both the estate tax and the inheritance tax are decreasing and convex.

The case where the welfare function for children $\hat{W}_1$ for the generation of children is Rawlsian deserves special attention. In this case, the social planning problem is to maximize W_0 subject to the resource constraint (3), the incentive-compatibility constraints (4) and

$$u_1(\theta_0) \geq \underline{u}_1 \qquad \text{for all } \theta_0 \qquad (13)$$

for some $\underline{u}_1$. Let $\underline{c}_1$ be the corresponding consumption level: $\underline{c}_1 = u^{-1}(\underline{u}_1)$.[6]

Let $u_1^* = \min\{u_1(\theta_0)\}$ in the social planning problem just outlined but when the last constraints (13) are ignored. When $\underline{u}_1 > u_1^*$, then there exists $\underline{\theta}_0 > 0$ such that equation (13) is binding for all $\theta_0 \leq \underline{\theta}_0$. Then for all $\theta_0 \geq \underline{\theta}_0$, the implicit estate tax is zero. When $\theta_0 < \underline{\theta}_0$,

---

[6] Atkeson and Lucas (1995) studied a similar program in a setup with taste-shocks, where c_0 and c_1 are interpreted as consumption for the same individual.

the implicit estate tax is given by

$$\tau(\theta_0) = 1 - \frac{u'(c_0(\theta_0))}{\beta R u'(\underline{c}_1)}$$

which is always negative and progressive. In fact, the threshold $\underline{\theta}_0$ is such that $c_1(\underline{\theta}_0) = \underline{c}_1$ and $1 - \frac{u'(c_0(\underline{\theta}_0))}{\beta R u'(\underline{c}_1)} = 0$.

Our next result shows that efficient allocations in the Rawlsian case can be implemented by imposing an intergenerational debt constraint that binds for some agents. Under this implementation, an agent of type θ_0 faces the following program:

$$\max_{c_0,c_1,n_0} \left(u(c_0) + \beta u(c_1) - h\left(\frac{n_0}{\theta_0}\right) \right)$$

subject to

$$c_0 + b = n_0 - \hat{T}^y(n_0),$$
$$c_1 = Rb$$
$$b \geq \underline{c}_1/R.$$

Proposition 4. *Suppose that the welfare function for the children's generation is Rawlsian. There exists an income tax function T^y together with a debt constraint of the form $b \geq \underline{c}_1/R$ that implements the optimal allocation.*

Note that for those parents that find the debt constraint strictly binding, the intertemporal Euler equation holds with strict inequality $u'(c_0(\theta_0)) > \beta R u'(c_1(\theta_0))$. Thus, they face an implicit estate subsidy $\tau(\theta_0) < 0$. Parents with low enough θ_0 would like to borrow more against their kids, but the implementation precludes it. The lower the productivity θ_0, the lower $c_0(\theta_0)$ and the stronger is this borrowing motive. As a result, the shadow subsidy is strictly increasing in θ_0 over the range of parents that are at the debt limit. This is how the debt limit implements the progressivity in implicit marginal tax rates.

This implementation captures one feature of actual economies. In the US, for example, the system can be characterized by two features: (a) estates are taxed at a flat rate above an exemption level; and (b) children are not legally liable for the debt of their parents. Although, discussions about the progressivity of estate taxes usually revolve around (a), our model, in the Rawlsian case, highlights (b).

The Rawlsian case, and the debt limit it implies, can be fruitfully thought of as a limiting case of the previous welfare function. As the welfare function $\hat{W}_1$ becomes more concave, the social planning solution converges to that of the Rawlsian case. Similarly, the progres-

sive estate tax schedule $T^b(b)$ converges to a function that punishes low levels of bequests, effectively imposing a intergenerational debt constraint.

4.2 Educational Subsidies

In the previous subsection we reviewed how the standard policy of not allowing parents to leave debt to their children actually generates negative implicit marginal taxes on estates. In light of this, our result on negative marginal tax rates may seem less radical than it does at first sight. In this subsection we point out that another common policy explicitly subsidizes intergenerational transfers.

In our simple model, bequests were the only available means to transfer between one generation and the next. In reality, parents make transfers to their children in a number of ways. Human capital investments are a main form of giving, especially for all but the most affluent. Subsidies on education is one of the most pervasive government policies around the world. This example of a subsidy on intergenerational transfers suggests thinking about the conclusion of our simple model in broader terms. Interestingly, one rationale often offered for educational subsidies is that they help provide equality of opportunity for the next generation. This direct concern for the welfare of future generations is precisely what our model captures.

We now formalize this idea by incorporating the simplest form of human capital investments. Let x denote the investment and $H(x)$ denote the level of human capital achieved by this investment, where H is a differentiable, increasing and concave function satisfying the standard Inada conditions $H'(0) = \infty$ and $H'(\infty) = 0$. Each unit of human capital produces a unit of the consumption good, so that the resource constraint is

$$\int_0^\infty \left(c_0(\theta_0) + \frac{c_1(\theta_0)}{R} \right) dF(\theta_0) \leq \int_0^\infty \left(n_0(\theta_0) + \frac{H(x(\theta_0))}{R} - x(\theta_0) \right) dF(\theta_0)$$

Preferences are

$$v_0(\theta_0) = u^0(c_0(\theta_0)) - h\left(\frac{n_0(\theta_0)}{\theta_0} \right) + \beta v_1(\theta_0),$$

$$v_1(\theta_0) = u^1(c_1(\theta_0), H(x(\theta_0)))$$

where u^1 is differentiable, increasing and concave in both arguments, and satisfies standard Inada conditions. This structure of preferences preserves the weak separability assumption required for the Atkinson-Stiglitz benchmark result. The assumption that human capital enters the utility function is a convenient way of ensuring that not everyone makes the same

human capital investment. Indeed, we will assume human capital is a normal good, so that richer parents invest more.

The formula for the implicit estate tax is unaffected by the introduction of human capital

$$\tau(\theta_0) = -\frac{1}{\beta}\frac{\nu}{\mu}u^{0\prime}(c_0(\theta_0)) \quad \text{or} \quad \frac{\tau(\theta_0)}{1-\tau(\theta_0)} = -R\frac{\nu}{\mu}u^{1\prime}(c_1(\theta_0)).$$

We now turn to the implicit human capital tax. Consider the following perturbation of the optimal allocation at a particular point θ_0 let

$$x^\varepsilon(\theta_0) = x(\theta_0) + \varepsilon\frac{u^1_{c_1}(c_1(\theta_0), H(x(\theta_0)))}{u^1_H(c_1(\theta_0), H(x(\theta_0)))}\frac{1}{H'(x(\theta_0))}$$

and define $c_0^{\tilde{\varepsilon}}(\theta_0) = c_0$ and $c_1^{\tilde{\varepsilon}}(\theta_0)$ as the solution of

$$u^1(c_1^{\tilde{\varepsilon}}(\theta_0), H(x^\varepsilon(\theta_0))) = u^1(c_1(\theta_0), H(x(\theta_0))).$$

This perturbation leaves utility for both the parent and the child unaffected but impacts the resource constraint. This leads to the following first order condition:

$$H'(x(\theta_0)) = R\left(\frac{u^1_H(c_1(\theta_0), H(x(\theta_0)))}{u^1_{c_1}(c_1(\theta_0), H(x(\theta_0)))} + 1\right)^{-1} \tag{14}$$

Equation (14) is also the first order condition of

$$V_1(e) = \max_{c_1,x} u^1(c_1, H(x)) \qquad \text{subject to} \qquad c_1 + Rx - H(x) = e. \tag{15}$$

The quantity $c_1 - H(x)$ is the financial bequest received by the child, and x is human capital investment. Equation (14) implies that it is optimal not to distort the choice between these two forms of transfers from parent to child. In what follows, we assume that financial bequest and human capital investment are both normal goods. That is, the optimal $c_1 - H(x)$ and x in (15) are increasing in e.

