Taxation of Transfers and Wealth

Wojciech Kopczuk
Columbia University

Draft in preparation for Handbook of Public Economics
Preliminary and unfortunately still quite incomplete, comments welcome

December 5, 2011

Abstract

The main messages of this chapter may be summarized as follows. Empirical evidence on bequest motivations and responses to estate taxation is spotty and much remains be done, but what we know points in the direction of (1) mixed motives (2) heterogeneity of preferences and (3) importance of retaining control over wealth. These patterns are important for analyzing taxation toward the top of the distribution. Theoretical work should further focus on understanding implications of inequality of inherited wealth: the topic that has been neglected in the past, even though it is closely related to — more carefully studied but arguably much less important in practice — externalities from giving. On the other hand, potential negative externalities from wealth accumulation and concentration are yet to be seriously addressed.

Contents

1 Introduction

The objective of this chapter is to provide an introduction to and review of economic literature devoted to taxation of transfers and wealth. As will become clear in what follows, the focus will be primarily on taxes imposed on intergenerational transfers. Those taxes take many different forms. Transfers that occur at death may take a form of estate taxation — they may be imposed based on the total amount of wealth left by the decedent. They may take the form of an inheritance tax in which case they are imposed on the donee's side, based on the amount of the transfer to that particular individual.[1] If taxes were imposed only at death, the simplest form of avoidance would involve transfers during lifetime and hence transfer taxation systems almost always include a tax on gifts as well. Taxes on estates are a form of tax on wealth — some countries (e.g., France and Norway) impose annual taxes of that kind. There are many dimensions of differences in transfer taxation across countries, states and over time. On the basic design level, the estate tax may treat preferentially transfers to the spouse or charity, inheritance tax may vary depending on the relationship to the donor, gifts may be taxed on annual or lifetime basis, they may be integrated (or not) with taxation at death, the details of interaction with capital gains taxation regime may vary. Details of the implementation of the tax may matter greatly — some examples are valuation rules, preferences for particular types of assets, treatment of transfers shortly before death, joint vs community property, treatment of charity, treatment of transfers that skip generations. The purpose of this chapter is not to discuss all of these issues, although in Section2 I will provide a short overview of history and international differences in transfer tax systems. Instead, my objective is to focus on the economics of transfer taxation and to provide an overview of related theoretical and empirical research. This research is of course largely motivated by the existing forms of taxation, in particular empirical work naturally relies on what can be observed in practice and some of it is directly motivated by important current policy questions. Most, though not all, of research on these topics took place in the United States and hence the "bias" toward evidence (and salient policy questions) from the U.S. is likely to be present. However, the focus of the chapter is on taxation of transfers in general, with the U.S. being just a (prominent) example. I will discuss taxation of wealth briefly, to the extent that it relates to taxation of transfers rather than being a form of tax on capital incomes that are discussed elsewhere in this Handbook.

This is not the first survey of literature on transfer and their taxation. Gale et al., eds (2001) and Boadway et al. (2010) provide useful background to the issues surrounding the design of transfer taxation. Cremer and Pestieau (2006) discuss some theoretical contributions to the literature on taxation of bequests. Laitner (1997), Laferrère and Wolff (2006) and Arrondel and Masson (2006) discuss theoretical and empirical work on intergenerational linkages. Davies and Shorrocks (2000) and Cagetti and De Nardi (2008) cover work on wealth distribution. Luckey (2008) provides an excellent overview of the history of estate tax legislations in the United States.

[1] Inheritance taxation may in principle be integrated with personal income taxation, see Batchelder (2009) for a discussion and a proposal for the reform of the U.S. transfer tax system along these lines.

Before dwelling into details, it is worth emphasizing the structure, major themes and conclusions of this review.

I will begin with an overview of how taxation of this kind works and varies in practice — across countries, across states in the United States, and over time.

In Section 3, I will discuss evidence on bequest motives and basic implications of different bequest motives for thinking about taxation. This is the primary building block for theoretical analysis of taxation of transfers and the literature that has not settled on a clear answer to the question about the nature of bequest motivations. I emphasize heterogeneity of two different kinds. First, the quest for *the* bequest motive is unlikely to be fruitful — saving plays dual role of protecting against lifetime risk and increases transfers to others. Different motivations for transfers are not mutually exclusive — the same person may be altruistic and yet interested in controlling wealth or engaged in strategic interactions with children. Second, I emphasize the evidence suggesting that preferences are heterogeneous and do not necessarily cut across predictable lines (such as having kids). The primary conclusion of this section is that theoretical work should either be agnostic about the nature of the bequest motive or explicitly account for heterogeneity. I finish this section with discussion of the role of transfers and their taxation in wealth accumulation and shaping the overall distribution of wealth.

In Section 4, I focus on the main theoretical framework for analyzing bequests taxation that builds on the Mirrlees and Atkinson-Stiglitz approach to taxation of commodities in the presence of nonlinear income taxation. This approach of course incorporates redistributive motives. The basic insight is that transfers can be modeled as a form of consumption, albeit one with two unusual but related features. First, transfers directly benefit someone else beyond the donor. Second, the presence of such a benefit may generate a form of externality from giving. An externality from giving is natural to consider in this context and provides a reason for subidizing rather than taxing bequests. I discuss the logic of corrective taxation of externalities in a context with individualized rather than atmospheric externality and conclude that externalities from giving are not important for thinking about taxation at the top of the distribution — arguably, where taxation of transfers is relevant in practice. Furthermore, I point out that implications of inequality in received inheritances are not yet fully understood and are likely to lead to arguments for positive taxation of bequests. I then discuss work on capital income taxation more generally and point out its relationship to transfer taxation.

In Section 5, I begin a review of the empirical evidence about the effects of taxation of transfers. The focus of that section is on "real" responses — changes in the volume and timing of actual transfers, effect on labor supply, capital gains realizations, charity and transfer and survival of businesses. I follow up in Section 6 with the discussion of responses that fall along the avoidance margin. That section is focused primarily on evidence that applies to people with significant net worth. The key message of this discussion is that transfer tax planning involves a trade-off between tax minimization and control over wealth.

Section 7 is devoted to a few other topics that do not naturally fall elsewhere. I first provide a discussion of potential negative externalities from wealth accumulation and concentration. Then, I discuss evidence related charity, mobility, state vs federal issues and political economy of this type taxation. The final section concludes.

The main messages of this chapter may be summarized as follows. Empirical evidence on bequest motivations and responses to estate taxation is spotty and much remains be done, but what we know points in the direction of (1) mixed motives (2) heterogeneity of preferences and (3) importance of retaining control over wealth. Incorporating these components of empirical evidence into theoretical analysis is crucial, especially so for thinking about taxation toward the top of the distribution. Theoretical work should further focus on understanding implications of inequality of inherited wealth: the topic that has been neglected in the past, even though it is closely related to — more carefully studied but arguably much less important in practice — externalities from giving. On the other hand, potential negative externalities from wealth accumulation and concentration are yet to be seriously addressed.

2 Overview of wealth and estate taxation historically, internationally and in the U.S. in particular

incomplete...

2.1 The role of transfer taxation

Revenue potential

Administrative convenience

Redistribution and equality of opportunities

Backstop for tax avoidance of other types of taxes

3 Bequest motives and taxation

3.1 Single generation

In order to systematize the discussion of theoretical arguments related to transfer taxation, it is instructive to start from a single individual utility maximization problem. Consider an individual maximizing utility $u(C, B)$ defined over consumption (C) and transfers to a beneficiary (B), subject to the budget constraint $C + pB = y$ where y is income, p is the relative price of transfers and the price of consumption is normalized to one. Individuals will naturally set $u_B/u_C = p$ and changes in p and y will give rise to price and substitution responses.

Let's denote the pre-tax relative price of bequests by R (absent taxation we have $p = R$; one natural interpretation is as $R = (1 + r)^{-1}$ where r is the rate of return). The base for the estate tax is $y - C$ and denoting the tax rate by t it yields the budget constraint of $(1 - t)C + RB = (1 - t)y$ or, equivalently, $C + \frac{R}{1-t}B = y$. The estate tax increases the relative price of bequests and stimulates negative substitution response and further reduction of bequests via income effect (unless bequests are an inferior good). Imposing instead a tax on the basis of the amount of gift to the beneficiary would instead correspond to the tax liability of $t^G B$, the budget constraint of $C + R(1 + t^G)B = y$ and identical predictions about the direction of the response.

This simple formulation is an example of a particular type of bequest motive — the "joy-of-giving" or "warm-glow" — and is the baseline reduced-form approach used in analyzing implications of taxation of bequests or charitable contributions when the focus is on the donor only (Andreoni, 1990).

The assumption in this simple formulation is that bequests are just as any other consumption good in that they deliver utility to the giver and correspondingly respond to price incentives. This is an assumption that may not hold in practice. The simplest alternative is to consider a situation in which a taxpayer does not care about the amount received by the recipient but is instead concerned with the amount of wealth W that she contributes. Using the simplest estate-tax formulation, $B = (1 - t)W/R$. The simplest wealth-in-utility formulation $u(C, W)$ is equivalent to the joy-of-giving motive, except for the implications of taxation: the budget constraint remains $C + W = y$ in the presence of taxation and changes in taxation have no implications for individual behavior.

This approach (also referred to as "capitalistic spirit," following Weber, 1958) has been advocated in the literature modeling the top end of wealth distribution as a suitable way of representing motives of high net worth individuals(Carroll, 2000; Reiter, 2004; Francis, 2009), the topic to which we will return below. A justification for considering wealth-in-utility is either as an intrinsic utility from accumulating wealth or as a proxy for unmodeled benefits of wealth holding — power that it allows to exert over others, relaxing of borrowing constraints, precautionary benefits or using wealth as a measure of relative status (eg. due to positional externalities as in Frank, 2008).

An alternative model of bequests that do not yield utility to the donor is "accidental" bequest approach. In the life-cycle framework with uncertain lifespan (Yaari, 1965), individuals save for future consumption but, except for the last possible period of life, may die with positive wealth holdings. Whether that occurs depends on actuarial fairness of the market for annuities: if annuities are fairly priced, all consumption should be effectively annuitized and bequests would not occur (instead, insurance company would gain *ex post* in case of early death). If the annuity market is imperfect, as the empirical evidence suggests (Friedman and Warshawsky, 1990; Mitchell et al., 1999), people die with positive wealth. Hence, bequests are unintended and stochastic. As with wealth-in-utility approach, taxation has no effect on the size of bequests.

One often-repeated statement about taxation of accidental bequests is that 100% tax is efficient because it elicits no response. While the latter part of this statement is true, the former requires

qualifications. The tax on accidental bequests in a representative individual context indeed has the benefit of reducing "waste" — bequests would otherwise not be available for consumption purposes. Naturally this argument does not survive considering a more realistic context when bequests instead flow to some other party and hence are not assumed to be wasted. In that case, the tax on accidental bequests becomes simply equivalent to lump-sum taxation on the beneficiary. More subtly, accidental bequests reflect the presence of an underlying imperfections in the market for annuities. A taxpayer would clearly be better off by selling the right to a bequest conditional on dying at some time t (that has no value to him) and using proceeds for consumption at any other period. While confiscating a bequest of this kind yields no harm, it also does not directly address the underlying market failure. The first-best policy would instead allow for complete consumption smoothing via annuities and imply no accidental bequest.

Kopczuk (2003b) shows that the estate tax itself may play an annuity role: the insight is that given interest rate r and sequence of effective tax liabilities conditional on dying in period i of T_i, surviving from period i to period $i+1$ implies savings in lifetime tax liability of $T_i - \frac{T_{i+1}}{1+r}$. The presence of this implicit annuity increases the value of the tax on accidental bequests — by using confiscatory tax on bequests, one reduces tax payments relative to the alternative of unconditional lifetime taxation with the same present value — but it becomes of more interest when individuals have additionally an explicit bequest motive where it can be shown that (1) a small estate tax is welfare improving because of its annuity role and (2) under strong enough bequest motive, sufficiently flexible estate tax can implement the first-best solution.[2]

3.2 Intergenerational linkages

The discussion so far abstracted from the recipients of transfers. From the point of view of understanding bequest behavior the recipients may matter because the donor may respond to their characteristics or behavior. Furthermore, transfers — actual or expected — may also change the behavior of a recipient. For normative analysis, understanding implications of transfers for welfare of the recipient is important.

The most influential way of modeling intergenerational linkages is by introducing altruistic preferences à la Barro (1974). It is assumed that prior generation cares directly about welfare of the following generation(s). With just two generations (parents and children) to begin with, preferences of the parents can be expressed as $u^P(C^P, C^K) = v^P(C^P) + \rho u^K(C^K)$ where C^P is a vector of consumption goods of the parent, C^K is a vector of consumption goods of the child, v^P is utility of the parent from own consumption and u^K is the utility of a child from own consumption. The parent is assumed to care about welfare of a child but discount it at some rate ρ (presumably with $\rho < 1$). Of course, this is a workhorse model used in hundreds of papers with many variants and extensions that are beyond the scope of this chapter (Laitner, 1997, provides a good survey

[2]Even more generally, one can think of the estate tax as serving insurance role against other types of risks — such as investment risk — that would affect the value of estate at death.

of theoretical aspects of altruistic preferences). In its simplest variant, one abstracts from overlap between generations (C^P occurs now, C^K in the future) and considers maximization subject to the common resource constraint

$$y^P + Ry^K = C^P + RC^K \tag{1}$$

where y^P is income of parents and y^K is income of children. In this formulation, bequests are equal to the unconsumed resources of the parent $y^P - C^P$. [3] The standard result is that re-allocating resources in a lump-sum fashion between period P and period C has no effect on the budget constraint (1) — the Ricardian equivalence result, with bequests adjusting to offset. This implication has been tested in the context of bequests (Altonji et al., 1992; Wilhelm, 1996; Altonji et al., 1997; Laitner and Ohlsson, 2001) and soundly rejected. Another way of describing the implication is by noting that it calls for smoothing of marginal utility profile $v_C^P = \rho R u_C^K$. With multiple potential beneficiaries (e.g., multiple children), this condition should hold for *any* beneficiary — a conclusion that is not consistent with the pattern of equal bequest splitting documented in the literature (Menchik, 1980; Menchik and David, 1983; Light and McGarry, 2004)[4]

To understand implications for bequests, it is useful to explicitly consider a single period of life so that C^P and C^K are scalars and the parent's optimization problem is $\max_B v^P(C^P) + \rho u^K(y^K + B)$ subject to the constraint $C^P + pB = y$, where $p = \frac{R}{1-t}$ is the after-tax cost of a dollar transfer to the beneficiary. This formulation makes it clear that when the focus is on donors' behavior only, there is a close connection between this model and the warm-glow one: parents care about their own consumption and bequests, except that the marginal value of bequests depends on the income of a child. Abel and Warshawsky (1988) build on this argument to show how intensity of altruism relates to the strength of the joy-of-giving bequest motive in a model with infinite horizon (although the connection they establish is not invariant to the changes in taxation).

An alternative approach to bequests treats them as part of a transaction between parents and children with bequests compensating children for services that they provide to their parents such as direct help, attention, access to grandchildren etc. (Bernheim et al., 1985; Cox, 1987; Perozek, 1998).[5] As with altruism, evidence in support of the exchange motive is mixed, see Arrondel and Masson (2006) and Laferrère and Wolff (2006) for recent surveys.

The conclusion that arises in the most recent work on bequest motives is that searching for *the* bequest motive is unlikely to be successful. This is for two reasons. First, different motives are not exclusive — in the presence of uncertainty, the precautionary/accidental and intentional motives naturally co-exist (Dynan et al., 2002, 2004); a person may also have a mix of altruistic and

[3] Perhaps augmented by investment returns R^{-1}, depending on the assumed convention about whether transfer occurs at the end of period when P generation is alive or the beginning of period when the C generation is active.

[4] Bernheim and Severinov (2003) propose an explanation for equal splitting that is based on the assumption that children care about altruistic parent's affection and infer it based in part on observable bequests. Dunn and Phillips (1997) and McGarry (1999) provides evidence that (presumably harder to observe) inter vivos gifts are compensatory while bequests are split equally.

[5] Exchange motives give rise to "strategic" interactions between parents and children, but strategic interactions naturally arise in the multi-period altruistic context as well (Bruce and Waldman, 1991; Coate, 1995).

exchange motivations, or simultaneously put weight on both wealth and bequests for example. Second, different individuals may have different motives.

For example, Light and McGarry (2004) document heterogeneity in preference for leaving bequests based on verbatim answers given to a question about reasons for planning not to split bequests equally in National Longitudinal Surveys of Young Women and Mature Women. Laitner and Juster (1996), based on a survey of TIAA-Cref participants, show that the intention to leave a bequest is not universal and, in fact, does not seem to be even remotely close to being well explained by having children — 45% people with children consider bequests important relative 21% among childless ones.

Hurd (1987) shows that wealth profiles of people with and without children are similar. Using cross-sectional AHEAD/HRS data and a structural approach to modeling wealth profiles, Hurd (1989) allows for accidental and intended bequests and tests for the presence of a bequest motive by assuming that people with kids have one and those without them do not and rejects that it is present. Kopczuk and Lupton (2007) build on his framework but they exploit longitudinal information and allow kids to be one of the potential indicators for the presence of the motive and conclude that bequest motive is present but not deterministically related to having children. Ameriks et al. (2011) model saving for long-term care and bequest motives; they use very similar switching regression strategy as Kopczuk and Lupton (2007) to conclude that both public long-term care aversion and bequest motives are important. Both of these papers find evidence supporting heterogeneity of the *presence* of the motive and they both find that bequests are a mix of accidental and intentional ones. The intentional bequests are effectively modeled as luxury good but they become important far from the top of the wealth distribution.

Kopczuk (2007) shows that wealth accumulation for the very wealthy continues until the onset of a terminal illness but that tax avoidance is responsive to that event, supporting the notion that people value both lifetime wealth and bequests. In the survey of literature on bequest motives, Arrondel and Masson (2006) advocate a mix of altruistic and strategic motives. The literature on the determinants of savings and wealth distribution grappled with this question as well. A strand of this literature assumes away the presence of a bequest motive (Hubbard et al., 1994, 1995; Scholz et al., 2006) but it has problems explaining the very top of the wealth distribution. Adding an explicit bequest motive helps (De Nardi, 2004; Cagetti and De Nardi, 2008; Reiter, 2004) but the standard in the literature approach of assuming altruism is not able to generate sufficient skewness within the top 1% or so. Hence, researchers are often resorting to reduced form specifications of the wealth-in-utility or warm-glow kind.

3.3 Normative issues

The lack of consensus about the nature of a bequest motive makes reaching definitive theoretical conclusions about the impact of taxation difficult and it makes normative analysis hard because it

(often) requires taking a stand on the unknown nature of the bequest motive.[6]

Before engaging in a normative analysis, it is worthwhile to pause to understand what role taxation might play. To do so, consider a parent with utility of $u^P(C, W, B, X)$ where W is wealth, B is effective bequest, X are other variables describing interaction with the child (attention, services, non-monetary transfers); and a child with a reduced-form utility of $u^K(B, X)$. Suppose that the social planner is interested in maximizing the weighted sum of utilities

$$u^P + \beta u^K \tag{2}$$

subject to the relevant resource constraint. The nature of the transfer motive, details of the household bargaining problems, strategic interactions between parents and children influence the value of the objective. The outcome need not be efficient in general — for example, addressing the Samaritan's dilemma problem may require commitment on the part of the parents. The outcome need not also be fair — in the exchange context, the market power may be on the side of the parent or on the side of the child and need not reflect social preferences. Hence, there may be a conceivable role justifying an intervention to address the potential inefficiency or redistribute resources within family. While not dismissing the relevance of such concerns, tax treatment of bequests or gifts is a blunt instrument for addressing them. In what follows, I will abstract from the issues that would call for the government intervention into family problem unless they explicitly relate to bequests. In particular, I will assume that the government respects the outcome of the family problem as efficient, unless explicitly stated otherwise.

The main reasons for a departure of government objective from respecting the maximization of family objective function considered in the literature has to do with the potential presence of externalities from giving. A dollar of bequests provides utility to both parents and children. From the social point of view, the benefit of bequest is given by $u^P_B + \rho u^K_B$ but when maximizing her own utility the parent is only taking into account her own marginal benefit u^P_B and ignores the ρu^K_B component giving rise to an interpersonal externality. This externality is there regardless of the bequest motive *if* one accounts for welfare of a child beyond its effect on parent's utility. In many cases, this is a natural approach. For example, when the bequest motive has the warm-glow structure, the parent does not care about the utility of a child and instead is assumed to derive the utility from the value of a gift itself. Naturally then one is inclined to consider bequests to as being under-provided: the benefit that they deliver to the donee is not taken into account by the donor.

Selecting the normative criterion under altruistic model is more controversial because parent's preferences already explicitly depend on child's utility. Writing, as before, the parental utility as $u^P(C^P, C^K) = v^P(C^P) + \rho u^K(C^K)$, the social planner's objective that accounts for both utility of the parent and the child becomes $u^P + \beta u^K = v^P + (\rho + \beta)u^K$. In the special case when $\beta = 0$,

[6] Kopczuk (2001) and Cremer and Pestieau (2006) analyze bequest taxation using models that have different types of bequest motives as special cases.

the social planner simply maximizes parental welfare. This is of course the standard approach of focusing on dynastic welfare. If instead $\beta > 0$, it corresponds to social planner putting an extra weight on welfare of children beyond what parents do.[7]

The key thing to observe is that normative analysis requires taking a stand on the presence of such an externality. In standard cases such as altruistic preferences or joy-of-giving bequest motive, the externality is caused by bequests and it is positive. As the result, its presence calls for corrective policies that would address the external effect. The Pigouvian subsidy to bequests that corrects the parental incentive to internalize the externality is the optimal policy in the first-best. In the second-best Ramsey commodity tax problems, it calls for adjustments to the tax structure but, as Sandmo (1975) shows, these adjustments should be targeted to the source of the externality — i.e., lead to a subsidy to bequests. Kopczuk (2003a) shows that the "targeting principle" logic applies to general tax problems with atmospheric externality (i.e., an externality that is generated by aggregate consumption) as long as the source of the externality can be taxed directly.

Considering an externality from giving is a normative assumption. Showing that it gives rise to subsidies to bequests is a straightforward consequence to keep in mind when evaluating normative tax exercises even when analytics of obtaining that conclusion is complicated. Having said that, the externality of that kind does come up naturally. Diamond (2006) provides a normative discussion of arguments for including the warm-glow motive in the social welfare function. In other words, the question he poses is not whether the benefit to the donee should be explicitly counted (as arises when one considers the altruistic case), but rather whether the benefit to the donor from the act of giving should be accounted for. The main argument for accounting for the warm glow is obviously that warm-glow preferences are presumed to determine behavior and hence should be accounted for by the social planner. The main counter-arguments have to do with reduced-form of such preferences that may miss other benefits or costs, and with consequences of accounting for the utility from the process (giving) rather than consumption of resources. For example, under a naive interpretation, two parties exchanging gifts of the same value would increase the utility of both parties with no change in ultimate consumption. Hence, a policy subsidizing such gifts might increase welfare. Alternatively, a policy that would substitute one-for-one bequests for direct government transfers to donees would reduce welfare by depriving donors of the warm glow.

Phelan (2006) and Farhi and Werning (2007) explicitly analyze placing an extra weight on future generations in an altruistic context. Considering altruism has an advantage over reduced form motives for bequests in that it avoids placing a value on the act of giving and instead focuses squarely on the final allocation of resources. The disadvantage is weak empirical support for these types of preferences especially when considering the very top of the distribution that estate taxation in practice is about.

Assuming that a form of an externality from giving is to be considered, there are a few additional

[7] One could also imagine $\beta < 0$ — social planner discounting welfare of children more than parents do — the case that has been considered in political economy models.

things to note.

First, as mentioned before, targeting prescription for dealing with externalities relies on the presence of an instrument that can target the source of an externality directly. The standard case is an "atmospheric externality" when the identity of the person taking action generating the externality is irrelevant. More generally, the social planner should target directly any source of the externality in proportion to the damage. Since with an atmospheric externality every source has the same impact on the social welfare, the tax does not to be differentiated. This is not the case with bequest externality: the externality is interpersonal and, with sufficient heterogeneity, marginal impact of bequests by different individuals will be different. This would then call for differentiating subsidies to bequests and whether it is feasible depends on available tax instruments.

Second, and relatedly, the importance of accounting for the giving externality may vary with the context considered. For example, one may place a high value on welfare of low-income children but correcting for inadequate gifts by wealthy parents to their wealthy children does not sound as an important policy objective. We will return to this issue when considering estate taxation in a redistributive context.

To conclude, a giving externality is often a component of the normative analysis of estate taxation. Its presence tilts the policy in the direction of subsidies to giving. The assumed nature of the bequest motive influences the nature of this externality but a normative choice can often be explicitly made: for example, one can ignore the warm-glow or make a decision about the extra weight, if any, to be put on welfare of future generations. The best theoretical practice is to be explicit about the presence of such an externality and its precise consequences; in particular about the consequences of varying its strength or complete elimination.

3.4 The role of inheritances and taxation in shaping wealth distribution and intergenerational mobility/transmission of inequality.

Kotlikoff and Summers (1981)Modigliani (1988)Davies and Shorrocks (2000)Stiglitz (1978)(Cagetti and De Nardi, 2009, 2008)Laitner (1992, 1988)Dynan et al. (2002, 2004); Hubbard et al. (1994, 1995)Piketty et al. (2003)Kopczuk and Saez (2004) Scholz (2003)Scholz et al. (2006)Gale and Scholz (1994)Piketty et al. (2003)Kopczuk and Saez (2004)Roine and Waldenström (2009)Edlund and Kopczuk (2009)Piketty (2011)

4 Redistribution

In the previous section, I abstracted from redistributive motives. Taxation of estates in practice is about redistribution. For example, according to the Piketty and Saez (2007) assessment of the overall progressivity of the U.S. tax code in 1970, the estate, gift and wealth taxes contributed 23.4 percentage points to the overall 74.6% average effective tax rate applying to the top 0.01% of the

distribution, while by 2004 contribution of these taxes fell to just 2.5 points out of the 34.7% total — according to that study, the decline in this type of taxation accounted for half of the change in effective tax burden of the wealthiest over that period. Clearly, analyzing the taxes that apply predominantly to those with high net worth and has the potential to make such a difference in the overall progressivity cannot ignore redistributional issues.

Building on the standard optimal income tax model of Mirrlees (1971), Kaplow (2001) provides the starting point for thinking about redistribution and estate taxation. Focusing on the donors, consider a society consisting of individuals maximizing utility given by $u(C, L, B)$ where C is consumption, L is labor supply and B is bequest. As in optimal income tax literature, assume that every individual is characterized by skill level w that remains private information. The planner can observe income wL and bequests B and can impose tax liability based on that information so that individuals are maximizing utility subject to the budget constraint $C + B \leq wL - T(wL, B)$. Denoting by $w(\cdot)$ a concave welfare function, one is interested in finding the tax schedule $T(\cdot)$ that maximizes welfare $\int w(u(C, L, B))$ subject to the revenue and incentive compatibility constraints.

This basic framework assumes that bequests are just like any other good. It also assumes that skills are the only source of heterogeneity. It implies that bequests are deterministic function of labor income. It also assumes away heterogeneity in tastes that led McCaffery (1994) to argue against estate taxation on the basis of its horizontally inequitable treatment of savers vs spenders.

This framework is of course a special case of Atkinson and Stiglitz (1976) and leads to the classic result that tax on commodities — bequests in this case — is redundant when utility has weakly separable structure $u(v(C, B), L)$.[8] The intuition for this result can be seen by appealing to the informational content of potential tax base. The unobservable piece of information is w. Under weak separability, one can consider a subproblem of maximizing utility from regular consumption and bequests given labor income wL: $\max_{C,B} v(C, B)$ subject to $C + B \leq wL - T(wL, B)$ that yields a solution $(C(wL), B(wL))$: consumption and bequests are a function of labor income and do not depend on wage rate directly. In other words, individuals with different wages will select the same level of consumption and bequests *if* their incomes are the same. As the result, distorting price of bequests does not provide any additional information about wages beyond that already contained in income and hence is redundant.

There are many limitations of this exercise of course, but it illustrates one of the components of the analysis of the estate tax: its interaction with lifetime redistribution. Viewed in this way, analysis of bequest taxation is analogous to analyzing desirability of capital taxation. That literature focused on understanding implications of preference heterogeneity (Saez, 2002a; Diamond and Spinnewijn, 2010; Golosov et al., 2010) and shows that uniform tax on capital income may be desirable even

[8] Laroque (2005) and Kaplow (2006) show that commodity taxation is redundant even when income tax is not optimally selected.

under the weak separability assumption if higher ability individuals have a lower discount rate,[9,10] while the tax on savings of just high ability individuals may be optimal under weaker assumption. Treating bequests as a form of saving and allowing for heterogeneity of preference for bequests would be a natural extension of this framework.

The natural next step is to explicitly consider multiple generations. Let us consider first the case when generations are altruistically linked, since this is the most common specification in the literature. The simplest approach builds on the Atkinson-Stiglitz framework and continues to abstract from decisions of children, instead assuming that there are two generations with parents choosing labor supply and consumption, and children selecting consumption level given the transfer. The dynastic utility is given by

$$u_P(C^P, L) + \rho u_K(C^K) \tag{3}$$

I will use this formulation in what follows. Note that bequests are present here as $B = C^K$, because bequests are the only source of income for the young generation. Denoting the pre-tax estate as $E = wL - C^P$, the general budget constraint of the parents (and the dynasty) can be written without loss of generality as

$$C^P + B/R = wL - T(wL, E) \qquad \Longleftrightarrow \qquad C^P + C^K/R = wL - T(wL, wL - C^P) \tag{4}$$

where $T(\cdot, \cdot)$ is a general tax function that depends on the two observable pieces of information wL and E. In particular, observing C^P and C^K is redundant since they can be recovered based on the values of wL and E.

When welfare is based on aggregating dynastic utilities, this model is again an example of the standard Atkinson-Stiglitz framework with two consumption goods C^P and C^K. Further assuming additive separability $u_P(C^P, L) = u_p(C^P) - v(L)$ to guarantee the weak separability assumption (as does the paper of Farhi and Werning, 2010, discussed in more details below), the model implies no tax distortions beyond income tax, the point previously made by Kaplow (2001).

One might argue that the social planner should account for both utility of parents and children. One way to introduce it is by putting an extra weight $\nu \geq 0$ on a child's utility when evaluating welfare of a given dynasty

$$u_P(C^P, L) + (\rho + \nu) u_K(C^K) \tag{5}$$

[9] Banks and Diamond (2010) discuss empirical evidence consistent with this pattern, while Gordon and Kopczuk (2010) test directly for a weaker condition necessary for deviation from the Atkinson-Stiglitz result — ability of capital income to predict wages conditional on labor income — and find support for it.

[10] Cremer and Pestieau (2001) show in the appendix desirability of bequest taxation in a two-type model that violates the Atkinson-Stiglitz assumption. See also Cremer et al. (2003) who consider the context where inheritance is not observable and show desirability of using an additional instrument (a tax on capital income) that is informative about the unobserved inheritances.

This approach may be interpreted as social planner disagreeing with dynastic preferences. Farhi and Werning (2010) consider a planner that puts an extra weight on the future generation but they take a slightly different tack. They set up their problem in terms of maximization of the welfare of the first period generation

$$\int u_p(C^P) - v(L) + \rho u_K(C^K) \tag{6}$$

subject to the lower bound on welfare of the second generation

$$\int u_K(C^K) \geq \underline{\mathrm{V}} \tag{7}$$

with $\underline{\mathrm{V}}$ indexing the problem. When $\underline{\mathrm{V}}$ is low enough, the constraint is not binding and the standard Atkinson-Stiglitz no-estate-tax result applies. When the constraint is binding, the problem is equivalent to maximizing $\int u_p(C^P) - v(L) + \rho u_K(C^K) + \nu u_K(C^K)$ as in equation (5) where (with some abuse of the notation) ν is the optimum value of the multiplier on the welfare constraint for the second generation.