We consider an implementation with three separate non-linear tax schedules: a nonlinear income tax schedule T^y, a non-linear estate tax T^b and non-linear human capital tax T^x. The parent maximizes

$$u^0(c_0) - h(n_0/\theta_0) + \beta u^1(c_1, H(x))$$

subject to

$$c_0 + b + x = n_0 - T^y(n_0) - T^b(b) - T^x(x),$$
$$c_1 = Rb + H(x).$$

We then have the following result.[7]

Proposition 5. *Assume that financial bequests and human capital investment are normal goods in (15). There exists three separate non-linear tax schedules T^y, T^b and T^x that implement the optimal allocation. In addition T^b and T^x are decreasing and convex. Moreover the marginal tax rates on bequests and human capital investment are equalized.*

The proposition shows that human capital taxation take the same form as estate taxation. Marginal tax rates are progressive and negative. This is because the choice between bequests and human capital should be undistorted. This requires human capital subsides and estate subsidies to be equalized:

$$T^{x\prime}(x(\theta)) = T^{b\prime}\left(\frac{c_1(\theta) - H(x(\theta))}{R}\right) = \frac{\tau(\theta)}{1 - \tau(\theta)}.$$

Indeed, many countries do employ various policies towards education and other forms of human capital that effectively help finance their investments. Typically basic education is provided for free, while higher levels of education are only partly subsidized, especially if one considers the opportunity cost of time component of investment and the quality of education. Hence, all parents are subsidized in their human capital investments, but parents that invest face a lower marginal subsidy.[8]

4.3 Estate vs. Inheritance Taxes

In the basic model each household has a single child. While this abstraction is immaterial to discuss the basic insights behind our main results, it does not allow us to compare estate versus inheritance taxes. In this section, we allow for more children and explore two new dimensions: differential altruism towards children within a household and variable number of children across households.

---

[7] If human capital does not enter utility, then all parents would make the same human capital investment, and equation equation (14) would reduce to $H'(x(\theta_0)) = R$. One cannot implement this with three separate tax schedules. An alternative implementation is to total wealth from tax human capital and bequests jointly, so that the child pays taxes as a function of total wealth $Rb + H(x)$.

[8] Richer parents may get more total subsidies, or even more average subsidies. The relevant comparison here is whether they face lower marginal subsidies on potential additional investments.

Suppose that parents can either have any number of children $m \in 1, 2, \ldots$. To simplify we assume that fertility is exogenous with joint distribution for fertility and productivity $F(\theta_0, m)$. We conjecture that our results carry over to an extension with endogenous fertility choice. The model is just like before, except that each child enters the utility of the parent through the altruism coefficient β_m, which allows parents with more children to care relatively less about each child (Becker and Barro, 1988). That is, a parent with m children and productivity θ_0 wishes to maximize total utility $u(c_0) - h(n_0/\theta_0) + \sum_{j=1}^{m} \beta_m u(c_{1,j})$ (optimal allocations will be symmetric across children within a family, so that $c_{1,j}$ is independent of j). The planning problem incorporates the welfare of the children's generation through the utilitarian criterion $\int u(c_1(\theta_0, m)) dF(\theta_0, m)$.

Because the same variational arguments that we used before can be applied conditioning on (θ_0, m), the implicit estate tax is again given by

$$\frac{\tau(\theta_0, m)}{1 - \tau(\theta_0, m)} = -R \frac{\nu}{\mu} u'(c_1(\theta_0, m)). \tag{16}$$

The marginal distortion reflects the planner's desire to insure children against two sources of risk—the parent's productivity and family size. However, because the origin of a child's luck is irrelevant in the welfare criterion, the optimal estate subsidy depends only on the child's consumption $c_1(\theta_0, m)$.

In this context, it is impossible to implement the optimal allocation with a non-linear income tax $T^{y,m}$ and an estate tax T^b that does not depend on the number of children m. To see this, suppose parents were confronted with such a system and that an estate of size b is equally divided among children to provide them each with consumption $c_1 = Rb/m$. Then, in such a system, families with different number of children m leaving the same estate level b would face the same marginal tax rate. But this contradicts equation (16) which shows that the marginal tax rate should be lower (i.e. a greater subsidy) for the larger family.

It is possible to implement the optimal allocation if the estate tax schedule is allowed to depend on family size m, so that parents face a tax schedule $T^{b,m}$. However, since the implicit tax in equation (16) depends on θ_0 and m only trough $c_1(\theta_0, m)$, it is possible to do the same with an inheritance tax that is independent of family size m. In this implementation, a parent with m children faces the budget constraint $c_0 + \sum_{j=1}^{m} b_j \leq n_0 - \hat{T}^{y,m}(n_0)$. Each child is then subject to the budget constraint $c_{1,j} \leq Rb_j - T^b(Rb_j)$.

Proposition 6. *There exists two separate non-linear tax schedules, a income tax $\hat{T}^{y,m}$ that depends on family size and an inheritance tax $\hat{T}^b$ independent of family size, that implement the optimal allocation. In addition $\hat{T}^b$ is decreasing and convex.*

The proof of this proposition proceeds exactly as the proof of Proposition 3 and Corollary

[3](#) with the inheritance tax defined as before by the ODE in equation ([11](#)).

We also explored an extension where parent care more about one child than the other; we omit the details but discuss the main features briefly. In this model, we assumed parents had two children indexed by $j \in \{L, H\}$ and let the altruism coefficient for child j be β^j, with $\beta^H \geq \beta^L$. The preference for one child over another may reflect the effects of birth order, gender, beauty, physical or intellectual resemblance. Our results are easily extended to this model. Once again, the model favors inheritance taxes over estate taxes because with an estate tax, the marginal tax on the two children would be equalized, even when their consumption is not. However, just as in equation ([16](#)) marginal tax rates should depend on the child's consumption.

4.4 Heterogeneous Altruism

In this section, we return to a setting with one child per household and explore an extension where parents are heterogeneous in their altruism β. The idea is to investigate whether the optimal estate tax should depends on the degree of altruism. One might imagine that the bequests of low-altruism parents should be more heavily subsidized, but our analysis shows that this is not optimal and that the estate tax can be made independent of the parent's altruism.