Whether the problem is set up by appealing to an externality from giving on the individual level or whether it introduces it on the generational level, makes no difference in the utilitarian case. The two approaches depart from each other when applying a non-linear welfare function — in one case, the welfare should be evaluated as

$$\int W(u_P(C^P, L) + (\rho + \nu)u_K(C^K)) \tag{8}$$

while in the other the welfare function is applied to parent's and child's utility separately, possibly using different welfare criteria: $W_1(u_p(C^P) - v(L) + \rho u_K(C^K))$. As the result, given the multiplier ν on the constraint (7), the planner's objective is

$$\int W_1(u_p(C^P) - v(L) + \rho u_K(C^K)) + \nu W_2(u_K(C^K)) \tag{9}$$

While objective functions (8) and (9) represent slightly different problems, we will see shortly that the difference in the welfare criterion has no implications for qualitative solutions.

The objective function of the social planner does not coincide here with that of the parent generation. Instead, it puts an extra weight on the utility of the next generation. From the point of view of evaluating social welfare, there is a *positive* externality associated with children consumption. Since in this model bequests play the sole role of determining consumption of children, there is then a positive externality associated with bequests.

4.1 Estate taxation with externalities from giving — intuition

To gain the intuition for implications of externalities from giving note the following.

First, as has been known since (Pigou, 1920), the presence of externalities in the first best world calls for internalizing the externality via the Pigouvian tax. Writing the dynastic budget constraint as $C^P + C^K/R = wL - T(wL, wL - C^P)$, individuals set $\rho R u'_K = (1 - T_2)u'_P$. The social optimum needs to satisfy $(\frac{W'_2}{W'_1}\nu + \rho)Ru'_K = u'_P$ and setting the marginal bequest tax rate to the value of $T_2 \equiv \frac{W'_2}{W'_1}\frac{u'_K}{u'_P}$ (with the right hand side evaluated at the social optimum) brings incentives in line. With sufficiently flexible instruments (ability to pursue individualized lump-sum taxation and to set the marginal tax rates on bequests for each individual at the corresponding Pigouvian level) to address the underlying heterogeneity, it allows for implementing the first best allocation.

Second, the prescription for dealing with *atmospheric* externalities i.e. externalities stemming from aggregate consumption of some dirty good $\int D$ when first-best taxation is not available but a tax on D is possible is only a slight modification of the Pigouvian taxation. For simplicity, suppose that the effect of externality on social welfare is additive and given by $g(\int D)$. The logic of the targeting principle (Sandmo, 1975; Cremer et al., 1998; Kopczuk, 2003a) is straightforward. The problem with an externality is equivalent to the one without one but with the price of a dirty good adjusted (via the marginal tax rate tax τ) to internalize the otherwise ignored social cost of increasing $\int D$, and the revenue requirement modified by the amount collected by that tax ($\tau \int D^*$) at the allocation one wishes to implement. As the result, the presence of an externality modifies the qualitative structure of the solution only by the tax on the dirty good.[11]

This result calls for a linear tax at the rate that internalizes the externality. It is easy to see that with an atmospheric externality the social damage due to anyone's consumption of D is the same and given by $g'(\int D)$ so that the rate is indeed expected to be constant. What is that rate? The social planner weighs the resource cost of $\int D$ against any other uses and the shadow price reflects the multiplier on the resource constraint μ. As the result, the corrective rate can be shown to be equal to $\tau = g'/\mu$. The multiplier μ reflects the cost of public funds and its value need not be equal to u_p for any particular individual, so that the correction departs from person-by-person Pigouvian correction of externality and, in fact, it will usually depart on average because μ also reflects the distortionary cost of taxation.[12]

Third, it is not important that the externality is aggregate, rather what is important is that there is an instrument that can target it directly. In particular, if the dirty good is consumed by a subset

[11] See Kopczuk (2003a) for the precise statement and the proof.

[12] The second-best Pigouvian rate can also be written as $\tau = \frac{1}{\text{MCF}}\frac{g'}{\lambda}$ where λ is some weighted average of individual utilities and MCF $= \frac{\mu}{\lambda}$ is the marginal cost of public funds. Writing the optimal tax schedule from the standard optimal income tax model as $T(y) - G$ (with $T(0) = 0$ as the normalization), the perturbation argument with respect to a small change in the demogrant component dG (an increase in the lump-sum transfer for everyone) implies $\int \frac{\partial u}{\partial G} = \mu(1 - \int T' \frac{\partial y}{\partial G})$ and defining $\lambda = \int \frac{\partial u}{\partial G}$ yields MCF$=\frac{1}{1-\int T' \frac{\partial y}{\partial G}}$. Interestingly, as shown by Sandmo (1998)m when $T' > 0$ and income is a normal good this means that MCF < 1 — raising funds can be accomplished more cheaply than using a lump-sum tax. This is because lump-sum taxation is a potential instrument here, but it is revealed to have an interior solution at the optimum due to redistributive considerations.

of individuals $\tilde{D}$, so that it's given by $g(\int_{\tilde{D}} D)$, the optimal correction remains g'/μ. This applies even when there is a single individual consuming the good: the correction of an externality weighs on one hand its social cost and on the other hand revenue constraint implications. Furthermore, multiple externalities need not be a problem if they can be targeted independently.[13]

Coming back to the externalities from transfers, the complication is that there is not a single atmospheric externality here and instead one can think of the problem as involving a *continuum* of externalities given by $\nu W_1(u_K(C^K))$ for each dynasty. The straightforward application of the targeting principle would then call for a *continuum* of taxes targeting each of these externalities separately at the rate of $-R\frac{\nu W_1' u_K'}{\mu}$ (with the minus sign, because it is a positive rather than a negative externality and with R reflecting the price of C^K relative to the numeraire C^P). If it is possible to implement such a scheme that would force each individual to internalize the giving externalities that she causes, and if the externality does not interact with other considerations (most importantly, with incentive constraints), the optimal prescription follows the principle of targeting: forcing individuals to internalize the externality turns the problem into the standard one without an externality present. In particular, adding the giving externality on top of the Atkinson-Stiglitz setup should yield a tax targeting its source (if feasible to implement) with no qualitative modifications to the optimal tax schedule implications otherwise. In particular, under weak separability assumption, the sole role of distortions to bequest decisions would then stem from internalization of the externality.

4.2 Estate taxation with giving externalities — results

The intuition described in the previous section applies directly to the analysis of Farhi and Werning (2010) who allow for imposing an extra weight on the welfare of future generations and embed the analysis in the optimal income/consumption tax problem. Their central result is indeed that the optimal "implicit" marginal estate tax rate is given by $t^E = -R\frac{\nu W_1' u_K'}{\mu}$ or, reinterpreting, it is equal to the optimal estate tax rate when the externality is not present (trivially, equal to zero because of the weak separability assumption) plus the Pigouvian correction. Under their assumptions, the marginal estate tax rate t^E is only a function of bequest (or child's consumption):

$$t^E(B) = -\frac{R\nu}{\mu} W_1'(u_K(B)) u_K'(B) \tag{10}$$

(obviously, R, μ and ν are constant). They show that the size of bequests (B) and the estate ($wL - C^P$) are increasing function of wages, so that this desired marginal incentive may be implemented using either a tax on estates or a tax on inheritances that is separable from the income tax. This can be seen by integrating $t^E(B)$ over B to obtain the tax liability that a person who at the optimum leaves the bequest of B should face:

[13] See for example Green and Sheshinsky (1976) and Micheletto (2008) for explorations of corrective taxation when externalities are not uniform and cannot be targeted using independent instruments.

$$\tilde{T}(w) = \tilde{c} - \frac{R\nu}{\mu} \int_0^{B(w)} W_1'(u_K(x))u_K'(x)\,dx = c - \frac{R\nu}{\mu} W_1(u_k(B(w))) \tag{11}$$

where $B(w)$ is bequest left by person w and c is an arbitrary constant. Finally, denoting estate of person w as $\tilde{E}(w)$ yields the estate tax schedule implementing the correct marginal incentives of $T(E) = \tilde{T}(\tilde{E}^{-1}(E))$. Implementing such a scheme requires an adjustment to the optimal income tax schedule as well, but again this is feasible to implement.

While imposing continuum of corrective taxes to deal with continuum of interpersonal externalities from giving may have seemed like a daunting task, it turns out feasible under the assumptions of one-dimensional heterogeneity implying that estates and bequests increase with the type. Furthermore, in this case, the implementation takes a form of estate taxation.

There are a few interesting features of this result. First, it is an explicit characterization of the optimal estate tax structure. Second, the rates are negative everywhere — $t^E(B) < 0$. This should not come as a surprise because the sole role that the tax plays here is addressing the externality from giving. Third, quite unusually in the optimal tax literature, there is a clear result about the profile of the marginal tax rates: the tax wedge that individuals are facing — characterized by equation (10) — is decreasing with the size of bequest (and, by implication, also with the type and with the size of estate because all three are monotonically related). This is because the marginal tax rate is a function of $W_1'(u_K(B))u_K'(B)$ that is declining due to concavity of welfare function and utility (a slightly weaker assumption of $W_1(u(\cdot))$ being concave would also be sufficient).

Farhi and Werning (2010) describe this result as demonstrating progressivity of the optimal estate tax schedule. Progressive subsidies are not the first thing that comes to mind when thinking about treatment of estate taxation for redistributive purposes. As should be clear from the discussion above, this result is purely driven by the assumed externality from giving. The role of the estate tax is to facilitate redistribution across generations rather than within generations — the latter role is played by income taxation. Hence, the message does not fall far from the basic Atkinson-Stiglitz logic — redistribution across members of the same generation or across dynasties does not call for estate taxation.

There is one important feature of the solution should be pointed out: asymptotically, the marginal estate tax rate goes to zero because the marginal utility tends to zero as wages and bequests increase. As Kopczuk (2009) suggested and Kaplow (2009) further elaborates, for the purpose of evaluating marginal tax rates at the top of the distribution, externalities from giving are irrelevant. The intuition is simply that even if such an externality present and recognized by the social planner, it involves transfers between wealthy parents and wealthy children. Hence, the marginal impact on overall welfare is negligible when bequests are large but, as seen before, the cost of addressing externalities is driven by the overall cost of public funds.

This analysis could be easily extended to incorporate other types of bequest motives. For example,

analysis with the joy-of-giving motive is almost identical. Consider using $u(C^P, L) + v(B)$ rather than preferences in (3) and continue to assume that the utility of the child is $u(C^K)$. In this formulation, $B = C^K$, so that the objective of the planner that puts weight ν on the next generation can be expressed as $u(C^P, L) + v(C^K) + \nu u(C^K)$: the only difference relative to objective under altruism (formula 5) is the joy-of-giving component replacing $u(C^K)$. In particular, the externality term remains exactly the same, and exactly the same analysis as before goes through.

4.3 Accounting for inheritance received

The model considered so far involved two generations — parents and children only.

Kopczuk (2001) considers a different extension of the Mirrlees model by allowing for both bequest decisions *and* heterogeneity in inheritance received. He imposes the steady state restriction that the distribution of bequests received should be the same as that of bequests left. Under the simplifying assumption of perfect correlation in skills across generation, it implies that on the individual level bequest received and left should be the same. He does not consider dynamics and transition to this steady state and instead maximizes over steady state allocations — a "golden rule"-like exercise. In other words, he searches for the welfare-maximizing incentive compatible allocation that remains stable i.e. meets steady-state criteria.[14]

The main result can be seen by considering a special case of the more general utility that he considers: the joy of giving formulation of the form $u(C) + v(L) + g(B)$ where C is consumption, L is labor supply, B is bequest left. Denoting by X the bequest received, the budget constraint is of the form $X + wL - T(B, wL) = C + RB$.[15] The bequest received, X, is taken as given by an individual but the planner's problem imposes the steady state constraint that $B = X$ for all individuals. Changing variables as $D = C - X$, the individual problem can be expressed as $u(D + X) + v(L) + g(B)$ subject to $wL - T(B, wL) = D + RB$. In this (equivalent to the original one) formulation, B, L and D are the choice variables and the externality acts through the utility rather than the budget constraint.

As before, this is a modification of the Atkinson-Stiglitz setup with an externality from bequests. However, contrary to the cases previously considered, the external effect has a very specific form: it generically interacts with consumption. Absent externality, labor income tax would be sufficient. The targeting principle discussed previously calls for the corrective subsidy of $-\frac{Ru'(X+D)}{\mu}$. Contrary to the two period model exemplified by the analysis of Farhi and Werning (2010), the Pigouvian tax rate is not just a function of X but rather it interacts with the level of consumption. For the same reason and more importantly, X also interacts with incentive constraints.[16] It turns out

[14] He does not show that the economy will converge to that allocation and ignores welfare implications of the transition.

[15] R is endogenized by considering constant returns to scale technology in aggregate bequests and aggregate labor.

[16] Interestingly, all specifications considered by Farhi and Werning (2010) do not have this feature. They consider a two period model with perfect correlation of skills so that the steady state interaction between bequests received and left is not present. They also consider (discussed later) an infinite horizon model with i.i.d. skills. In that case,

that it is important: the optimal formula for estate taxation includes three additive components: an aggregate term,[17] Pigouvian correction (that can be negative) and a term proportional to the product of the multiplier on the incentive constraint and $\frac{\partial}{\partial X}\left\{\frac{g'(B)}{u'(D+X)}\right\} > 0$ that gives rise to a positive contribution to the tax rate. Kopczuk (2001) suggests that one way of thinking about it is that in the presence of an externality from giving, bequests are a form of "income" that carries informational content about individual skill level and therefore should be taxed. This is obviously a very stylized model, in particular it ignores dynamics that is considered in the next section and instead focuses on the question of the properties of a "golden rule" steady-state without accounting for the transition to it. However, it allows for incorporating the dual role of bequests in a very tractable way.

The key points so far are applicability of the Atkinson-Stiglitz logic to bequests and clarifying the role of externalities from giving. Externalities from giving tend to imply subsidies to bequests, while the baseline Atkinson-Stiglitz case implies no tax on bequests. Combining the two in the intergenerational context that does allow for the inheritance received to be partially "exogenous" seems to point in the direction of estate taxation playing a role.

In a very recent paper, Piketty and Saez (2011) follow the approach of focusing on the welfare of a steady state generation to characterizing the optimal Mirleesian policy and stress the effect of receiving an inheritance. They consider a model that allows for imperfect correlation of abilities across generations. In doing so, they relax the assumption of one-dimensional differences between individuals that was made in the papers discussed so far. They find the role for bequest taxation and argue that it is driven by multi-dimensionality of the steady state distribution: both labor income and inheritances are (partially) independent sources of information about individual circumstances.

A few extensions of this basic framework have been considered in the literature. Blumkin and Sadka (2004) assume altruistic parents, introduce mortality risk and assume away the existence of annuity markets; they also allows for double-counting of children utility. They analyze linear (income, capital and estate) taxation only. As discussed before, as long as the estate tax declines with age at death, estate taxation provides insurance benefit and a small estate tax is welfare improving (Kopczuk, 2003b). This is true in the model of Blumkin and Sadka (2004). Bequests in their model are a mix of accidental and intentional and the estate tax can be shown to decline with the strength of the bequest motive reflecting the trade-off between distortions to intentional bequests and insurance benefit.[18] They also consider shutting down the altruism of parents in the model and investigate optimality of 100% tax on purely accidental bequests and argue that it rests on the effect on aggregate labor supply of the second generation: if (on the margin), revenue neutral reduction in the estate tax coupled with an increase in the lump tax on the second generation results in higher aggregate labor supply, the 100% tax need not be optimal. Note that the bequests are

the externality from giving externality is present and the steady state distributions of bequests left and received have to coincide, but each bequest is distributed over the whole population so that the externality is a function of X and aggregates: $-\frac{R\mathrm{E}[u'(X+D)|X]}{\mu}$ and does not affect incentive constraints either.

[17] In the presence of a positive externality from giving, there is an incentive increase the flow of bequests in aggregate.

[18] This result is shown assuming logarithmic preferences

lump-sum income from the young generation's point of view so that this exercise is effectively about tweaking the distribution of lump-sum transfers to the young generation — the lump-sum tax adjustment is a uniform tax, while the impact of the estate tax adjustment varies with the size and timing of accidental bequests. It is these distributional differences that are key for the result. Blumkin and Sadka (2004) provide a numerical example in which the effect is strong enough to make 100% estate tax not optimal. It remains unclear whether this mechanism would survive in the nonlinear income tax context.

Farhi and Werning (2010) also consider a number of extensions of their basic framework. Variation in the number of children requires the estate tax schedule to vary with the number of children to restore the equivalence between inheritance and estate taxation. They also consider imposing a non-negativeness restriction on the estate tax rate and allowing the rate of return to be endogenous. In that case, they show that the positive estate tax rate above a threshold is optimal: the intuition for this result is that reduction in bequests raises the rate of return and this effect serves as a substitute for an explicit bequest subsidy.

4.4 Dynamic issues and relationship to optimal capital taxation

Taxation of wealth and bequests is a form of a tax on capital. In this section, I briefly review the main results about capital income and wealth taxation in general. Models of capital taxation in finite or infinite setting with altruistically linked individuals can usually be interpreted as very simplified models of taxation of bequests, with each period corresponding to a different generation. I do not review here work on capital taxation with overlapping generations that are not explicitly linked by some form of bequest considerations or work on capital income taxation that is explicitly within the lifetime (e.g., focusing on age-dependent features).

Literature on optimal capital income taxation in the long run is vast. Chari and Kehoe (1999) provide an able survey of the older work on the topic that considered capital income taxation in a growth framework with linear restrictions on available instruments and redistributive issues ignored.[19] The key result (Chamley, 1986; Judd, 1985) is that the optimal capital income tax rates (or, more generally, tax rates on any accumulated factors including human capital, see Milesi-Ferretti and Roubini, 1998) should converge asymptotically to zero. The intuition for this result is that any non-zero capital income tax imposes a distortion between consumption in different periods that is increasing with the distance and that cannot be optimal. Atkeson et al. (1999) provide an excellent exposition of this argument.

Another key contribution is due to Aiyagari (1995) who introduced non-trivial heterogeneity. He assumes that markets are incomplete and allows for uninsured idiosyncratic risk and borrowing constraints. In this context, he showed that there is a role for capital income taxation. The intuition for this result can be seen by considering the Euler equation for an unconstrained individual:

[19] This is referred to as Ramsey taxation — it builds on Ramsey growth model framework *and* is an extension of the Ramsey commodity taxation problem.

$$u'(C_i) = \rho(1 + r(K_i))\mathrm{E}[u'(C_{i+1})] \tag{12}$$

where $r(K_i)$ is the rate of return rate in period i, expressed as a function of the level of capital stock.

In the ergodic steady state, the distribution of consumption in each period is the same. Suppose then that we integrate equation (12) over the whole population: this would result in the population expectation of $u'(C)$ on both sides and yield $\rho(1+r(K_i)) = 1$ if that expectation was finite. It cannot be finite however. To see it note that $u'(C_i) = \mathrm{E}[u'(C_{i+1})]$ implies $u'(C_{i+1}) = u'(C_i) + \varepsilon_i$ where $\varepsilon_{i+1} \perp u'(C_i)$ and $\mathrm{E}[\varepsilon_{i+1}^2] > 0$, so that $\mathrm{var}(u'(C_{i+1})) > \mathrm{var}(u'(C_i))$ which violates the ergodicity assumption. As Atkeson and Lucas (1992) demonstrate, efficiency in fact requires immiseration so that inequality is ever increasing and $\mathrm{E}[u'(C_i)]$ grows over time. By allowing for liquidity constraints, Aiyagari (1995) breaks the Euler equation for some individuals:

$$u'(C_i) \geq \rho(1 + r(K_i))\mathrm{E}[u'(C_{i+1})] \tag{13}$$

with inequality for individuals whose borrowing is constrained. As the result, he demonstrates that a stationary steady state exists but that it implies $\rho(1 + r(K_i)) < 1$ so that the rate of return is below and the capital stock is above the golden rule level. Intuitively, precautionary saving acts to increase saving on the individual level and leads to overaccumulation of capital and Aiyagari (1995) shows that positive capital income tax is welfare improving.

Saez (2002b) introduces heterogeneity in the Ramsey model a different way. He assumes that dynasties differ permanently with respect to their initial wealth (and assumes away other heterogeneity) and considers the capital income tax that applies above a certain threshold — mimicking the structure of the U.S. estate tax. He shows that the optimal tax of that kind has a finite threshold under reasonable assumptions about the shape of wealth distribution (sufficiently thick tail) and intertemporal elasticity that guarantee that tax distortions are not too strong[20] Interestingly, a policy of this kind reduces wealth accumulation of high wealth holders to the threshold level but does not distort long-run capital accumulation (Piketty, 2001) because asymptotically all remaining wealth is untaxed.

Work in this tradition has an important flaw: it imposes ad hoc restrictions on the set of available tax instruments. Most obviously, linearity of capital income and labor income taxation is unrealistic. More subtly, assumptions about government commitment and its ability to save play an important role.

The new dynamic public finance literature initiated by Golosov et al. (2003) seeks to remedy that by embedding dynamic capital income taxation questions in the Mirrleesian framework. It

[20] As in optimal income tax literature,raising marginal tax rates for high income population has mostly inframarginal effect (Diamond, 1998; Saez, 2001) when the tail of the distribution is thick.

considers individuals with unobserved and stochastically evolving ability. One way of thinking about the standard Mirrlees model is that the undistorted allocation is efficient but inequitable and the planner's problem is to address such an inequity. Another is to interpret income taxation as insurance against lifetime risk: such insurance does not result in the full information first-best allocation because inability to observe state of the world (individual skill type) generates the moral hazard problem (reduced effort). In a way, there is a market failure in the standard Mirrleesian framework too — asymmetric information does not allow for implementing the first-best lifetime insurance scheme — but interpretation of this framework as insurance rather than redistribution is very stylized. Some extensions of this framework as lifetime income insurance have been considered in the literature (see e.g., Eaton and Rosen, 1980).

The new dynamic public finance literature adds to the picture dynamics so that incompleteness of insurance markets explicitly kicks in within the time frame of the model rather than behind the veil of ignorance. The role of the policy is to insure, the cost is moral hazard due to reduced incentives to exert costly effort in the presence of insurance. It turns out that intertemporal distortions are required to mitigate that disincentive effect. The basic result builds on insights of Diamond and Mirrlees (1978) and Rogerson (1985) and can be illustrated as follows. Consider a single individual and two period model. Suppose that the utility is additively separable in consumption and effort

$$u(C_0) + \rho \mathrm{E}[u(C_1(\theta)) - v(e(\theta))] \tag{14}$$

and that the resource constraint is given by $(1+r)C_0 + \mathrm{E}[C_1(\theta)] = \mathrm{E}[\theta e(\theta)]$ where the rate of return is given by r and θ (state of the world, interpreted as skills) is unobservable. The new dynamic public finance adopts the mechanism design approach to characterizing Pareto efficient allocations. Suppose that an individual chooses to report its type as θ'. A change in reported type from θ' to θ'' results in change in utility by $u(C_1(\theta'')) - u(C_1(\theta'))$. If we modify the profile C_1 to $\tilde{C}_1$ so that utility in every state of the world changes by exactly the same amount ε, $u(\tilde{C}_1(\theta)) \stackrel{\mathrm{df}}{=} u(C_1(\theta)) + \varepsilon$, it will not change the report of the individual because $u(C_1(\theta'')) - u(C_1(\theta')) = u(\tilde{C}_1(\theta'')) - u(\tilde{C}_1(\theta'))$ for all θ' and θ''. In other words, the incentive compatibility constraint is unaffected by such a modification. Put differently, we are considering a change in second period tax-and-transfer scheme that results in no behavioral response — a uniform lump-sum adjustment would distort labor supply decisions but a transfer combined with offsetting marginal rate adjustment can keep it intact. The modification for small values of ε is given by $\tilde{C}_1(\theta) = C_1(\theta) + \frac{\varepsilon}{u'(C_1(\theta))}$. At the optimum, shifting consumption from period 0 to period 1 in an incentive compatible way should have no impact on welfare so that, accounting for the storage technology that allows for transferring consumption between periods at the rate of $1 + r$, implementing this variation implies

$$u'(C_0)\mathrm{E}\left[\frac{\varepsilon}{u'(C_1(\theta))}\right] - (1+r)\rho \mathrm{E}[u'(C_1) \cdot \frac{\varepsilon}{u'(C_1(\theta))}] = 0$$

and this equation can be simplified as

$$\frac{(1+r)\rho}{u'(C_0)} = \mathrm{E}\left[\frac{1}{u'(C_1(\theta))}\right] \tag{15}$$

This is the inverse Euler equation formula. It represents a necessary condition for the allocation to be Pareto efficient given incentive constraints — it was obtained by appealing to a shift from consumption in period zero to consumption in period one in a way that implies no resource cost and has no effect on relative utility comparisons between states of the world in period 1 thereby leaving incentive constraints intact. This formula differs subtly but importantly from the standard Euler equation (12), that can be written as $\frac{(1+r)\rho}{u'(C_0)} = \frac{1}{E[u'(C_1(\theta))]}$ — the expectations in equation (15) are of $1/u'$ rather than u'. The concavity of $1/x$ implies that $\frac{1}{E[u'(C_1(\theta))]} \leq \mathrm{E}\left[\frac{1}{u'(C_1(\theta))}\right]$ (with equality if and only if there is no uncertainty) so that the inverse Euler equation implies that consumption is should be distorted toward period 0. In particular, it implies that individuals should not be allowed to invest at the rate r but rather should face a positive "wedge" between current and future consumption, hence introducing a rationale for capital taxation.

The inverse Euler equation is a necessary condition for the constrained Pareto efficient allocation in a dynamic setting: it describes the optimal program in the presence of private information. There is a trade-off between insurance and incentives to provide effort: provision of insurance (equalizing marginal utility across states) weakens incentives to provide effort. The way to (partially) restore work incentives is to discourage saving.[21] The lesson carries over to multiple periods and infinite horizon settings with arbitrary data-generating process for skills (Golosov et al., 2003).

Applied in the bequest context, the model would call for discouraging bequests in order to stimulate effort of the younger generation — it seems (though has not been seriously explored) that the case for the importance of this channel should hinge on the empirical effect of inheritance on labor supply of children. We will discuss such evidence in Section 5.5.

The implementation of the optimum turns out to be more complex than simply coming up with the deterministic marginal capital income tax rate t that would make the solution to the standard Euler equation governing individual choice $\frac{(1+r(1-t))\rho}{u'(C_0)} = \frac{1}{E[u'(C_1(\theta))]}$ satisfy the inverse Euler formula, equation (15). In the stochastic setting, there are in principle more ways to impose the wedge because the tax rate may vary depending on the state in the second period. In terms of implementation, it means that the tax rate on saving may depend on (current and past) labor income and, in fact, it turns out that this is optimal.

One lesson from this line of work is that there are many ways to implement the optimal allocation. Of course, this is also true without uncertainty and/or dynamics: for example, Fullerton (1997) considers implications of various normalizations in the context of taxation of externalities and large literature analyzes relationship between income and consumption taxation, (see e.g., Auerbach, 2006). The dynamic context allows for a rich set of instruments that includes current labor and capital income, assets as well as history of these tax bases and uncertainty adds richness of interac-

[21] The model builds in normality of second-period leisure via the assumption of additive separability.

tions, so that the quest is for a "simple" and realistic implementation. For example Kocherlakota (2005) proposes an implementation that has zero wealth tax rate on *average*, but that rate depends in general depends on the history of labor income reports (the wealth tax in each state of the world can be linear in wealth). Albanesi and Sleet (2006) propose an implementation that depends on current assets and income only but it applies only in a setting with shocks that are i.i.d., i.e. it rules out persistence of productivity over time. Golosov and Tsyvinski (2006) highlight the importance of asset testing in the empirical implementation.

As discussed by Kocherlakota (2010) and Salanié (2011, chapter 6), the implementation requires intertemporal distortion to be negatively correlated with the level of labor income or, in other words, the marginal tax rate on capital or assets to decline with labor income. This has two noteworthy implications for thinking about estate or wealth taxation. First, the optimal taxes are non-trivially joint functions of income and assets, possibly involving the full history of these variables. This is in contrast to the actual estate and gift taxation in the United States that operates independently from income taxation. It is also in contrast to the important types of capital taxation such as capital gains and (currently) dividend tax that impose linear tax rates with relatively minor interactions with the rest of the tax system. It is also in contrast to corporate taxation that, while complex and to some extent nonlinear, does not account for other taxes paid by shareholders. Second, all proposed implementations feature taxation capital or wealth taxation that either explicitly or on average fall with the current income/productivity.[22] As the result, this line of work appears to provide arguments for capital taxation at the bottom of the distribution — for example, asset testing in the context of disability/welfare programs — rather than lessons for understanding potential optimality of capital taxation at high income levels.

This vibrant literature delivers new and interesting insight but has remained somewhat stylized in its empirical applications. In particular, there were no attempts to express the results in terms of empirically estimable quantities along the lines of Saez (2001) (in the context of optimal income tax). This approach relies on the presence of private information but also on individuals placing value on insurance. Indeed, trivially, there would be no inefficiency due to the presence of private information and no role for taxation if individuals were risk neutral. More subtly, the literature has not (yet?) attempted to carefully disentangle the implications of uncertainty, risk attitudes and incentives for the shape of the optimal tax schedule in this context. The absolute utility gains from insurance are high when marginal utility is high so that the planner's objective should be to deliver utility in those case *and*, by the logic of the inverse Euler equation, distort accumulation decisions in situations that are correlated with experiencing high marginal utility. In the contexts considered by the new dynamic public finance literature, this implies little role for distorting intertemporal margin in states of the world that correspond to experiencing low marginal utility from consumption. As

[22]The exception is a recent paper by Werning (2011) that proposes an implementation that features history-dependent labor taxation and savings tax that is independent of the current shock (though history-dependent, although sufficient conditions for history independence are discussed). The unique feature of this implementation is that savings are always set at *zero* and the role of the savings tax is to implement that as an optimum with transfer of resources across periods taking instead the form of adjustments to the labor income tax schedule.

the result, as of now, this literature provides little constructive insight for thinking about tax policy toward the top of the distribution.

The only contribution in the new dynamic public finance that explicitly incorporates bequests is again Farhi and Werning (2010) who consider an infinite horizon version of their model basic model with the externality from giving, allowing for abilities uncorrelated across generations. They show that the presence of the giving externality modifies the inverse elasticity rule familiar from this literature precisely by the additive component reflecting the externality as before. Adapting the implementation in terms of linear (but history dependent) taxes on inherited wealth as in Kocherlakota (2005), they show that the average inheritance tax rate (that would be zero absent externality from giving) has exactly the same structure as given in equation (10) thereby preserving their conclusions about negativity and "progressivity."

5 Behavioral responses to transfer taxation

The empirical work on the impact of estate taxation on taxpayers' decisions is marred with empirical difficulties. On the conceptual side, the question is how to identify the effect of the tax that will apply at the time of death — which is uncertain and, *in expectations*, many years away — on current behavior. Alternatively, one might wonder how what is observed at the time of death might have been impacted by tax policy regime(s) earlier in life. Potentially long lag between behavioral response and the ultimate taxation,[23] makes it difficult to credibly establish the link between estate taxation and behavior. Still, certain aspects of taxpayer behavior can be studied over shorter-term. In particular, studying the effect of gift taxation and behavior very late in life can be more conclusively related to tax incentives.

The second important issue has to do with availability of data. With some exceptions, estate tax returns are not public information and, because the estate tax applies to high net worth population, standard surveys that do not focus on high net worth population are of limited use. Probate records and studies in other countries are additional sources of data.

Despite these difficulties, the literature has made some strides into understanding the impact of transfer taxation on behavior.

Perhaps the most basic question is about the effect of estate taxation on wealth accumulation. The simplest approach is to consider certainty framework that ignores dynamic dimension, with individual maximizing utility $u(C, L, B)$ subject to the constraint $C + RB = y - T(E)$ where $E = wL - C$ is the size of the estate. Following the approach of Feldstein (1995, 1999), the literature on responsiveness to income taxation (recently surveyed by Saez et al., 2011) focused on the "sufficient statistic" for behavioral response — the responsiveness of taxable income. A

[23] In one of the most famous examples of effective estate tax planning, Sam Walton set up a family limited partnership that allowed for great majority of his estate to pass tax free to his wife and children in 1953, 39 years before his death.

similar sufficient statistic argument can be applied in the estate tax context. While the estate tax can affect behavior on many possible margins (even in this simple formulation, labor supply and bequests may both respond), the welfare impact of the estate tax should be summarized by the impact of the tax on taxable base — in this case, the size of the estate. This would be so even when we add other types of margins, such as tax avoidance for example. This argument relies on considering a small change in the marginal estate tax rate dt above some threshold $\bar{E}$ (so that the tax is $T(E) + dt \cdot (E - \bar{E})$). By the envelope theorem, the impact of that change on the overall level of utility is $-u_C \cdot (E - \bar{E}) \cdot dt$ while the impact on revenue reflects the effect of the change in the tax rate on the size of the estate: $T'(E) \cdot dE + dt \cdot (E - \bar{E})$. In either case, it is the level and responsiveness overall estate rather than the composition of its response that matters. Hence, by extension, focusing on the effect of estate taxation on the size of estate at the time of death is a natural starting point for understanding the efficiency cost of estate taxation.