Assume that (θ_0, β) are jointly distributed with distribution function $F(\theta_0, \beta)$. We continue to assume that θ_0 is private information. In contrast, to keep things manageable we assume that β is observable to the planner; we later briefly discuss relaxing this assumption. The planning problem is then very similar to the case with fixed β, except that allocations must now be indexed by both θ_0 and β. In particular, a child's consumption $c_1(\theta, \beta)$ will generally depend on both variables: the parent's productivity θ_0 and the parent's altruism β. More fortunate children are born into more productive and caring parents. The arguments leading to equation ([8](#)) are unchanged and yield

$$\tau(\theta_0, \beta) = -\frac{1}{\beta}\frac{\nu}{\mu}\frac{u'(c_0(\theta_0, \beta))}{\beta} \qquad \text{and} \qquad \frac{\tau(\theta_0, \beta)}{1 - \tau(\theta_0, \beta)} = -R\frac{\nu}{\mu}u'(c_1(\theta_0, \beta)). \qquad (17)$$

This first formula makes clear that τ is decreasing in parental consumption c_0 and decreasing in altruism β. The richer or the more caring the parent, the lower is the estate subsidy. The second formula shows, however, that the marginal tax τ can be expressed as a function of the child's consumption c_1. The dependence on of θ_0 and β only enters indirectly.

Intuitively, newborns are insured against both sources of risk, θ and β. However, the origin of a child's luck is irrelevant, all that matters is the impact it has on the child's

consumption.

This observation allows us to implement the optimal allocation with the combination of an income tax $T^{y,\beta}(n_0)$ that depends on β and an estate tax $T^b(b)$ that does not, so that parents are facing the sequence of budget constraints

$$c_0 + b = n_0 - T^{y,\beta}(n_0) - T^b(b)$$
$$c_1 = Rb$$

Proposition 7. *There exists two separate non-linear tax schedules, a income tax $T^{y,\beta}$ that depends on the altruism parameter β and an inheritance tax T^b that is independent of β, that implement the optimal allocation. In addition T^b is decreasing and convex.*

The crucial part of this proposition states that the estate tax schedule is independent of β. This follows from the fact that the implicit marginal tax in equation (17) is a function of the child's consumption c_1, but does not depend directly on β. Given this remark, the proof of this proposition proceeds exactly as the proof of Proposition 3 with an estate tax function given by

$$T^{b\prime}(b) = -R\frac{\nu}{\mu}u'(Rb). \tag{18}$$

In general, the income tax schedule $T^{y,\beta}$ may depend on β. This dependence does two things. First, it allows redistribution across parents with different β. Second, it allows redistribution across productivity θ_0, within a group of parents with the same β, to depend on the value of β. It is important to note, however, that in some cases it is Pareto efficient to drop the dependence of $T^{y,\beta}$ on β, forgoing the conditioning of redistribution on altruism. More precisely, consider the general welfare analysis from subsection 4.1, allowing the welfare function $\hat{W}_0$ to depend on both θ_0 and β. Then there may exist a set of welfare functions $\hat{W}_0$ such that the optimal allocation is implementable with tax functions T^y and T^b that are independent of β and T^b is defined by equation (18).

How would things be different if the altruism coefficient β were unobservable? While we do not provide a complete answer to this question, we argue that the analysis may characterize efficient arrangements. That is, the analysis would is unaffected by the introduction of private information in precisely in the cases described above where the planner is willing to forgo conditioning redistribution on altruism when β is observable.[9]

---

[9]Characterizing the income tax schedules T^y that support the allocations at the intersection of the two frontiers is beyond the scope of this paper. Here we show how to reduce this problem to one that is studied in Werning (2008). Consider the Pareto problem with observable β and fix a value of the multiplier $\tilde{\nu}$ on the

5 A Mirrleesian Economy with Infinite Horizon

We now turn to a repeated version of this economy with an infinite horizon. We provide two alternative tax implementations and show that our main results regarding estate taxation carry through in both. We also discuss the role of estate taxation in shaping the dynamics of long run inequality.

5.1 An Infinite Horizon Planning Problem

An individual born into generation t has ex-ante welfare v_t with

$$v_t = \mathbb{E}_{t-1}[u(c_t) - h(n_t/\theta_t) + \beta v_{t+1}] = \sum_{s=0}^{\infty} \beta^s \, \mathbb{E}_{t-1}\big[u(c_{t+s}) - h(n_{t+s}/\theta_{t+s})\big], \qquad (19)$$

where θ_t indexes the agent's productivity type and $\beta < 1$ is the coefficient of altruism. We assume that the utility function satisfies the Inada conditions $u'(0) = \infty$, $u'(\infty) = 0$, $h'(0) = 0$ and $h'(\bar{n}) = \infty$, where $\bar{n}$ is the (possibly infinite) upper bound on work effort. We assume that types θ_t are independently and identically distributed across dynasties and generations $t = 0, 1, \ldots$ With innate talents assumed non-inheritable, intergenerational transmission of welfare is not mechanical linked through the environment but may arise to provide incentives for altruistic parents.

Since productivity shocks are assumed to be privately observed by individuals and their descendants we need to impose incentive compatibility. Each dynasty faces a sequence $\{c_t, n_t\}$, where $c_t(\hat{\theta}^t)$ and $n_t(\hat{\theta}^t)$ represent consumption and effective units of labor as a

---

utility of children. Let T^b be defined by equation (18) where $\tilde{\nu} = \hat{\nu}\,\mu$ and consider the following problem:

$$V(I; \beta) = \max u(c_0) + \beta u(c_1)$$

subject to $c_0 + c_1/R - T^b(c_1/R) \leq I$. Let $(c_0(I; \beta), c_1(I; \beta))$ be the solution of this program and define $e(I; \beta) \equiv c_0(I; \beta) + c_1(I, \beta)/R$. It is easy to establish that both c_0 and c_1 are normal goods i.e. that both $c_0(I; \beta)$ and $c_1(I; \beta)$ are strictly increasing in I. This in turn implies that $e(I; \beta)$ is strictly increasing in I. We can therefore perform a change of variable and define $\hat{V}(e, \beta) \equiv V(e^{-1}(e; \beta))$ where $e^{-1}(e; \beta)$ is the inverse of the function that maps I into $e(I; \beta)$. We can then restate the Pareto problem as

$$\min \int_0^{\infty} [e(\theta_0, \beta) - y(\theta_0, \beta)]\, dF(\theta_0, \beta),$$

subject to

$$\hat{V}(e(\theta_0, \beta)) - h(y(\theta_0, \beta)/\theta_0) \geq \underline{v}(\theta_0, \beta)$$

and

$$\hat{V}(e(\theta_0, \beta)) - h(y(\theta_0, \beta)/\theta_0) \geq \hat{V}(e(\theta'_0, \beta')) - h(y(\theta'_0, \beta')/\theta_0).$$

Werning (2008) gives conditions on primitives for the existence of a distribution of utility functions $\underline{v}(\theta_0, \beta)$ such that the corresponding optimal allocation is implemented by an income tax that does not depend on β. These allocations span the intersection of the unconstrained and constrained Pareto frontiers.

function of the history of reports $\hat{\theta}^t \equiv (\hat{\theta}_0, \hat{\theta}_1, \ldots, \hat{\theta}_t)$. A dynasty's reporting strategy $\sigma \equiv \{\sigma_t\}$ is a sequence of functions $\sigma_t : \Theta^{t+1} \to \Theta$ that maps histories of shocks θ^t into a current report $\hat{\theta}_t$. Any strategy σ induces a history of reports $\sigma^t : \Theta^{t+1} \to \Theta^{t+1}$. We use σ^* to denote the truth-telling strategy with $\sigma_t^*(\theta^t) = \theta_t$ for all $\theta^t \in \Theta^{t+1}$.