It is obvious that this simple framework misses a lot of things. Comprehensively applying the Feldstein's argument requires understanding the effect on the *overall* tax liability. Even within the simplest framework, it calls for estimating the effect on another main source tax revenue present — income taxation. Reducing labor supply is one obvious margin that the donor can respond on but by far not the only one. Investment decisions and occupational choice might respond to taxation and have implications for income and corporate tax base. Delaying capital gains realization due to step-up in basis at death reduces capital gains tax revenue. Taxpayers who might respond to the estate tax by increasing their charitable contributions, might do so via giving in life with income tax consequences. Avoidance strategies that rely on freezing the value of estate and transferring ownership to beneficiaries might shift taxable income (not just estate) to other individuals. The margins of interest for understanding tax consequences of changes in transfer taxation are responsiveness of inter vivos gifts, and life-time taxes such as due to the impact on income, capital gains or corporate tax base. The response of transfers also naturally has implications for the behavior and tax liabilities of the beneficiaries.

Similarly as in the case of taxable income, focusing on the tax base has limitations. In the presence of tax evasion or other situations where revenue or welfare spillovers are present, decomposing the response into "real" and "shifting" component is important as pointed out by Slemrod (1998) and elaborated by Chetty (2009). Additionally, the response of the tax base to tax incentives is informative about implications of narrow reforms modifying tax rate structure but not about implications of reforms that might affect the base or other non-rate aspects of the system that may themselves affect the magnitude of the behavioral response (Slemrod and Kopczuk, 2002). Last, but certainly not least, the effect of taxation on real quantities is the relevant parameter to know for many non-tax related questions.

In what follows, I review empirical evidence on major types of responses to transfer taxation but begin with clarifying the magnitude of distortions that are caused by these types of taxation.

5.1 Magnitude of distortions

Taxes on transfers are related to other forms of taxation. Most obviously, a tax on estates or bequests is a form of a tax on wealth. To see that, consider first the following asset accumulation equation:

$$A_{k+1} = (1 - t^w)(1 + (1 - t^k)r)\,(A_k + S_k) \tag{16}$$

where A_k represents assets in period k, t^w is the annual wealth tax, t^k is the tax imposed on the returns to savings, r is in the interest rate and S_k is new saving as of the end of period k.

For simplicity of exposition, I am assuming here that an individual is a saver so that $S_i \geq 0$ for all i. In the event of death at the beginning of period $t+1$, after tax bequests are given by $B_{k+1} = (1-t)A_{k+1}$. Iterating equation 16, yields

$$B_{k+1} = (1-t)\left(R^{k+1}A_0 + \sum_{i=0}^{k} R^{k-i}S_i\right) \qquad \text{where} \qquad R = (1-t^w)(1+(1-t^k)r) \tag{17}$$

Assuming first no saving, $S_i = 0$ for all i, the tax on estate is equivalent to an annual wealth tax given by $1-t^w = (1-t)^{1/(\mathrm{k}+1)}$. For example, assuming $k+1 = 50$ and $t = 0.35$ implies $t^w = 0.0086$: obviously, the longer the horizon the lower the equivalent wealth tax. Similarly, there is equivalence here between capital income tax and the estate tax: $1 + (1-t^k)r = (1+r)(1-t)^{1/(k+1)}$. Again, as an example, additionally assuming 5% rate of return, one obtains $t^k = 0.18$. It is straightforward to show that the equivalent capital income tax rate is a decreasing function of the rate of return: the return on the taxed amount does not accrue to the taxpayer and this effect is more important when the rate of return is higher.

Adding saving to the picture complicates the analysis: the equivalent wealth or capital income tax rate depends on when saving takes place — the shorter the horizon, the lower the rates. The main point here is that the horizon and rate of return matter: the estate tax is an infrequent tax so that it is mechanically less burdensome relative to annual taxes as horizon increases and it provides a deferral advantage that grows with the horizon and the rate of return.

This discussion illustrates how one might compare the effective tax rate under the estate tax to that under other types of capital taxation, it does not appropriately describe the distortion induced by the tax. The 35% estate tax rate from the example above has different incentive effects than annual capital income tax rate of 18%: holding bequest constant, estate tax does not distort lifetime consumption profile while capital income tax does.

Evaluating the marginal estate tax rate that applies at the time of death is a reasonably well-defined exercise although even that is not always straightforward. The actual tax liability is affected by many factors including the reliance on the marital and other types of deductions, valuation

discounts, interaction with state taxation and deferred payment schedule among other things.[24] We will delay talking such complications until reviewing related empirical papers.

The actual marginal tax rate at the time of death is directly relevant for the decision that take time around that point — deathbed estate tax planning or decisions made by the executor of the estate. The decision of the donor should be governed by *expectations of* the estate tax rate at the time of death rather than its actual value.

Poterba (2000a) asks how the marginal investment decisions are affected by the presence of estate taxation. He points out that saving rate of return between period i and $i+1$ is not affected by the presence of the estate tax if there is no mortality risk. Denoting mortality rate by m_i, the expected rate of return is given by $(1-m_i)(1+(1-t^k)r)+m_i(1-t)(1+(1-t^k)r) = (1-m_it)(1+(1-t^k)r)$: on the margin, the estate tax is equivalent to a wealth tax at the rate of m_it_i. Equivalently, the rate of return is $1+(1-t^k)r-m_it(1+(1-t^k)r)$. An inspection of this formula reveals that the effect of the estate tax on the marginal rate of return declines with mortality rate but otherwise is not very dependent on the rate of return because $1+(1-t^k)r \approx 1$. Hence the estate tax is relatively more important than capital income taxation when the rate of return is low and less important when the rate of return is high.

Note though that this is marginal analysis that applies to investments over short horizon. Suppose that the taxpayer is saving with leaving a bequest in mind. In that case, as the first pass, the mortality rate does not matter. To see it, consider period N that is sufficiently distant to guarantee that a taxpayer dies by then. The value of a marginal dollar of saving as of period N is given by $(1-t)(1+r(1-t^k)^N$: regardless of the timing of death, the estate tax will be paid once, by period N. As long as the marginal increase in tax liability is neutral with respect to the timing of death and the utility from bequest is not a function of age at death, the mortality risk should not enter. Neither of these are assumptions is automatic: the timing of capital gains step-up matters and welfare neutrality of the timing of bequests would depend on the presence of liquidity constraints and precise assumptions about the nature of the bequest motive. Ultimately, the extent to which mortality risk interacts with estate taxation is an empirical (and, as of now, unresolved) question.

Saving/investment decision is just one of the margins that can be distorted by the estate tax. In particular, taxpayers might choose to make transfers in life instead of making them at death. This has nuanced tax consequences under the U.S. (and other countries tax law). The marginal tax rates applying to both are usually different, as we discuss elsewhere in this chapter. Denote the gift tax rate by t^G and contine to denote the estate tax rate as t. For simplicity of the exposition, I assume here that they are applied to the same base — total amount of the transfer — although the actual practice in the U.S. is different.[25] Let's continue to assume that taxpayers are interested in maximizing the total after tax amount of transfers. In the certainty, one shot context, one

[24] The estate tax liability attributable to a qualified business may qualified for payment in installments over a 10 year period.

[25] Gifts are taxed on the tax exclusive basis so that, to make it comparable to the estate tax rate, the marginal gift tax rate should be converted as $t^G = \frac{\tilde{t}^G}{1+\tilde{t}^G}$ where $\tilde{t}^G$ is the statutory rate applying to gifts.

should select the mode of transfer that corresponds to the lower marginal tax rate. In practice, $t^G < t$, so that there is a presumption that gifts are advantageous. Now consider the same decision but instead assume that the transfer at death will take place n years from now (for now assumed certain) and the asset accumulates at the rate of return r. The value of transfer at period n is $(1+r)^n t^G$ or $(1+r)^n t$ so that the comparison is unchanged. However, in practice, the tax schedule is often nonlinear so that (with the abuse of the notation to allow for t and t^G to vary with the base), the correct comparison is between $(1+r)^n t^G \left(B^G + G\right)$ and $t\left(B^E + (1+r)^n G\right)$ where B^E and B^G are bases for the gift, estate tax respectively and G is the value of the transfer considered. The gift changes the base for the gift tax by its pre-accumulation value while it changes the base for the estate tax by its (higher) after-accumulation value. This effect tends to make gift taxation further preferred when the nominal rate of return is high and the horizon is long. In the U.S. case, B^E and B^G are (imperfectly) tied to each other and the case for pre-paying the estate tax via gift taxation is further strengthened.

Now suppose that fraction β of the asset to be transferred corresponds to appreciated taxable gain to be taxed at realization at some rate t^K (assume that this is the rate that accounts for the benefits of the deferral). In the U.S., capital gains transferred at death benefit from step-up in basis but gifts do not. Hence, the cost consequences of the bequest are unaffected by this modification but those of the gift are. The basis of the asset is adjusted for gift taxes paid that are attributable to the capital gains liability so that there is an incremental tax liability of $(\beta - \beta t^G)t^K$. One can go further by modeling the benefits of deferral, implications of various liquidation strategies. There are additional complications in case of gifts shortly (3 years) before death that may be treated as bequests and further opportunities to benefit from step up by first transferring the asset to a surviving spouse. See, for example, Joulfaian (2005a) for further discussion and illustrative calculations. Another aspect that has not been accounted for here are differences in the marginal tax rates of the donors and donees that create additional opportunities for tax arbitrage (Stiglitz, 1985; Agell and Persson, 2000). It has been argued that it has important implications in the context of the estate tax (Bernheim, 1987).

5.2 Effect on wealth accumulation and reported estates

A number of papers attempted to relate estate taxation to wealth or estates at death. Kopczuk and Slemrod (2001) use estate tax returns covering selected years between 1916 and 1996.[26] They first pursue aggregate and micro-based analysis using aggregate variation in top marginal tax rate and rates at 40 or 100 average wealth as instruments for the individual marginal tax rate at death. Given lack of convincing identification strategy, the results are not particularly robust. They propose some specifications that attempt to exploit cross-sectional variation. To do so, they note that it is the lifetime tax rate that should matter and they propose using the imputed marginal

[26]The IRS has complete micro data for period 1916-1945 and samples for 1962, 1965, 1969, 1972, 1976 and annually starting from 1982 (though with varying coverage and design). See McCubbin (1994) for a description of the pre-1945 data and Johnson (1994) for the discussion of post-1982 datasets.

tax rate at age 45 as a proxy (instrumented using tax rates at the fixed point of the distribution as before). This approach introduces variation in the marginal tax rate at any particular point in time that is driven by variation in age of decedents, and this variable turns out to dominate other measures of tax burden in "horse-race" specifications It corresponds to a significant net-of-tax elasticity of net worth at death of about 0.16. Holtz-Eakin and Marples (2001) use Health and Retirement Survey to estimate the effect of estate taxation on wealth of the *living* population. They primarily rely on cross-sectional variation in state inheritance and estate tax rates to identify the effect. This is arguably a more credible source of variation than age-variation considered by Kopczuk and Slemrod (2001) and focusing on the living individuals allows for interpreting the results as the response to estate taxation expected in the future (under the assumption that current rates are a prediction of future tax rates). However, the HRS data contains few high net worth individuals and cross-sectional variation may not deal adequately with location-based heterogeneity. Joulfaian (2006a) pursues a systematic attempt to exploit time variation to explain the size of estates. Rather than using the marginal tax rate at death, he uses an ("representative") equivalent marginal tax rate 10 years before death. That rate is constructed using a stylized procedure that follows the insight of Poterba (2000a): that rate is obtained by solving for t^k the equation given by $(1+r)^n(1-t) = (1+r(1-t^k))^n$, where$t$ is the tax rate that applies 10 years before death, r is linked to the growth rate of S&P 500 and constant life-expectancy of $n = 15$ or $n = 20$ years is assumed.

As can be inferred from this brief summary, as of yet, the literature has not been able to come up with a convincing empirical strategy to estimate this key dimension of the response. It is worth noting though that all these papers estimate similar baseline elasticity of net worth/reported estate estimates with respect to the net-of-tax rate of between 0.1 and 0.2.[27] The baseline specifications in each case attempt to shed a light on the response to incentives over the lifetime (rather than those that are ultimately realized at death), but use different dependent variables: in Holtz-Eakin and Marples (2001) it is wealth at some point while alive measured in survey data, while the other two papers focus on estate at death as reported on tax returns. Taken at face value, these results would be consistent with the notion that tax avoidance is not the main driver of response, the topic to which we will return below.

5.3 Inter vivos giving

In contrast to the work on the response of wealth and estates, the literature has made more significant strides in estimating the effect of taxation on giving while alive in a more convincing fashion. The U.S. and many other countries tax gifts and estates separately, creating opportunities for tax avoidance (see Nordblom and Ohlsson, 2006, for a theoretical analysis). In the U.S., estate and gift taxes have operated completely independently since 1932 (when the gift tax was introduced)

[27] Both Holtz-Eakin and Marples (2001) and Joulfaian (2006a) do not estimate the elasticity directly but discuss converting their estimates to those obtained from the "standard" log-log specification used by Kopczuk and Slemrod (2001).

until 1976. Since 1977, the gift and estate taxation have been integrated, that is gifts reduce the size of exemption available at the time of death. Since the very beginning, the gift tax applied to lifetime gifts, that is gifts made in the past are accumulated to provide a lifetime basis. Also, since the beginning of its existence, the gift tax rates have been lower than estate tax rates: explicitly initially and through the distinction between tax-inclusive and tax-exclusive basis since.[28] Gift taxation allows for annual exemptions[29] and interacts in non-trivial way with step-up in basis at death. The final component of the system in the United States is the Generation-Skipping Transfer Tax that's imposed on transfers that skip a generation.[30] See Joulfaian (2004) for the history of changes in gift tax provisions.

The tax advantaged nature of gifts usually provides an incentive to transfer inter vivos rather than in death. This incentive is particularly strong in the case of assets that are expected to appreciate. On the other hand, gifts generally do not benefit from the step up in basis and hence may trigger capital gains tax liability.

Joulfaian (2004) focuses on aggregate gift tax revenue and documents massive spikes corresponding to changes in gift tax rates. In particular, gifts in 1976 — in anticipation of integration of gifts and estates and, simultaneously, an increase in the top gift rate from 57.5% to 70% — quadrupled compared to the previous year, only to decline well below pre-1976 levels for another decade or so. This is further supported by more formal aggregate time-series that converts striking salient features of the time series into large and very significant estimates of elasticity to current and anticipated tax rates. The aggregate evidence strongly indicates that some (presumably large) gifts are very responsive. Ohlsson (2011) documents similar dynamics in Sweden in 1948 just before Sweden instituted a temporary estate tax on top of existing inheritance and gift taxation.

Bernheim et al. (2004) provide a micro-based evidence. They use data from the Survey of Consumer Finances between 1989 and 2001 and focus on the impact of increases in estate tax exemption. The increase should have no effect on people who never expected to be above the exemption, discourage gifts for those who are phase out of the tax reach and possibly increase gifts for very high net worth individuals via wealth tax.[31] They crudely classify individuals into groups that may fall into each category based on their current net worth and do find patterns of gift-giving that are very supportive of the presence of response: gifts for the middle category decline relative to other while gifts for the top category (insignificantly) increase. Page (2003) relies on cross-sectional variation in state estate tax rates and also finds a relationship between the marginal tax rate and the size

[28] The estate tax applies on the tax inclusive basis, so that a dollar of estate yields the tax liability of T' and the gift of $1 - T'$. The gift tax applies on the tax exclusive basis so that the gift of a dollar entails additional liability of T'. As the result, a one dollar expense (to be comparable to estate) results in an after tax gift of $\frac{1}{1+T'}$ and the tax liability of $\frac{T'}{1+T'}$ (which is obviously smaller than T').

[29] $13,000 in 2011, for each donee separately.

[30] In particular, it applies to related individuals who are more than one generation apart (such as grandchildren) and to unrelated parties that are younger by 37.5 or more years.

[31] There were other changes, such as the decline in the capital gains rate, that may have increased the advantage of making gifts by taxable individuals. They argue that the effect is likely small and otherwise would work against finding an effect for the most interesting middle group.

of gifts in the SCF. Joulfaian (2005a) revisits the same question but focuses more carefully on the role of capital gains taxation. The capital gains tax applies to gifts but not to estates that benefit from the step up at death. This matters more when appreciated assets constitute a large part of the estate and when horizon is short. He shows the magnitude of this effect and demonstrates that the advantage of gifts over estates is not universal. Using the relative tax price that accounts for the capital gains considerations and relying on state tax variation, he also finds that gifts are responsive to tax considerations. Arrondel and Laferrère (2001) also document that gifts in France responded to major changes in fiscal incentives.

A number of papers (McGarry, 2000; Poterba, 2001; McGarry, 2001; Joulfaian and McGarry, 2004) focused on a different outcome: the reliance of taxpayers on the annual gift tax exclusion. This is estate tax planning 101: taxpayers can transfer completely tax free up to $13,000 (in 2011) per donee, to as many people as they wish. For example, a married couple with two children can make four such transfers every year (each spouse can make a gift to each child) so that, let's say over 20 year horizon, they could transfer over $1 million tax free (even before adjusting for the rate of return) . Before 2000, which is the period that these studies used, the exclusion was $10,000 and the estate tax threshold $600,000 so that the potential for reliance on this strategy to effectively eliminate tax liability for many otherwise taxable taxpayers was very high. Even with higher exemptions, this continues to be the basic planning strategy. Yet, the key finding in these studies is that this strategy is significantly underutilized by potential estate taxpayers.[32] Poterba (2001) concludes that the results imply that taxpayers fail to minimize tax liability [33]

Joulfaian and McGarry (2004) report, based on linked estate and gift tax data for 1992 decedents, that relatively few (1/3) ultimate estate taxpayers make taxable gifts over their lifetime and that such gifts are infrequent; ultimately the volume of lifetime taxable gifts is of the order of 10% of the estate.

Hence, the literature does find that gifts are very responsive to tax considerations but it also finds that, despite responsiveness, gifts appear to be significantly underutilized as a tax planning tool

5.4 Unrealized capital gains

Poterba and Weisbenner (2001), Wolff (1996), Eller et al. (2001) Auten and Joulfaian (2001)

Poterba and Weisbenner (2003)

[32]Poterba (2001) uses SCF that oversamples high net-worth population and shows that relatively weak gift giving strategy extends to individuals with net worth several times the exemption limit. McGarry (2000, 2001) relies on HRS/AHEAD data that focuses on elderly population and reaches similar conclusion for elderly households that are on average closer to the estate threshold than the SCF sample.

[33]This is contrast to predictions from a stylized frictionless model that has been used as a benchmark elsewhere in the literature. A small literature considered that question previously using aggregate information and illustrative marginal tax rate calculations (Adams, 1978; Kuehlwein, 1994).

5.5 Labor supply of donees

A number of papers have considered the effect of receiving inheritance on the behavior of recipients, mainly on labor supply. This is one of the important dimensions necessary for understanding the efficiency cost of transfer taxation. It is of relevance because it represents an incentive effect and has revenue consequences: one needs to be able to trace the effect of changes in the tax on all sources of revenue in order to understand its efficiency cost. It is also the response that is potentially linked to externalities from bequests.

At its simplest form, inheritance is the type of exogenous, non-wage, income for the donee. Express the utility of the donee as $u^K(C, L)$ and the budget constraint as $C = y + B(t) + wL$, where $B(t)$ is the bequest received given donor's tax rate t and y represents income from other exogenous sources. The labor supply can be written as $L(w, y + B(t))$, making it clear that when B is taken as given by the donee the receipt of inheritance should generate income effect response on labor supply. Under the standard assumption that leisure is a normal good, inheritance should reduce labor supply.

An increase in the estate tax should (other things constant) generate a response in labor supply proportional to the effect the tax has on the size of inheritance. Combined with the effect of taxation on the size of inheritance, the effect of taxation on labor supply of donees could be pinned down if one had a credible estimate of the income effect on labor supply from elsewhere could be sufficient for evaluating the effect of taxation on labor supply of the donee's. As we have seen though, estimating the effect of taxation on inheritance is not trivial (and credible direct estimates of the income responsiveness of labor supply are not abound either). Furthermore, this simple reasoning requires at least three important qualifications.

First, the assumption that inheritance is "exogenous" need not be correct. In particular, the basic prediction of the altruistic model is that the size of inheritance should respond to individual characteristics. In that case, using the simple labor supply framework as before, the bequest itself should be a function of individual characteristics such as the wage rate, $B(t, w)$. When this is the case, simply regressing donee's labor supply on the size of inheritance is not necessarily informative about the tax implications because bequests are correlated with individual characteristics. Tax driven variation remains appealing for the purpose of identifying the labor supply response.

The second qualification to the basic view of inheritance as an exogenous income effect is that bequests are not unexpected (although their timing often might be). Individuals who anticipate a bequest may respond before the actual receipt of inheritance, making it difficult to estimate the effect of inheritance itself. Additionally, a natural concern is the potential presence of strategic interactions. When that is the case, bequests remain endogenous and, furthermore, labor supply of the donee may be affected by characteristics of the donor beyond the size of the bequest itself. The natural way of considering this is in the context of a dynamic framework along the lines considered in the Samaritan's dilemma (Bruce and Waldman, 1991; Coate, 1995) results that breaks Becker's "rotten kid" theorem (Becker, 1974; Bruce and Waldman, 1990; Bergstrom, 2008). Casting it in

an empirical framework, labor supply can be written as $L(w, w^P, y + B(t, w, w^P))$, where w^P is the strategic effect of wage (or other characteristics) of the parent and it is assumed that upon the receipt of inheritance the source of income does not matter. Considering strategic considerations further complicates interpretation of any estimated effect of inheritance, but variation in tax incentives remains a natural source of identification and the reduced form effect of tax on labor supply is in principle the parameter of interest for understanding revenue consequences. However, since the behavior of family can no longer be assumed to be efficient, understanding the strength of strategic interactions is of relevance too for thinking about policy implications.

Thirdly, the effect of inheritance may vary with the characteristics/situation of the recipient. The presence of liquidity constraints is of natural interest for thinking about implications of taxation because in that case estate taxation interacts with other market imperfections.[34]

Holtz-Eakin et al. (1993) framed the question about the effect of inheritance on labor supply as "Carnegie conjecture." Famously, Andrew Carnegie suggested that inheritance makes donees less productive members of society. Anecdotal evidence abound of course and labor supply is not the only margin that may be considered as being "less productive" but it is certainly the one with important economic consequences. Holtz-Eakin et al. (1993) used information from income tax returns of inheritance recipients to study the effect on their labor force participation and earnings and found robust negative effect on participation and some evidence suggesting earnings declines. Joulfaian and Wilhelm (1994) study the same question using PSID and find smaller participation responses in that sample.

Brown et al. (2010) focus on an older population using Health and Retirement Survey. This older sample is more likely to receive inheritance than the general public. The nature of this sample allows them to focus on retirement decision rather than general labor force participation or hours response — this is important, because evidence suggests that labor supply responses are much stronger on the extensive margin and retirement in particular. Importantly, HRS includes question about expected inheritance and, by relying on the first wave of the survey, they construct measures of whether the inheritance was expected and how its size compares to expectations. They confirm the finding of negative participation effect and further show that the effect is stronger for unexpected inheritances.[35]

Elinder et al. (2011) use Swedish tax register data and confirm that the receipt of inheritance reduces labor income (they do not decompose the response into extensive and intensive margins). [36] They also find a short-lived increase in capital income, possibly suggesting temporary consumption increases. This is also consistent with findings of Joulfaian and Wilhelm (1994) who document

[34]Presumably by aggravating them although recall the result of Aiyagari (1995) that calls for capital taxation in the presence of liquidity constraints due to general equilibrium implications of "excessive" precautionary saving.

[35]They also argue that the effect is not driven by grieving — the labor supply estimate is unaffected by inclusion of the dummy for death of one's parent (rather than the parent of the spouse).

[36]They provide weak evidence of potential anticipation effect. This is not necessarily evidence of a strategic response though: reduction in labor supply prior to inheritance receipt may be due to devoting time to taking care of an ill family member.

small consumption increases following receipt of inheritance and Joulfaian (2006b) who documents that wealth response much less than one for one to the receipt of inheritance.

There are other estimates of the effect of unearned income on labor supply and consumption, using variation in lottery winnings (Imbens et al., 2001; Kuhn et al., 2011) and stock market wealth (e.g., Poterba, 2000b; Coile and Levine, 2006), although the results do not paint fully consistent picture. As discussed before, their applicability in the context of bequests requires the strong assumption of inheritance being equivalent to exogenous income. Hence direct studies of the impact of inheritance are of independent interest (and they can always be viewed as a type of unearned income shock). The literature has provided evidence suggestive of negative effects on labor supply of donor, though much remains to be done. None of this work has provided evidence derived from tax variation so that the tax policy relevant effect — the impact of the transfer tax on labor supply of the donee that requires accounting for the effect of the tax on the size of inheritance — has not yet been directly studied.

5.6 Entrepreneurship, family firms and inherited control

Studies that document the effect on labor supply participation tend to find the negative effect of inheritance receipt. On the other hand, Holtz-Eakin et al. (1994b,a) focus on the effect of inheritance receipt on entrepreneurship and survival of existing small businesses. They find that inheritance matters for both and conclude that liquidity constraints are important. Tsoutsoura (2011) uses a 2002 repeal of inheritance taxation in Greece to show the effect of the tax on investment in transferred firms and provides some evidence consistent with the importance of financial constraints in driving this effect. This suggests that the impact of taxation on behavior of the next generation may be substantially more nuanced than negative labor supply effects would suggest. If negative labor supply effects reflect the presence of liquidity constraints, welfare implications of increasing estate taxation would need to account for exacerbating the distortion on this margin. Additionally, the positive effects on entrepreneurship may not have immediate revenue consequences. Furthermore, the evidence about the link between inheritance and lifting liquidity constraints is not bullet proof: Hurst and Lusardi (2004) show that both past and future inheritances predict entrepreneurship suggesting that they may capture either anticipation effects (inconsistent with liquidity constraints) or other factors correlated with both entrepreneurship and inheritance such as preferences or habits (Charles and Hurst, 2003).

While recipients of inheritance may set up a new business, continuing a family firm is another possible and common outcome. It is popularly claimed that forcing beneficiaries to sell a business is an undesirable effect of estate taxation. The economic evidence on this topic is much less clear: a number of papers found that inheritance of control in family firms reduces performance (Pérez-González, 2006; Villalonga and Amit, 2006; Bloom and Van Reenen, 2007).[37] Evidence on whether

[37]Grossmann and Strulik (2010) analyze theoretically whether family firms should face preferential transfer tax treatment. The trade-off they consider is between the cost of firm dissolution and lower management quality. In their

the estate tax has any effect on transfer of control is scarce. Brunetti (2006) uses probate records from San Francisco in 1980-1982 in order to study the effect of the decline in federal and state estate tax rates on the likelihood that decedent's business is sold and finds small positive effects of the tax on the likelihood of selling a business. The results are based on a small sample and variety of imperfect diff-in-diff strategies, but are intriguing. However, if this effect is there and is undesirable, entrepreneurs should pursue strategies to reduce its likelihood. Holtz-Eakin et al. (2001) study life insurance purchases of entrepreneurs and conclude that they do not take full advantage of opportunities to protect their firm from being sold in order to meet the estate tax liability.

6 Tax avoidance responses

6.1 Trade-off between tax minimization and control

The discussion so far made no serious distinction between responses that involve "real" behavior — wealth accumulation, labor supply, magnitude of transfers — and those that are solely intended to reduce tax liability with no real consequences.

As usual in the tax-related contexts, drawing a line between "real" and "avoidance" is difficult. Consider for example an extreme type of response that has been discussed in the literature: Kopczuk and Slemrod (2003) show that during two weeks before/after major estate tax changes, the likelihood of dying in the low tax regime is positively correlated with the magnitude of tax savings; Gans and Leigh (2006) and Eliason and Ohlsson (2008) show similar evidence surrounding the repeals of transfer taxes in Australia and Sweden respectively. The response may be real — the will to live may be strengthened by the benefits to one's beneficiaries (or dislike of the government). Another possible explanation is tax evasion — perhaps death certificate can be forged (possibly more likely with pre-1945 reforms studied by Kopczuk and Slemrod (2003) then in recent years in Australia or Sweden). The response may also be due to avoidance: there may be some control over the timing of disconnecting life support. Sorting out these possibilities is very hard in practice.

It does not require much convincing that estate tax planning does take place in practice and taxpayers are in fact interested in reducing their tax liability — the existence of estate tax planning industry is a prima facie evidence of that. How effective can tax planning be?

In a very influential paper, Cooper (1979) dubbed the estate tax a "voluntary tax." His argument was that with sufficient planning, taxpayers can significantly reduce and perhaps even eliminate tax liability. Some of the strategies he described are no longer available but estate tax planning remains an active arena. The extent of tax avoidance is controversial and naturally hard to estimate.[38] *Elaborate based on IRS studies*

model, preferential taxation can induce a pooling equilibrium in which low ability children (inefficiently) continue a firm.

[38] An ingenious approach of Wolff (1996)and Poterba (2000a) was to compare the estate tax returns to the data

Schmalbeck (2001) argues that the "voluntary" nature of the estate tax ignores an important consequence of all strategies identified by Cooper (1979) (and many others): in order to reduce the estate, the taxpayer has to give up control over assets. Hence, the right framework for thinking about tax planning is not as tax minimization but rather as the trade-off between reducing tax liability and control over assets. A taxpayer who does not value control may be able to significantly reduce tax liability, while taxpayers with significant preference for retaining control will choose not to do so.

Perhaps the most direct evidence in favor of this trade-off is provided by Kopczuk (2007). Relying on (publicly available) estate tax returns filed in 1977, he first shows that the size of estate at death in this very wealthy elderly sample is strongly correlated with age, indicating that wealth accumulation continues until very old age (there is sufficient data to show upward sloping wealth profiles even for people in their 90s). However, since the estate tax return used to contain (reported by the executor of the estate) information about the length of terminal illness, it is possible to evaluate the effect of terminal illness on the size of the estate. It turns out that the effect is very strong — 15 to 20% drop with an illness lasting "months to years." After evaluating alternative explanations, he concludes that the most likely one is tax avoidance (in particular, the composition of assets and deductions reported on tax returns changes in a way indicating tax avoidance). Taken together, it suggests that despite continuing accumulation, wealthy taxpayers underinvest in tax planning until the onset of a terminal illness but they reveal that they value tax reduction by their actions at the end of life. Because tax planning is much more effective when done early, it suggests that taxpayers also forgo significant tax savings. This pattern of behavior requires a combination of the desire to leave a bequest and some form of a reason not to part with wealth while alive. Some form of benefit from controlling wealth is a natural candidate.

The desire to retain control is also consistent with previously discussed evidence about responsiveness of gifts to tax incentives coupled with the level of giving that is grossly insufficient for the purpose of minimizing tax liability. Such behavior could be naturally explained by the simultaneous desire to reduce taxation coupled with control motive.

An alternative explanation, not yet seriously explored in this context, is the possibility of inattention — taxpayers may not be paying attention to tax consequences. While possible, this is also a population that is financially sophisticated and one that in most cases has professional assistance in place. Still, inadequate life insurance holdings by business owners (Holtz-Eakin et al., 2001) could potentially be explained by this motivation.[39]

from SCF *weighted by mortality risk*, with the difference interpreted as reflecting the extent of tax evasion. However, their estimates vary from 70% of tax loss to the very small amount. One difficulty is to estimate the mortality risk for the population of the estate taxpayers; furthermore, as Eller et al. (2001) elaborate, this procedure is very sensitive to assumptions about mortality differences between married and single individuals and about the distribution of charitable bequests.