Given a sequence $\{c_t, n_t\}$, the utility obtained from any reporting strategy σ is

$$U\big(\{c_t, n_t\}, \sigma; \beta\big) \equiv \sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \beta^t [u\big(c_t\big(\sigma^t(\theta^t)\big)\big) - h(n(\sigma^t(\theta^t))/\theta_t)] \Pr(\theta^t).$$

Incentive compatibility amounts to requiring that truth-telling be optimal: $U\big(\{c_t, n_t\}, \sigma^*; \beta\big) \geq U\big(\{c_t, n_t\}, \sigma; \beta\big)$ for all σ.

We identify dynasties by their initial utility entitlement v with distribution ψ in the population. An allocation is a sequence of capital stocks $\{K_t\}$ and a sequence of functions $\{c_t^v, n_t^v\}$ for each v. For any given initial distribution of entitlements ψ, we say that an allocation $(\{c_t^v, n_t^v\}, \{K_t\})$ is *feasible* if: (i) $\{c_t^v, n_t^v\}$ incentive compatible and delivers expected utility v

$$v = U\big(\{c_t^v, n_t^v\}, \sigma^*; \beta\big) \geq U\big(\{c_t^v, n_t^v\}, \sigma; \beta\big) \quad \text{for all } \sigma, v. \tag{20}$$

and (ii) it satisfies the resource constraints

$$C_t + K_{t+1} \leq F(K_t, N_t) \qquad t = 0, 1, \ldots \tag{21}$$

where $C_t \equiv \int_0^{\infty} \sum_{\theta^t} c_t^v(\theta^t) \Pr(\theta^t) \, d\psi(v)$ and $N_t \equiv \int \sum_{\theta^t} n_t^v(\theta^t) \Pr(\theta^t) \, d\psi(v)$ are aggregate consumption and labor, respectively. We assume the production function $F(K, N)$ is strictly increasing and continuously differentiable in both of its argument, exhibits constant returns to scale and satisfies the usual Inada conditions.

We now consider efficient allocations indexed by a distribution of initial utility entitlements ψ for the first generation and a minimum average welfare level $\underline{V}$ for future generations. Efficiency allocations minimize the resource cost of delivering these promises. More precisely, given ψ and $\underline{V}$, efficient allocations solve the following planning problem:

$$\min K_0 \tag{22}$$

over $\big(\{c_t^v, n_t^v\}; \{K_t\}\big)$ subject to (20), (21) and the admissibility constraints

$$\int U\big(\{c_{t+s}^v, n_{t+s}^v\}_{s \geq 0}, \sigma^*; \beta\big) d\psi(v) \geq \underline{V} \qquad t = 1, 2, \ldots \tag{23}$$

This is a Pareto problem between current and future generations. Let $V^* \equiv (u(0) -$

$\mathbb{E}[h((\bar{n})/\theta)])/(1-\beta)$ be the welfare associated with misery.

When $\underline{V} = V^*$, the admissibility constraints (23) are slack and future generations are taken into account only through the altruism of the first generation. This is the case studied by Atkeson and Lucas (1995) and Kocherlakota (2005). When $\underline{V} > V^*$, the admissibility constraints are binding at times.

Let $\beta^t \mu_t$ and $\beta^t \nu_t$ denote the multipliers on the resource constraints (21) and the admissibility constraints (23). At an interior solution, the first order necessary conditions for consumption and capital can be rearranged to give

$$\frac{1}{u'(c^v(\theta^t))} = \frac{1}{\beta F_K(K_{t+1}, N_{t+1})} \mathbb{E}_t \left[\frac{1}{u'(c^v(\theta^{t+1}))} \right] - \frac{\nu_{t+1}}{\mu_t}. \tag{24}$$

When $\nu_{t+1} = 0$ this optimality condition is known as the Inverse Euler equation. Consequently, we refer to equation (24) as the Modified Inverse Euler equation. It generalizes equation (7) to incorporate uncertainty regarding the descendants consumption. Later, we shall explore the implications of this condition for the dynamics of optimal allocations. First, we study its consequences for estate taxation.

In the next subsection we construct an explicit tax system that implements the allocation. Before doing that it is useful to study the intergenerational wedge $\chi^v(\theta^t)$ defined by

$$u'(c^v(\theta^t)) \equiv (1 - \chi^v(\theta^t)) \beta F_K(K_{t+1}, N_{t+1}) \mathbb{E}_t \left[u'(c^v(\theta^{t+1})) \right]. \tag{25}$$

In words, χ is the implicit wedge between the rate of return on capital and that riskless rate of return that would make the agent's standard consumption Euler equation hold. Then using equation (24) we obtain:

$$\chi^v(\theta^t) = \frac{1}{\beta F_K(K_{t+1}, N_{t+1})} \left(\mathbb{E}_t \left[\frac{u'(c^v(\theta^t))}{u'(c^v(\theta^{t+1}))} \right] - \frac{u'(c^v(\theta^t))}{\mathbb{E}_t[u'(c^v(\theta^{t+1}))]} \right) - u'(c^v(\theta^t)) \frac{\nu_{t+1}}{\mu_t}.$$

As long as there is uncertainty in next period's consumption, Jensen's inequality implies that the first term on the right hand side is positive, contributing towards a positive intergenerational wedge $\chi > 0$. This is precisely the positive distortion emphasized by Golosov et al. (2003). However, as long as the admissibility constraint binds, so that $\nu_{t+1} > 0$, the second term on the right hand side contributes towards a negative intergenerational wedge $\chi < 0$. In addition, because of the presence of $u'(c^v(\theta^t))$ this second term is increasing in current consumption $c^v(\theta^t)$, contributing towards the progressivity of the intergenerational wedge χ. It follows that in general we the sign of the intergenerational wedge is ambiguous.

In the two period economy, the first term was zero and the intergenerational wedge

was determined exclusively by the second term. In that case, the intergenerational wedge coincides with the explicit marginal tax rate from our implementation. In the infinite horizon economy, implementations are necessarily more complex and may also break the direct link between the intertemporal wedge, defined above, and an explicit marginal tax rate. Next, we study two such explicit tax implementations.

5.2 Two Tax Implementations

We now explore two tax implementations and focus on their implications for estate taxation.

Linear inheritance taxes. Our first implementation is along the lines of Kocherlakota (2005) and features linear taxes on inherited wealth. These tax rates generally depend on the entire history of reports, including the current shock's report. The dependence on the current report makes the net-of-tax return on capital risky. As shown by Kocherlakota, this feature is sufficient to discourage double deviations in reports and savings.

Take any incentive compatible, feasible allocation $\{c_t^v(\theta^t), n_t^v(\theta^t)\}$. In each period, conditional on the history of their dynasty's reports $\hat{\theta}^{t-1}$ and any inherited wealth, individuals report their current shock $\hat{\theta}_t$, produce, consume, pay taxes and bequeath wealth subject to the budget constraints

$$c_t(\theta^t) + b_t(\theta^t) \leq W_t n_t^v(\hat{\theta}^t) - T_t^v(\hat{\theta}^t) + (1 - \tau_t^v(\hat{\theta}^t))R_{t-1,t}b_{t-1}(\theta^{t-1}) \tag{26}$$

and initially $b_{-1} = K_0$. Individuals are subject to two forms of taxation: a labor income tax $T_t^v(\hat{\theta}^t)$, and a proportional tax on inherited wealth $R_{t-1,t}b_{t-1}$ at rate $\tau_t^v(\hat{\theta}^t)$.[10]

Given a tax policy $\{T_t^v(\theta^t), \tau_t^v(\theta^t)\}$, an equilibrium is a sequence of wages and interest rates $\{W_t, R_{t,t+1}\}$, an allocation for consumption, labor and bequests $\{c_t^v(\theta^t), n_t^v(\theta^t), b_t^v(\theta^t)\}$; and a reporting strategy $\{\sigma_t^v(\theta^t)\}$ such that: (i) $\{c_t^v, b_t^v, \sigma_t^v\}$ maximize dynastic utility subject to (26), taking wages and interest rates $\{W_t, R_{t-1,t}\}$ and tax policy $\{T_t, \tau_t\}$ as given; (ii) in each period t, aggregate capital K_t and labor N_t maximize profits $F(k, n) - R_{t-1,t}k - W_t n$ taking the wage and interest rate as given, or equivalently $W_t = F_N(K_t, N_t)$ and $R_{t-1,t} = F_K(K_t, N_t)$; (iii) markets clear: the resource constraints (21) are satisfied with equality. We seek a tax policy that implements efficient allocations as a competitive equilibrium with

---

[10]In this formulation, taxes are a function of the entire history of reports, and labor income n_t is mandated given this history. However, if the labor income histories $n^t \colon \Theta^t \to \mathbb{R}^t$ being implemented are invertible, then by the taxation principle we can rewrite T and τ as functions of this history of labor income and avoid having to mandate labor income. Under this arrangement, individuals do not make reports on their shocks, but instead simply choose a budget-feasible allocation of consumption and labor income, taking as given prices and the tax system. See Kocherlakota (2005).

truth-telling.

For any feasible, incentive-compatible allocation $\{c_t^v,\, n_t^v\}$ with strictly positive consumption we construct a tax policy that induces an equilibrium where all agents bequeath $b_t = K_t$. First, using the budget constraint with equality, let

$$T_t^v(\theta^t) = W_t n_t^v(\theta^t) + (1 - \tau_t^v(\theta^t))R_{t-1,t}K_t - c_t^v(\theta^t) - K_{t+1}.$$

Second, following Kocherlakota (2005), set the linear tax on inherited wealth to

$$\tau_t^v(\theta^t) = 1 - \frac{1}{\beta R_{t-1,t}}\frac{u'(c_{t-1}^v(\theta^{t-1}))}{u'(c_t^v(\theta^t))}. \tag{27}$$

These choices work because for any reporting strategy σ, the agent's consumption Euler equation

$$u'\big(c_t^v\big(\sigma^t(\theta^t)\big)\big) = \beta R_{t,t+1}\sum_{\theta_{t+1}} u'\big(c_{t+1}^v\big(\sigma^{t+1}(\theta^t,\theta_{t+1})\big)\big)\big(1 - \tau_{t+1}^v\big(\sigma^{t+1}(\theta^t,\theta_{t+1})\big)\big)\Pr(\theta_{t+1})$$

holds. Since the budget constraints hold with equality, this bequest choice is optimal regardless of the reporting strategy σ. The allocation is incentive compatible by hypothesis, so it follows that truth telling σ^* is optimal. Resource feasibility ensures that the markets clear.[11]

For efficient allocations, the assignment of consumption and labor at in any period depends on the history of reports in a way that can be summarized by the continuation utility $v_t(\theta^{t-1})$. Therefore, the estate tax $\tau^v(\theta^{t-1}, \theta_t)$ can be expressed as a function of $v_t(\theta^{t-1})$ and θ_t; abusing notation we denote this by $\tau_t(v_t, \theta_t)$. Similarly write $c_{t-1}(v_t)$ for $c_{t-1}^v(\theta^{t-1})$. The average estate tax rate $\bar\tau_t(v_t)$ is then defined by

$$\bar\tau_t(v_t) \equiv \sum_\theta \tau(v_t, \theta)\Pr(\theta).$$

Using the modified inverse Euler equation (24) we obtain

$$\bar\tau_t(v_t) = -\frac{\nu_t}{\mu_{t-1}}u'\big(c_{t-1}(v_t)\big). \tag{28}$$

Proposition 8. *Efficient interior allocations can be implemented by a combination of income and linear estate taxes. The optimal average estate taxes $\bar\tau_t(v_t)$ defined by (28) is negative*

---

[11] A version of Ricardian equivalence holds, so that the same allocation can be implemented with the same estate taxes, but adjusting the income taxes and bequests. In particular, it is possible to have agents with higher v_t leaving higher bequests. (This is actually the case in the next implementation.)

and increasing in promised continuation utility v_t.

Formula (28) is the exact analog of equation (8). Note that in the Atkeson-Lucas benchmark where the welfare of future generations is only taken into account through the altruism of the first generation, the average estate tax is equal to zero exactly as in Kocherlakota (2005). Both the negative sign and the progressivity of average estate taxes derive directly from the desire to insure future generations against the risk of being born to a poor family. The progressivity of the estate tax results from the resolution of the tradeoff between this desire to insure children and the provision of incentives for altruistic parents.

A Nonlinear Estate Tax Implementation. We now propose a two-stage implementation that is closer to the one used in the two period version of the model. In the first stage, before the descendant's productivity is realized, bequests are taxed according to a nonlinear schedule that is decreasing and convex. This estate tax reflects the features of our main results, both in terms of the progressivity and the sign of taxation. In the second stage, a linear wealth or inheritance tax is implemented exactly as in Kocherlakota (2005). In particular, tax rates are zero on average, but vary to deter double deviations in reports and bequests. If there were no uncertainty in the descendant's skill, then the second-stage linear wealth tax would be identically zero, just as in the two period version of the model.

Each period, individuals report their current shock $\hat{\theta}_t$, produce, consume, pay taxes and bequeath wealth subject to the budget constraints

$$c_t(\theta^t) + b_t(\theta^t) \leq W_t n_t^v(\hat{\theta}^t) - T_t^{y,v}(\hat{\theta}^t) - T_t^b(b_t) + (1 - \tau_t^{b,v}(\hat{\theta}^t))R_{t-1,t}b_{t-1}(\theta^{t-1}) \quad (29)$$

Individuals are subject to three forms of taxation: an income tax $T_t^{y,v}(\hat{\theta}^t)$, an estate tax $T_t^b(b_t)$ and a proportional tax on inherited wealth $R_{t-1,t}b_{t-1}$ with rate $\tau_t^{b,v}(\hat{\theta}^t)$. Given a tax policy $\{T^{n,v}, T^b, \tau^{b,v}\}$, a competitive equilibrium is defined exactly as before, but replacing the budget constraint (26) with (29). Once again, we will construct a tax policy that implements efficient allocations as a competitive equilibrium with truth-telling.