[39]Motivated by Becker (1973), Kopczuk and Slemrod (2005) model theoretically the "denial of death" behavior with agents rationally repressing information about their mortality (as in Carrillo and Mariotti, 2000; Bénabou and Tirole, 2002) in order to reduce the psychic cost due to high mortality risk.

While the literature has been exploring explanation for wealth accumulation due to precautionary motives — longevity, health care costs, long-term care insurance — such evidence does not appear as applicable for thinking about high net worth individuals who are subject to the estate tax and who one might think have sufficient wealth for such precautionary considerations not to be important. As discussed previously, some form of utility from holding on to wealth appears necessary for successfully explaining the upper tail of wealth distribution and direct microeconomic evidence on this topic remains limited.

6.2 Valuation issues (minority discounts, family limited partnerships, small busineses)

Johnson et al. (2001) and other IRS studies

6.3 Tax evasion

incomplete — IRS and anecdotal evidence on tax evasion or borderline avoidance

7 Other topics

7.1 Wealth-accumulation and concentration externalities

Positional externalities

Superstars, rent-seeking

Wealth concentration and political economy

Equal opportunities

Nepotism, family firms

Carnegie conjecture, Doepke and Zilibotti (2008) — endogenizing preference for leisure among the wealthy

7.2 Charity

Auten and Joulfaian (1996); Boskin (1976); Joulfaian (1991, 1998, 2000, 2005b); McNees (1973)

Bakija et al. (2003) Brunetti (2005)

7.3 Family/marital issues

7.4 Mobility, state vs federal taxation, tax competition

Conway and Rork (2004)Bakija and Slemrod (2004) Conway and Rork (2006)

7.5 political economy of the estate tax

Gale and Slemrod (2001), Graetz and Shapiro (2005), Bertocchi (2011), Scheve and Stasavage (2010)

8 Summary and conclusions

References

Abel, Andrew B. and Mark Warshawsky, "Specification of the Joy of Giving: Insights from Altruism," *Review of Economics and Statistics*, February 1988, *70* (1), 145–149.

Adams, James D., "Equalization of true gift and estate tax rates," *Journal of Public Economics*, 1978, *9* (1), 59–71.

Agell, Jonas and Mats Persson, "Tax Arbitrage and Labor Supply," *Journal of Public Economics*, October 2000, *78* (1-2), 3–24.

Aiyagari, S. Rao, "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints and Constant Discounting," *Quarterly Journal of Economics*, December 1995, *103* (6), 1158–1175.

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

Altonji, Joseph G., Fumio Hayashi, and Laurence J. Kotlikoff, "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data.," *American Economic Review*, 1992, *82* (5), 1177–1198.

_ , _ , and _ , "Parental Altruism and Inter Vivos Transfers: Theory and Evidence," *Journal of Political Economy*, December 1997, *105* (6), 1121–66.

Ameriks, John, Andrew Caplin, Steven Laufer, and Stijn Van Nieuwerburgh, "The Joy of Giving or Assisted Living? Using Strategic Surveys to Separate Public Care Aversion from Bequest Motives," *Journal of Finance*, 2011, *66* (2), 519–561.

Andreoni, James, "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *Economic Journal*, June 1990, *100* (401), 464–477.

Arrondel, Luc and André Masson, "Altruism, exchange or indirect reciprocity: what do the data on family transfers show?," in Serge-Christophe Kolm and Jean Mercier Ythier, eds., *Handbook on the Economics of Giving, Reciprocity and Altruism*, Vol. 2, Elsevier, 2006, chapter 14, 971–1053.

_ and Anne Laferrère, "Taxation and wealth transmission in France," *Journal of Public Economics*, 2001, *79* (1), 3–33. ISPE 1998: Bequests and Wealth Taxation.

Atkeson, Andrew and Robert E. Jr. Lucas, "On Efficient Distribution with Private Information.," *Review of Economic Studies*, 1992, *59* (3), 427–453.

_ , V. V. Chari, and Patrick J. Kehoe, "Taxing Capital Income: A Bad Idea," *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 1999, *23* (3), 3–17.

Atkinson, Anthony B. and Joseph E. Stiglitz, "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics*, July-August 1976, *6* (1-2), 55–75.

Auerbach, Alan J., "The Choice Between Income and Consumption Taxes: A Primer," Working Paper 12307, National Bureau of Economic Research June 2006.

Auten, Gerald and David Joulfaian, "Charitable Contributions and Intergenerational Transfers," *Journal of Public Economics*, January 1996, *59* (1), 55–68.

_ **and** _ , "Bequest taxes and capital gains realizations," *Journal of Public Economics*, 2001, *81* (2), 213–229.

Bakija, Jon and Joel Slemrod, "Do the Rich Flee from High Tax States? Evidence from Federal Estate Tax Returns," Working Paper 10645, National Bureau of Economic Research July 2004.

Bakija, Jon M., William G. Gale, and Joel B. Slemrod, "Charitable Bequests and Taxes on Inheritances and Estates: Aggregate Evidence from across States and Time," *American Economic Review Papers and Proceedings*, 2003, *93* (2), 366–370.

Banks, James and Peter A. Diamond, "The Base for Direct Taxation," in James A. Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Steve Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth D. Myles, and James Poterba, eds., *Dimensions of Tax Design: the Mirrlees Review*, Oxford University Press, 2010, chapter 6.

Barro, Robert J., "Are Government Bonds Net Wealth?," *Journal of Political Economy*, November-December 1974, *82* (6), 1095–1117.

Batchelder, Lily L., "What Should Society Expect from Heirs? A Proposal for a Comprexensive Inheritance Tax," *Tax Law Review*, Fall 2009, *63* (1), 1–111.

Becker, Ernest, *The Denial of Death*, New York: The Free Press, 1973.

Becker, Gary S., "A Theory of Social Interactions.," *Journal of Political Economy*, 1974, *82* (6), 1063–1093.

Bénabou, Roland and Jean Tirole, "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, August 2002, *117* (3), 871–915.

Bergstrom, Theodore C., "Rotten Kid Theorem," in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan, 2008.

Bernheim, B. Douglas, "Does the Estate Tax Raise Revenue?," in Lawrence H. Summers, ed., *Tax Policy and the Economy*, Vol. 1, Chicago: National Bureau of Economic Research; Cambridge, Mass.: MIT Press, 1987, 113–138.

_ **and Sergei Severinov**, "Bequests As Signals: An Explanation for the Equal Division Puzzle," *Journal of Political Economy*, August 2003, *111* (4), 733–764.

_ , **Andrei Shleifer, and Lawrence H. Summers**, "The Strategic Bequest Motive," *Journal of Political Economy*, December 1985, *93* (6), 1045–76.

_ , **Robert J. Lemke, and John Karl Scholz**, "Do Estate and Gift Taxes Affect the Timing of Private Transfers?," *Journal of Public Economics*, December 2004, *88* (12), 2617–34.

Bertocchi, Graziella, "The Vanishing Bequest Tax: The Comparative Evolution of Bequest Taxation in Historical Perspective," *Economics & Politics*, 2011, *23*, 107–131.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

Blumkin, Tomer and Efraim Sadka, "Estate Taxation with Intended and Accidental Bequests," *Journal of Public Economics*, January 2004, *88* (1-22), 1–21.

Boadway, Robin, Emma Chamberlain, and Carl Emmerson, "Taxation of Wealth and Wealth Transfers," in James A. Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Steve Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth D. Myles, and James Poterba, eds., *Dimensions of Tax Design: the Mirrlees Review*, Oxford University Press, 2010, chapter 8.

Boskin, Michael, "Estate Taxation and Charitable Bequests," *Journal of Public Economics*, January-February 1976, *5* (1-2), 27–56.

Brown, Jeffrey R., Courtney C. Coile, and Scott J. Weisbenner, "The Effect of Inheritance Receipt on Retirement," *Review of Economics and Statistics*, 2010, *92* (2), 425–434.

Bruce, Neil and Michael Waldman, "The Rotten-Kid Theorem Meets the Samaritan's Dilemma," *Quarterly Journal of Economics*, 1990, *105* (1), 155–165.

_ **and** _ , "Transfers in Kind: Why They Can be Efficient and Nonpaternalistic," *American Economic Review*, December 1991, *81* (5), 1345–1351.

Brunetti, Michael J., "The Estate Tax and Charitable Bequests: Elasticity Estimates Using Probate Records.," *National Tax Journal*, 2005, *58* (2), 165–188.

_ , "The Estate Tax and the Demise of the Family Business," *Journal of Public Economics*, November 2006, *90* (10-11), 1975–93.

Cagetti, Marco and Mariacristina De Nardi, "Wealth Inequality: Data and Models," *Macroeconomic Dynamics*, 2008, *12*, 285–313.

_ **and** _ , "Estate Taxation, Entrepreneurship, and Wealth," *American Economic Review*, 2009, *99* (1), 85–111.

Carrillo, Juan D. and Thomas Mariotti, "Strategic Ignorance as a Self-Disciplining Device," *Review of Economic Studies*, 2000, *67*, 529–544.

Carroll, Christopher D., "Why Do the Rich Save So Much?" In Slemrod (2000).

Chamley, Christophe, "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, May 1986, *54* (3), 607–22.

Chari, V. V. and Patrick J. Kehoe, "Optimal Fiscal and Monetary Policy," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, number 1C, Elsevier, 1999, 1671–1745.

Charles, Kerwin Kofi and Erik Hurst, "The Correlation of Wealth acrross Generations," *Journal of Political Economy*, December 2003, *6* (111), 1155–1182.

Chetty, Raj, "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," *American Economic Journal: Economic Policy*, August 2009, *1* (2), 31–52.

Coate, Stephen, "Altruism, the Samaritan's Dilemma, and Government Transfer Policy," *American Economic Review*, March 1995, *85* (1), 46–57.

Coile, Courtney C. and Phillip B. Levine, "Bulls, Bears, and Retirement Behavior," *Industrial and Labor Relations Review*, 2006, *59* (3), 408–429.

Conway, Karen Smith and Jonathan C. Rork, "Diagnosis Murder: The Death of State Death Taxes.," *Economic Inquiry*, 2004, *42* (4), 537–559.

_ and _ , "State 'Death' Taxes and Elderly Migration–The Chicken or the Egg?," *National Tax Journal*, 2006, *59* (1), 97–128.

Cooper, George, *A Voluntary Tax? New Perspectives on Sophisticated Tax Avoidance* Studies of Government Finance, Washington D.C.: The Brookings Institution, 1979.

Cox, Donald, "Motives for Private Income Transfers.," *Journal of Political Economy*, 1987, *95* (3), 508–546.

Cremer, Helmuth and Pierre Pestieau, "Non-linear Taxation of Bequests, Equal Sharing Rules and the Tradeoff Between Intar- and Inter-Family Inequalities," *Journal of Public Economics*, January 2001, *79* (1), 35–54.

_ and _ , "Wealth transfer taxation: a survey of the theoretical literature," in Serge-Christophe Kolm and Jean Mercier Ythier, eds., *Handbook on the Economics of Giving, Reciprocity and Altruism*, Vol. 2, Elsevier, 2006, chapter 16, 1107–1134.

_ , Firouz Gahvari, and Norbert Ladoux, "Externalities and Optimal Taxation," *Journal of Public Economics*, December 1998, *70* (3), 343–364.

_ , Pierre Pestieau, and Jean-Charles Rochet, "Capital Income Taxation When Inherited Wealth Is Not Observable," *Journal of Public Economics*, 2003, *87* (11), 2475–2490.

Davies, James B. and Anthony F. Shorrocks, "The Distribution of Wealth," in Anthony B. Atkinson and François Bourguignon, eds., *Handbook of Income Distribution*, Amsterdam; New York: Elsevier, 2000.

De Nardi, Mariacristina, "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, July 2004, *71* (3), 743–768.

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

_ , "Optimal tax treatment of private contributions for public goods with and without warm glow preferences," *Journal of Public Economics*, May 2006, *90* (4-5), 897–919.

Diamond, Peter A. and James A. Mirrlees, "A Model of Social Insurance with Variable Retirement," *Journal of Public Economics*, 1978, *10* (3), 295–336.

Diamond, Peter and Johannes Spinnewijn, "Capital Income Taxes with Heterogeneous Discount Rates," January 2010. MIT and LSE, mimeo.

Doepke, Matthias and Fabrizio Zilibotti, "Occupational Choice and the Spirit of Capitalism.," *Quarterly Journal of Economics*, 2008, *123* (2), 747–793.

Dunn, Thomas A. and John W. Phillips, "The timing and division of parental transfers to children," *Economics Letters*, 1997, *54* (2), 135–137.

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes, "The Importance of Bequests and Life-Cycle Saving in Capital Accumulation: A New Answer," *American Economic Review*, May 2002, *92* (2), 274–78.

_ , _ , and _ , "Do the Rich Save More?," *Journal of Political Economy*, 2004, *112* (2), 397–444.

Eaton, Jonathan and Harvey S. Rosen, "Taxation, Human Capital and Uncertainty," *American Economic Review*, September 1980, *70* (4), 705–715.

Edlund, Lena and Wojciech Kopczuk, "Women, Wealth and Mobility," *American Economic Review*, March 2009, *99* (1), 146–78.

Eliason, Marcus and Henry Ohlsson, "Living to Save Taxes.," *Economics Letters*, 2008, *100* (3), 340–343.

Elinder, Mikael, Oscar Erixson, and Henry Ohlsson, "Carnegie visits Nobel: Do inheritances affect labor and capital income?," Working Paper 2011:5, Uppsala University, Department of Economics 2011.

Eller, Martha, Brian Erard, and Chih-Chin Ho, "The Magnitude and Determinants of Federal Estate Tax Noncompliance." In Gale et al., eds (2001) 375–410.

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

_ **and** _ , "Progressive Estate Taxation," *Quarterly Journal of Economics*, May 2010, *125* (2), 635–73.

Feldstein, Martin S., "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," *Journal of Political Economy*, June 1995, *103* (3), 551–572.

_ , "Tax Avoidance and the Deadweight Loss of the Income Tax," *Review of Economics and Statistics*, November 1999, *4* (81), 674–680.

Francis, Johanna L., "Wealth and the Capitalist Spirit.," *Journal of Macroeconomics*, 2009, *31* (3), 394–408.

Frank, Robert H., "Should public policy respond to positional externalities?," *Journal of Public Economics*, 2008, *92* (8-9), 1777–1786. Special Issue: Happiness and Public Economics.

Friedman, Benjamin M. and Mark Warshawsky, "The Cost of Annuities: Implications for Saving Behavior and Bequests," *Quarterly Journal of Economics*, February 1990, *105* (1), 135–54.

Fullerton, Don, "Environmental Levies and Distortionary Taxes: Comment," *American Economic Review*, March 1997, *87* (1), 245–51.

Gale, William G. and Joel Slemrod, "Rhetoric and Economics in the Estate Tax Debate.," *National Tax Journal*, 2001, *54* (3), 613–627.

_ **and John Karl Scholz**, "Intergenerational Transfers and the Accumulation of Wealth," *Journal of Economic Perspectives*, Fall 1994, *8* (4), 145–160.

_ **, James R. Hines Jr., and Joel Slemrod, eds**, *Rethinking Estate and Gift Taxation*, Brookings Institution Press, 2001.

Gans, Joshua S. and Andrew Leigh, "Did the Death of Australian Inheritance Taxes Affect Deaths?," *B.E. Journal of Economic Analysis and Policy: Topics in Economic Analysis and Policy*, 2006, *6* (1), 1–7.

Golosov, Mikhail, Aleh Tsyvinski, and Matthew Weinzierl, "Preference Heterogeneity and Optimal Commodity Taxation," 2010. Yale and Harvard, mimeo.

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

_ **, Narayana Kocherlakota, and Aleh Tsyvinski**, "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, July 2003, *70* (244), 569–587.

Gordon, Roger and Wojciech Kopczuk, "The Choice of Personal Income Tax Base," August 2010. UC-San Diego and Columbia, preliminary draft, mimeo, `http://www.columbia.edu/~wk2110/bin/choiceBase.pdf`.

Graetz, Michael J. and Ian Shapiro, *Death by a Thousand Cuts: The Fight over Taxing Inherited Wealth*, 1st ed., Princeton University Press, 2005.

Green, Jerry and Eytan Sheshinsky, "Direct versus Indirect Remedies for Externalities," *Journal of Political Economy*, August 1976, *84* (4), 797–808.