We have already argued that continuation utility v_t is a sufficient state variable for efficient allocations. The continuation utility v_t depends on the history of a dynasty's report θ^{t-1} and the initial welfare entitlement v, so we write $v_t = v_t(\theta^{t-1}, v)$ to emphasize this dependence. In our implementation, there will be a one to one mapping between bequests and continuation utility, so that we can keep track of the latter using the former.

First, select any sequence of strictly increasing functions $B_t(v_{t+1})$, normalized so that

$$\int \sum_{\theta^t} B_t(v_{t+1}(\theta^t, v)) \Pr(\theta^t) d\psi(v) = K_{t+1}.$$

Next, let the estate tax schedule $T_{t-1}^b(\cdot)$ for any $t = 1, 2, \ldots$ solve

$$T_{t-1}^{b\prime}(B_{t-1}(v_t)) = \frac{\bar{\tau}_t(v_t)}{1 - \bar{\tau}_t(v_t)}$$

with $T_{t-1}^b(K_t) = 0$. Since $B_{t-1}(v_t)$ is increasing in v_t and $\bar{\tau}_t(v_t)$ is negative and increasing in v_t, the estate tax schedule $T_t^b(\cdot)$ is decreasing and convex. Set the inheritance tax rate to

$$\tau_t^{b,v}(\theta^t) \equiv 1 - \frac{u'(c_{t-1}^v(\theta^{t-1}))}{\beta R_{t-1,t} u'(c_t^v(\theta^t))(1 - \bar{\tau}_t(v_t(\theta^{t-1}, v)))}.$$

Finally, the income tax schedule $T_t^y(\theta^t)$ is defined so that the budget constraint holds with equality at the proposed allocation and bequests are given by $b_{t-1}^v(\theta^{t-1}) = B_{t-1}(v_t(\theta^{t-1}, v))$.

Proposition 9. *Efficient interior allocations can be implemented by a combination of income T_t^y an a estate tax T_t^b and an inheritance tax $\tau_t^{b,v}$. The inheritance tax is linear and its average is equal to zero:*

$$\sum_{\theta_t} \tau_t^{b,v}(\theta^{t-1}, \theta_t) \Pr(\theta_t) = 0.$$

The estate tax schedule $T_t^b(\cdot)$ is negative and increasing in the size of the bequest.

The proof of the proposition is similar to that of the previous implementation. In particular, by construction, the agent's consumption Euler equation

$$u'(c_t^v(\sigma^t(\theta^t)))(1 + T_t^{b\prime}(B_t(v_{t+1})))$$
$$= \beta R_{t,t+1} \sum_{\theta_{t+1}} u'(c_{t+1}^v(\sigma^{t+1}(\theta^t, \theta_{t+1})))(1 - \tau_{t+1}^{b,v}(\sigma^{t+1}(\theta^t, \theta_{t+1}))) \Pr(\theta_{t+1})$$

holds for any reporting strategy σ. In addition, note that given any reporting strategy σ, the budget set is convex since the estate tax T_t^b is convex and the inheritance tax is linear. Thus, first order conditions are sufficient for optimality of consumption and bequest decisions, given σ. Hence, given a reporting strategy σ, the resulting consumption and labor allocation $\{c_t^v(\sigma^t(\theta^t)), n_t^v(\sigma^t(\theta^t))\}$ with bequests given by $\{B_t(v_{t+1}(\sigma^t(\theta^t), v))\}$ is optimal from the perspective of the agents. Since the original allocation is incentive compatible, it follows that truth-telling is optimal. The resource constraint together with the budget constraints then ensure that the asset market clears.

This implementation is appealing because it decouples a nonlinear estate tax schedule $T_t^b(\cdot)$, that parallels the analysis of the two period model, from the linear tax associated with the standard inverse Euler equation as studied in Kocherlakota (2005).

5.3   Discussion: Long Run Inequality and Estate Taxation

Starting from any initial distribution of welfare entitlements, the cross-sectional distribution of consumption, work effort and welfare evolve over time along the efficient allocation. It is well known that, in an economy with infinitely lived agents and private information, these distributions are not guaranteed to settle down to a steady state (Atkeson and Lucas, 1992). How are things different in an intergenerational context? What role may estate taxation have in ensuring convergence of these distributions to a steady state? We do not attempt to address these questions fully. Instead, this section works out a simple example with logarithmic utility that sheds some light on this issue.[12]

A steady state consists of a distribution of utility entitlements ψ^* and a welfare level $\underline{V}^*$ such that the solution to the planning problem (22) features, in each period, a cross-sectional distribution of continuation utilities v_t that is also distributed according to ψ^*. We also require the cross-sectional distribution of consumption and work effort and consumption to replicate itself over time. As a result, all aggregates are constant at a steady state. In particular, $K_t = K^*$, $N_t = N^*$ and $R_t = R^*$, etc.

In the rest of the section, we specialize to the logarithmic utility case, $u(c) = \log(c)$. This simplifies things because $1/u'(c) = c$, which is the expression that appears in the first-order optimality condition (24). Consider first the case where $\underline{V} = -\infty$. Suppose that there exists an invariant distribution ψ, and let R be the associated interest rate. The admissibility constraints are slack and $\nu_t = 0$ giving the standard Inverse Euler equation

$$c_t^v(\theta^t) = \frac{1}{\beta R^*}\mathbb{E}_t[c_{t+1}^v(\theta^{t+1})]. \tag{30}$$

Integrating over v and θ^t, it follows that $C_{t+1} = \beta R^* C_t$, which is consistent with a steady state only if $\beta R^* = 1$. However, equation (30) then implies that consumption is a positive martingale. By the Martingale Convergence Theorem, consumption must converge almost surely to a finite constant. Indeed, one can argue that $c_t \to 0$ and $v_t \to -\infty$.[13] We conclude that no steady state exists in this case, echoes the findings in Atkeson and Lucas (1992).

Now suppose that $\underline{V} > -\infty$. At a steady state, the admissibility constraints are binding and μ_t/ν_t is equal to a strictly positive constant. To be compatible with some constant average consumption $\bar{c}$, equation (24) requires $R^* < 1/\beta$ and can be rewritten as

$$\mathbb{E}_t[c_{t+1}^v] = \beta R^* c_t^v + (1 - \beta R^*)\bar{c},$$

---

[12]Phelan (2006) and Farhi and Werning (2007) explore this existence question in some depth for a models without capital.

[13] This follows because consumption c_t is a monotone function of v_{t+1}. However, if v_{t+1} converges to a finite value then the incentive constraints must be slack. This can be shown to contradict optimality.

Consumption is an autoregressive process, mean reverting towards average consumption $\bar{c}$ at rate $\beta R^* < 1$.

Just as in the two period case, the intergenerational transmission of welfare is imperfect. Indeed, the impact of the initial entitlement of dynasties dies out over generations and $\lim_{j\to\infty} \mathbb{E}_t c_{t+j} \to \bar{c}$. Indeed, one can show that a steady state may exist with bounded inequality. Moreover, at the steady state there is a strong from of social mobility in that, regardless of their ancestor's welfare position v_t, the probabilistic conditional distribution at t for v_{t+j} of distant descendants converges to ψ^* as $j \to \infty$.

6   Concluding Remarks

Our constrained efficient analysis delivers a strikingly simple result: the estate tax is progressive and negative. We have shown that this result is robust to a number of extensions. We now briefly discuss what we have omitted here.

Farhi, Kocherlakota and Werning (2005) explore some extensions, such as modeling life-cycle elements and allowing skills to be correlated across generations. The main result on progressive estate taxation holds. However, a number of issues are still unexplored. For example, the effects of endogenous and variable fertility, and of inter-vivo transfers all remain open questions for future research.

The focus in this paper was entirely normative. In an intergenerational context, questions of political economy and lack of commitment arise naturally. Within a capital taxation context, Farhi and Werning (2008) explores of a model similar to the one in this paper but with explicit political economy constraints. In that model, taxation remains progressive but the marginal tax rate may be positive.

Appendix

A   Proof of Proposition 3

Equation (10) implies that T^b is decreasing and convex, and that

$$T^{b\prime}(b) = \frac{1}{1 - \tau((c_1)^{-1}(Rb))} - 1. \tag{31}$$

Next define net income

$$I(\theta_0) \equiv c_0(\theta_0) + R^{-1} c_1(\theta_0) + T^b(c_1(\theta_0)/R)$$

We can express this in terms of output y by using the inverse of $n_0(\theta_0)$: $I^y(n) \equiv I(n_0^{-1}(n))$. Then we let $T^y(n_0) \equiv n_0 - I^y(n_0)$. Finally, let the consumption allocation as a function of net income I be: $(\hat{c}_0(I), \hat{c}_1(I)) \equiv (c_0(I^{-1}(I)), c_1(I^{-1}(I)))$.

We now show that the constructed tax functions, $T^y(y)$ and $T^b(b)$, implement the allocation. For any given net income I the consumer solves the subproblem:

$$V(I) \equiv \max\{u(c_0) + \beta u(c_1)\}$$

subject to $c_0 + R^{-1} c_1 + T^b(c_1/R) \leq I$. This problem is convex, the objective is concave and the constraint set is convex, since T^b is convex. It follows that the first-order condition

$$1 = \frac{\beta R}{1 + T^{b\prime}(c_1/R)} \frac{u'(c_1)}{u'(c_0)}$$

sufficient for optimality. Combining equation (6) and equation (31) it follows that these conditions for optimality are satisfied by $\hat{c}_0(I), \hat{c}_0(I)$ for all I. Hence $V(I) = u(\hat{c}_0(I)) + \beta u(\hat{c}_0(I))$.

Next, consider the worker's maximization over n_0 given by

$$\max_{y}\{V(I(n_0)) - h(n_0/\theta_0)\}.$$

We need to show that $n_0(\theta_0)$ solves this problem, which implies that the allocation is implemented since consumption would be given by $\hat{c}_0(I(n_0(\theta_0))) = c_0(\theta_0)$ and $\hat{c}_1(I(n_0(\theta_0))) = c_1(\theta_0)$. Now, from the previous paragraph and our definitions it follows that

$$n_0(\theta_0) \in \arg\max_{n_0}\{V(I(n_0)) - h(n_0/\theta_0)\}$$
$$\Leftrightarrow n_0(\theta_0) \in \arg\max_{n_0}\{u(\hat{c}_0(I(n_0))) + \beta u(\hat{c}_1(I(n_0))) - h(n_0/\theta_0)\}$$
$$\Leftrightarrow \theta_0 \in \arg\max_{\theta}\{u(c_0(\theta)) + \beta u(c_1(\theta)) - h(n_0(\theta)/\theta_0)\}$$

Thus, the first line follows from the last, which is guaranteed by the assumed incentive compatibility of the allocation, equation (4). Hence, $n_0(\theta_0)$ is optimal and it follows that $(c_0(\theta_0), c_1(\theta_0), n_0(\theta_0))$ is implemented by the constructed tax functions.

B Proof of Proposition 4

We can implement this allocation with an income with an income tax $T^y(n_0)$ and a borrowing constraint mandating that $c_1 \geq \underline{c_1}$. Let $I(\theta_0) \equiv c_0(\theta_0) + c_1(\theta_0)/R$. We can express this in terms of output by using the inverse of $n_0(\theta_0) : I^y(\hat{n}) = I(n_0^{-1}(\hat{n}))$. We then let $T^y(n_0) \equiv n_0 - I^y(n_0)$. Finally we define the allocation as a function of net income I: $(\hat{c}_0(I), \hat{c}_1(I)) = (c_0(I^{-1}(I)), c_1(I^{-1}(I)))$.

For a given net income I, agents maximize

$$V(I) = \max_{c_0, c_1} u(c_0) + \beta u(c_1)$$

subject to

$$c_0 + c_1/R \leq I$$

$$c_1 \geq \underline{c_1}.$$

This is a concave problem with solution $(\hat{c}_0(I), \hat{c}_1(I)) : V(I) = u(\hat{c}_0(I)) + \beta u(\hat{c}_1(I))$. The value function is concave and differentiable with respect to I with $\frac{dV(I)}{dI} = u'(\hat{c}_0(I))$.

The agent then chooses n_0 in order to maximize

$$W(\theta_0) = \max_{n_0} V(I^y(n_0)) - h(n_0/\theta_0)$$

Incentive compatibility of the optimal allocation then shows that the objective function is maximized for $n_0 = n_0(\theta_0)$. This completes the proof that the optimal allocation can be implemented by the combination of an income tax T^y and a borrowing constraint.

Note that net income $I(\theta_0)$ is increasing in θ_0. Hence $c_0(\theta_0)$ and $c_1(\theta_0)$ are increasing in θ_0. This in turn implies that the estate tax is zero as long as $c_1(\theta_0) \geq \underline{c_1}$, i.e. $\theta_0 > \underline{V}_0$. For $\theta_0 \leq \underline{V}_0$, the estate tax is negative and progressive.

C Proof or Proposition 5

We can separate the social planning problem into two steps: first, solve the optimal allocation in terms of the reduced allocation $\{c_0(\theta_0), e(\theta_0), y(\theta_0)\}$; second, solve $c_1(\theta_0)$ and $x(\theta_0)$ using the program (15). The reduced allocation $\{c_0(\theta_0), e(\theta_0), y(\theta_0)\}$ is the solution of the following planning program:

$$\max \int_0^\infty \left[u^0(c_0(\theta_0)) - h(y(\theta_0)/\theta_0) + \beta V_1(e(\theta_0)) \right] dF(\theta_0)$$

subject to the resource constraint

$$\int_0^\infty \left[c_0(\theta_0) + \frac{e(\theta_0)}{R} \right] dF(\theta_0) \leq \int_0^\infty y(\theta_0) dF(\theta_0),$$

the incentive compatibility constraints

$$u^0(c_0(\theta_0)) + \beta V_1(e(\theta_0)) - h\left(\frac{n_0(\theta_0)}{\theta_0}\right) \geq u^0(c_0(\theta_0')) + \beta V_1(e(\theta_0')) - h\left(\frac{n_0(\theta_0')}{\theta_0}\right) \quad \forall \theta_0, \theta_0'$$

and the promise keeping constraint

$$\int_0^\infty V_1(e(\theta_0))dF(\theta_0) \geq \underline{V_1}.$$

Note that since V_1 is increasing and concave, this program has the same properties as the simple social planning problem introduced in Section 2. In particular, c_0, e and y are increasing in θ_0. Therefore, richer parents invest more in the education of their kids and leave them higher financial bequests. This in turns implies that $x(\theta_0)$ and $c_1(\theta_0)$ are increasing in θ_0 at the optimal allocation.