Grossmann, Volker and Holger Strulik, "Should continued family firms face lower taxes than other estates?," *Journal of Public Economics*, 2010, *94* (1-2), 87–101.

Holtz-Eakin, Douglas and Donald Marples, "Distortion Costs of Taxing Wealth Accumulation: Income Versus Estate Taxes," Working Paper 8261, National Bureau of Economic Research April 2001.

_ , David Joulfaian, and Harvey S. Rosen, "The Carnegie Conjecture: Some Empirical Evidence.," *Quarterly Journal of Economics*, 1993, *108* (2), 413–435.

_ , _ , and _ , "Entrepreneurial Decisions and Liquidity Constraints.," *RAND Journal of Economics*, 1994, *25* (2), 334–347.

_ , _ , and _ , "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints.," *Journal of Political Economy*, 1994, *102* (1), 53–75.

_ , John W. Phillips, and Harvey S. Rosen, "Estate Taxes, Life Insurance, and Small Business," *Review of Economics and Statistics*, February 2001, *83* (1), 52–63.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes, "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving," *Carnegie-Rochester Conference Series on Public Policy*, June 1994, *40*, 59–125.

_ , _ , and Stephen Zeldes, "Precautionary Saving and Social Insurance," *Journal of Political Economy*, April 1995, *103* (2), 360–399.

Hurd, Michael D., "Savings of the Elderly and Desired Bequests," *American Economic Review*, June 1987, *77* (3), 298–312.

_ , "Mortality Risk and Bequests," *Econometrica*, July 1989, *57* (4), 779–813.

Hurst, Erik and Annamaria Lusardi, "Liquidity Constraints, Household Wealth, and Entrepreneurship.," *Journal of Political Economy*, 2004, *112* (2), 319–347.

Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote, "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review*, 2001, *91* (4), 778–794.

Johnson, Barry W., ed., *Compendium of Federal Estate Tax and Personal Wealth Studies*, Department of Treasury, Internal Revenue Service, Pub. 1773 (4-94), 1994.

_ **, Jacob M. Mikow, and Martha Britton Eller**, "Elements of Federal Estate Taxation." In Gale et al., eds (2001).

Joulfaian, David, "Charitable Bequests and Estate Taxes," *National Tax Journal*, June 1991, *44* (2), 169–180.

_ , "Charitable Bequests and Estate Taxes: Another Look at the Evidence," November 1998. Office of Tax Analysis, U.S. Department of Treasury, mimeo.

_ , "Estate Taxes and Charitable Bequests by the Wealthy," Working Paper 7663, National Bureau of Economic Research 2000.

_ , "Gift Taxes and Lifetime Transfers: Time Series Evidence," *Journal of Public Economics*, August 2004, *88* (9-10), 1917–1929.

_ , "Choosing between gifts and bequests: How taxes affect the timing of wealth transfers," *Journal of Public Economics*, 2005, *89* (11-12), 2069–2091.

_ , "Estate Taxes and Charitable Bequests: Evidence from Two Tax Regimes," Office of Tax Policy Analysis Working Paper 92, U.S. Department of Treasury 2005.

_ , "The Behavioral Response of Wealth Accumulation to Estate Taxation: Time Series Evidence," *National Tax Journal*, June 2006, *59* (2), 253–68.

_ , "Inheritance and Saving," WP 12569, National Bureau of Economic Research October 2006.

_ **and Kathleen McGarry**, "Estate and Gift Tax Incentives and Inter Vivos Giving," *National Tax Journal*, June 2004, *57* (2 (part 2)), 429–444.

_ **and Mark O. Wilhelm**, "Inheritance and Labor Supply.," *Journal of Human Resources*, 1994, *29* (4), 1205–1234.

Judd, Kenneth L., "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, October 1985, *28* (1), 59–83.

Kaplow, Louis, "A Framework for Assessing Estate and Gift Taxation." In Gale et al., eds (2001).

_ , "On the Undesirability of Commodity Taxation Even When Income Taxation is Not Optimal," *Journal of Public Economics*, August 2006, *90* (6-7), 1235–50.

_ , "On the Taxation of Private Transfers," *Tax Law Review*, 2009, *63* (1), 159–188.

Kocherlakota, Narayana R., "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 2005, *73* (5), 1587–1622.

_ , *The New Dynamic Public Finance*, Princeton and Oxford: Princeton University Press, 2010.

Kopczuk, Wojciech, "Optimal Estate Taxation in the Steady State," March 2001. University of Michigan, mimeo.

_ , "A Note on Optimal Taxation in the Presence of Externalities," *Economics Letters*, July 2003, *80* (1), 81–86.

_ , "The Trick is to Live: Is the Estate Tax Social Security for the Rich?," *Journal of Political Economy*, December 2003, *111* (6), 1318–1341.

_ , "Bequest and Tax Planning: Evidence from Estate Tax Returns," *Quarterly Journal of Economics*, November 2007, *122* (4), 1801–1854.

_ , "Economics of estate taxation: A brief review of theory and evidence," *Tax Law Review*, 2009, *63* (1), 139–157.

_ **and Emmanuel Saez**, "Top Wealth Shares in the United States, 1916-2000: Evidence from Estate Tax Returns," *National Tax Journal*, June 2004, *57* (2 (part 2)), 445–488.

_ **and Joel Slemrod**, "The Impact of the Estate Tax on the Wealth Accumulation and Avoidance Behavior of Donors." In Gale et al., eds (2001) 299–343.

_ **and** _ , "Dying to Save Taxes: Evidence from Estate Tax Returns on the Death Elasticity," *Review of Economics and Statistics*, May 2003, *85* (2), 256–265.

_ **and** _ , "Denial of Death and Economic Behavior," *Advances in Theoretical Economics*, 2005, *5* (1), Article 5. `http://www.bepress.com/bejte/advances/vol5/iss1/art5`.

_ **and Joseph Lupton**, "To Leave or Not to Leave: An Empirical Investigation of the Distribution of Bequest Motives," *Review of Economic Studies*, January 2007, *74* (1), 207–235.

Kotlikoff, Laurence J. and Lawrence H. Summers, "The Role of Intergenerational Transfers in Aggregate Capital Accumulation," *Journal of Political Economy*, August 1981, *89* (4), 706–732.

Kuehlwein, Michael, "The non-equalization of true gift and estate tax rates," *Journal of Public Economics*, 1994, *53* (2), 319–323.

Kuhn, Peter, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn, "The Effects of Lottery Prizes on Winners and Their Neighbors: Evidence from the Dutch Postcode Lottery.," *American Economic Review*, August 2011, *101* (5), 2226–2247.

Laferrère, Anne and François-Charles Wolff, "Microeconomic models of family transfers," in Serge-Christophe Kolm and Jean Mercier Ythier, eds., *Handbook on the Economics of Giving, Reciprocity and Altruism*, Vol. 2, Elsevier, 2006, chapter 13, 889–969.

Laitner, John, "Bequests, Gifts and Social Security," *Review of Economic Studies*, April 1988, *55* (2), 275–299.

_ , "Random Earnings Differences, Lifetime Liquidity Constraints and Altruistic Intergenerational Transfers," *Journal of Economic Theory*, December 1992, *58* (2), 135–170.

_ , "Intergenerational and Interhousehold Economic Links," in Mark K. Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, Vol. 1A, Amsterdam; New York: Elsevier/North Holland, 1997.

_ **and F. Thomas Juster**, "New Evidence on Altruism: A Study of TIAA-CREF Retirees," *American Economic Review*, September 1996, *86* (4), 893–908.

_ **and Henry Ohlsson**, "Bequest motives: a comparison of Sweden and the United States," *Journal of Public Economics*, 2001, *79* (1), 205–236. ISPE 1998: Bequests and Wealth Taxation.

Laroque, Guy R., "Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof," *Economics Letters*, April 2005, *87* (1), 141–44.

Light, Audrey and Kathleen McGarry, "Why Parents Play Favorites: Explanations for Unequal Bequests," *American Economic Review*, December 2004, *94* (5), 1669–81.

Luckey, John R., "A History of Federal Estate, Gift and Generation-Skipping Taxes," CRS Report for Congress 95-444, Congressional Research Service January 2008.

McCaffery, Edward J., "The Uneasy Case for Wealth Transfer Taxation," *Yale Law Journal*, November 1994, *104* (2), 283–365.

McCubbin, Janet G., "Improving Wealth Estimates Derived From Estate Tax Data." In Johnson (1994) 363–369.

McGarry, Kathleen, "Inter Vivos Transfers and Intended Bequests," *Journal of Public Economics*, September 1999, *73* (3), 321–51.

_ , "Inter Vivos Transfers or Bequests? Estate Taxes and the Timing of Parental Giving," *Tax Policy and the Economy*, 2000, *14*, 93–121.

_ , "The cost of equality: unequal bequests and tax avoidance," *Journal of Public Economics*, 2001, *79* (1), 179–204.

McNees, Stephen, "Deductibility of Charitable Bequests," *National Tax Journal*, March 1973, *26* (1), 79–98.

Menchik, Paul L., "Primogeniture, Equal Sharing, and the U. S. Distribution of Wealth.," *Quarterly Journal of Economics*, 1980, *94* (2), 299–316.

_ **and Martin David**, "Income Distribution, Lifetime Savings, and Bequests," *American Economic Review*, September 1983, *73* (4), 672–690.

Micheletto, Luca, "Redistribution and Optimal Mixed Taxation in the Presence of Consumption Externalities," *Journal of Public Economics*, October 2008, *92* (10-11), 2262–2274.

Milesi-Ferretti, Gian Maria and Nouriel Roubini, "On the Taxation of Human and Physical Capital in Models of Endogeneous Growth," *Journal of Public Economics*, November 1998, *70* (2), 237–54.

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

Mitchell, Olivia S., James M. Poterba, Mark J. Warshawsky, and Jeffrey R. Brown, "New Evidence on the Money's Worth of Individual Annuities," *American Economic Review*, December 1999, *89* (5), 1299–1317.

Modigliani, Franco, "The Role of Intergenerational Transfers and Life Cycle Saving in the Accumulation of Wealth," *Journal of Economic Perspectives*, Spring 1988, *2* (2), 15–40.

Nordblom, Katarina and Henry Ohlsson, "Tax avoidance and intra-family transfers," *Journal of Public Economics*, 2006, *90* (8-9), 1669–1680.

Ohlsson, Henry, "The legacy of the Swedish gift and inheritance tax, 1884–2004," *European Review of Economic History*, December 2011, *15* (3), 539–69.

Page, Benjamin R, "Bequest taxes, inter vivos gifts, and the bequest motive," *Journal of Public Economics*, 2003, *87* (5-6), 1219–1229.

Pérez-González, Francisco, "Inherited Control and Firm Performance," *American Economic Review*, 2006, *96* (5), 1559–1588.

Perozek, Maria G., "Comment: A Reexamination of the Strategic Bequest Motive.," *Journal of Political Economy*, 1998, *106* (2), 423–445.

Phelan, Christopher, "Opportunity and Social Mobility," *Review of Economic Studies*, April 2006, *72* (2), 487–504.

Pigou, Arthur C., *Economics of Welfare*, London: Macmillan and Co., 1920.

Piketty, Thomas, "Income Inequality in France 1901-98," DP 2876, CEPR 2001.

_ , "On the Long-Run Evolution of Inheritance: France 1820–2050," *Quarterly Journal of Economics*, 2011, *126* (3), 1071–1131.

_ **and Emmanuel Saez**, "How Progressive is the U.S. Federal Tax System? A Historical and International Perspective," *Journal of Economic Perspectives*, Winter 2007, *21* (1), 3–24.

_ **and** _ , "A Theory of Capital Income Taxation," October 2011. Paris School of Economics and UC Berkeley, mimeo.

_ **, Gilles Postel-Vinay, and Jean-Laurent Rosenthal**, "Wealth Concentration in a Developing Economy: Paris and France, 1807-1994," 2003. EHESS and UCLA, mimeo.

Poterba, James M., "The Estate Tax and After-Tax Investment Returns." In Slemrod (2000) 329–349.

_ , "Stock Market Wealth and Consumption," *Journal of Economic Perspectives*, 2000, *14* (2), 99–118.

_ , "Estate and Gift Taxes and Incentives for Inter Vivos Giving in the US," *Journal of Public Economics*, January 2001, *79* (1), 237–64.

_ **and Scott J. Weisbenner**, "Inter-asset Differences in Effective Estate-Tax Burdens," *American Economic Review*, May 2003, *93* (2), 360–365.

_ **and Scott Weisbenner**, "The Distributional Burden of Taxing Estates and Unrealized Capital Gains at the Time of Death." In Gale et al., eds (2001).

Reiter, Michael, "Do the Rich Save Too Much? How to Explain the Top Tail of the Wealth Distribution," 2004. Universitat Pompeu Fabra, mimeo.

Rogerson, William P., "Repeated Moral Hazard," *Econometrica*, 1985, *53* (1), 69–76.

Roine, Jesper and Daniel Waldenström, "Wealth Concentration over the Path of Development: Sweden, 1873-2006," *Scandinavian Journal of Economics*, 2009, *111* (1), 151–187.

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

_ , "The Desirability of Commodity Taxation under Non-Linear Income Taxation and Heterogeneous Tastes," *Journal of Public Economics*, 2002, *83* (2), 217–320.

_ , "Optimal Progressive Capital Income Taxes in the Infinite Horizon Model," Working Paper 9046, National Bureau of Economic Research July 2002.

_ **, Joel B. Slemrod, and Seth H. Giertz**, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 2011. Forthcoming.

Salanié, Bernard, *The Economics of Taxation*, 2nd ed., Cambrige, MA and London, England: The MIT Press, 2011.

Sandmo, Agnar, "Optimal Taxation in the Presence of Externalities," *Swedish Journal of Economics*, 1975, *77* (1), 86–98.

_ , "Redistribution and the Marginal Cost of Public Funds," *Journal of Public Economics*, December 1998, *70* (3), 365–382.

Scheve, Kenneth and David Stasavage, "Democracy, War and Wealth: Evidence from Two Centuries of Inheritance Taxation," 2010. Yale and NYU, mimeo.

Schmalbeck, Richard, "Avoiding Federal Wealth Transfer Taxes." In Gale et al., eds (2001).

Scholz, John Karl, "Wealth Inequality and the Wealth of Cohorts," 2003. University of Wisconsin, mimeo.

_ **, Ananth Seshadri, and Surachai Khitatrakun**, "Are Americans Saving "Optimally" for Retirement?," *Journal of Political Economy*, August 2006, *114* (4), 607–643.

Slemrod, Joel, "Methodological Issues in Measuring and Interpreting Taxable Income Elasticities," *National Tax Journal*, December 1998, *51* (4), 773–788.

_ **and Wojciech Kopczuk**, "The Optimal Elasticity of Taxable Income," *Journal of Public Economics*, April 2002, *84* (1), 91–112.

_ **, ed.**, *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, New York: Harvard University Press and Russell Sage Foundation, 2000.

Stiglitz, Joseph E., "Notes on Estate Taxes, Redistribution, and the Concept of Balanced Growth Path Incidence," *Journal of Political Economy*, April 1978, *86* (2), S137–S150.

_ , "The General Theory of Tax Avoidance," *National Tax Journal*, September 1985, *38* (3), 325–37.

Tsoutsoura, Margarita, "The Effect of Succession Taxes on Family Firm Investment: Evidence from a Natural Experiment," 2011. Chicago Booth, mimeo.

Villalonga, Belen and Raphael Amit, "How do family ownership, control and management affect firm value?," *Journal of Financial Economics*, 2006, *80* (2), 385–417.

Weber, Max, *The Protestant ethic and the spirit of capitalism*, New York: Scribner, 1958. Translated by Talcott Parsons; with a forword by R.H. Tawney.

Werning, Iván, "Nonlinear Capital Taxation," September 2011. MIT, mimeo.

Wilhelm, Mark O., "Bequest Behavior and the Effect of Heirs' Earnings: Testing the Altruistic Model of Bequests.," *American Economic Review*, 1996, *86* (4), 874–892.

Wolff, Edward N., "Discussant's Comments on Douglas Holtz-Eakin, 'The Uneasy Case for Abolishing the Estate Tax'," *Tax Law Review*, 1996, *51* (3), 517–22.

Yaari, Menahem E., "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies*, April 1965, *32* (2), 137–150.