This problem is the exact analog of our original social planning problem with $c_1(\theta_0)$ replaced by $e(\theta_0)$ and child utility $u(c_1(\theta_0))$ replaced by $V_1(e(\theta_0))$. Therefore we know that $c_0(\theta_0)$ and $e(\theta_0)$ are increasing in θ_0. We also know that $c_1(\theta_0) - H(x(\theta_0))$ and $x(\theta_0)$ are increasing in θ_0.

Use the generalized inverse of $x(\theta)$, where possible flat portions of $x(\theta)$ define discontinuous jumps, to define

$$T^{x'}(x) = \frac{1}{1 - \tau((x)^{-1}(x))} - 1$$

and normalize so that $T^x(0) = 0$.

Use the generalized inverse of $(c_1 - H(x))(\theta)$ to define

$$T^{b'}(b) = \frac{1}{1 - \tau\left((c_1 - H(x))^{-1}(Rb)\right)} - 1$$

and normalize so that $T^b(0) = 0$.

Note that by the monotonicity of $\tau(\theta)$, $x(\theta)$ and $(c_1 - H(x))(\theta)$, the functions T^b and T^x are convex. Next define net income

$$I(\theta_0) \equiv c_0(\theta_0) + R^{-1}c_1(\theta_0) + T^b((c_1 - H(x))(\theta_0)/R) + T^x(x(\theta_0)).$$

We can express this in terms of output y by using the inverse of $y_0(\theta_0)$: $I^y(y) \equiv I(y_0^{-1}(y))$. Then we let $T^y(y_0) \equiv y_0 - I^y(y_0)$. Finally, let the consumption and human capital allocation as a function of net income I be: $(\hat{c}_0(I), \hat{c}_1(I), \hat{x}(I)) \equiv (c_0(I^{-1}(I)), c_1(I^{-1}(I)), x(I^{-1}(I)))$.

We now show that the constructed tax functions implement the allocation. For any given net income I the consumer solves the subproblem:

$$V(I) \equiv \max\{u^0(c_0) + \beta u^1(c_1, H(x))\}$$

subject to $c_0 + R^{-1}c_1 + T^b((c_1 - H(x))/R) + x + T^x(x) \leq I$. This problem is convex, the objective is concave and the constraint set is convex, since T^b and T^x are convex. It follows

that the first-order conditions

$$1 = \frac{\beta R}{1 + T^{b\prime}(c_1 - H(x))} \frac{u_{c_1}^1(c_1, H(x))}{u_{c_0}^0(c_0)}$$

$$1 = \frac{\beta}{1 + T^{x\prime}(x)} \frac{u_H^1(c_1, H(x))}{u_{c_0}^0(c_0)}$$

sufficient for optimality. It follows from the construction of the tax functions T^b and T^x that these conditions for optimality are satisfied by $\hat{c}_0(I), \hat{c}_0(I), \hat{x}(I)$ for all I. Hence $V(I) = u(\hat{c}_0(I)) + \beta u(\hat{c}_0(I), \hat{x}(I))$.

Next, consider the worker's maximization over n_0 given by

$$\max_{n_0}\{V(I(n_0)) - h(n_0/\theta_0)\}.$$

We need to show that $n_0(\theta_0)$ solves this problem, which implies that the allocation is implemented since consumption would be given by $\hat{c}_0(I(n_0(\theta_0))) = c_0(\theta_0)$ and $\hat{c}_1(I(n_0(\theta_0))) = c_1(\theta_0)$. Now, from the previous paragraph and our definitions it follows that

$$n_0(\theta_0) \in \arg\max_{n_0}\{V(I(n_0)) - h(n_0/\theta_0)\}$$
$$\Leftrightarrow n_0(\theta_0) \in \arg\max_{n_0}\{u(\hat{c}_0(I(n_0))) + \beta u(\hat{c}_1(I(n_0))) - h(n_0/\theta_0)\}$$
$$\Leftrightarrow \theta_0 \in \arg\max_{\theta}\{u(c_0(\theta)) + \beta u(c_1(\theta)) - h(n_0(\theta)/\theta_0)\}$$

Thus, the first line follows from the last, which is guaranteed by the assumed incentive compatibility of the allocation. Hence, $n_0(\theta_0)$ is optimal and it follows that $(c_0(\theta_0), c_1(\theta_0), n_0(\theta_0))$ is implemented by the constructed tax functions.

References

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

Atkeson, Andrew and Robert E. Lucas Jr., "On Efficient Distribution with Private Information," *Review of Economic Studies*, 1992, *v59* (3), 427–453.

_ **and** _ , "Efficiency and Equality in a Simple Model of Unemployment Insurance," *Journal of Economic Theory*, 1995, *66*, 64–88.

Atkinson, A.B. and J.E. Stiglitz, "The Design of Tax Structure: Direct vs. Indirect Taxation," *Journal of Public Economics*, 1976, *6*, 55–75.

Becker, Gary S. and Robert J. Barro, "A Reformulation of the Economic Theory of Fertility," *The Quarterly Journal of Economics*, February 1988, *103* (1), 1–25.

Cremer, H. and P. Pestieau, "Non-linear taxation of bequests, equal sharing rules and the tradeoff between intra- and inter-family inequalities," *Journal of Public Economics*, 2001, *79*, 35–53.

Diamond, Peter A, "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 1998, *88* (1), 83–95.

Ebert, Udo, "A reexamination of the optimal nonlinear income tax," *Journal of Public Economics*, 1992, *49* (1), 47–73.

Farhi, Emmanuel and Iván Werning, "Inequality and Social Discounting," *Journal of Political Economy*, 2007, *115(3)*, 365–402.

_ **and** _ , "The Political Economy of Non-Linear Capital Taxation," 2008. mimeo.

_ , **Narayana Kocherlakota, and Iván Werning**, 2005. work in progress.

Golosov, Mikhail, Aleh Tsyvinski, and Iván Werning, "New Dynamic Public Finance: A User's Guide," *forthcoming in NBER Macroeconomics Annual 2006*, 2006.

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

Kaplow, Louis, "A Note on Subsidizing Gifts," *Journal of Public Economics*, 1995, *469* (58).

_ , "A Framewok for Assessing Estate and Gift Taxation," 2000. NBER Working Paper 7775.

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

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

Phelan, Christopher, "Opportunity and Social Mobility," *Review of Economic Studies*, 2006, *73* (2), 487–505.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, *68* (1), 205–29.

Seade, Jesus, "On the Sign of the Optimum Marginal Income Tax," *Review of Economic Studies*, 1982, *49* (4), 637–43.

Sleet, Christopher and Sevin Yeltekin, "Credible Social Insurance," 2005. mimeo.

Tuomala, Matti, *Optimal Income Taxation and Redistribution*, Oxford University Press, Clarendon Press, 1990.