Taxation of Intergenerational Transfers and Wealth

Wojciech Kopczuk

Department of Economics and School of International and Public Affairs and National Bureau of Economic Research, Columbia University

Contents

1. INTRODUCTION

The objective of this chapter is to provide an introduction to and review of economic literature related to taxation of transfers and wealth. As will become clear in what follows, the focus will be primarily on taxes imposed on intergenerational transfers. Such taxes take many different forms. Transfers that occur at death may be taxed in the form of estate taxation, i.e., the tax may be imposed on the total amount of wealth left by the decedent. They may take the form of an inheritance tax, in which case the base is defined on the level of the donee, and reflects the transfers to that particular individual.[1] If taxes were imposed only at death, the simplest form of avoidance would be to transfer resources inter vivos (during lifetime). Hence transfer taxation systems usually include a tax on gifts as well.[2] Taxes on estates are a form of a tax on wealth and, although it is rare, some countries (e.g., France and Norway) impose annual taxes of that kind.

Transfer and wealth taxes have unique features that make them different than other types of direct taxation. First of all, they inherently affect two-related parties so that the distortion affects a transaction (transfer) that is not arms-length and that may involve externalities. This makes understanding and assumptions about transfer motives central to both positive and normative analysis. Second, they are infrequent (at the extreme, occurring just at death), thereby allowing for a long period of planning, making expectations about future tax policy critical and empirical identification of the effect of incentives particularly hard. Potentially large amounts of money at stake make investment in tax avoidance worthwhile in the presence of fixed costs and incentives to do so are potentially quite salient. Third, in practice, this type of taxation often applies only to a small but important group of individuals—those with sufficiently high wealth at the time of taxation—and thus plays a potentially important role in overall distributional implications of taxation, both in the short and in the long term.

There are many dimensions of differences in the implementation of 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, details of interaction with capital gains taxation regime may vary. Many features of implementation of the tax may matter greatly—some examples are valuation rules, preferences for particular types of assets, treatment of transfers shortly before death, treatment of marital assets, implications of

---

[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 US transfer tax system along these lines.

[2] The actual implementation does vary though. For example, in the US, the estate tax operated without a gift tax from 1916 until 1932, and since then gifts made within three years of death may be subject to different treatment than those made earlier (if made in "contemplation of death"; they were automatically included in estate between 1976 and 1981, see Luckey (2008) for the history of provisions). In the UK, there is no tax on gifts made more than seven years before death (though there was one between 1975 and 1984, see Boadway, Chamberlain, and Emmerson (2010b) for the discussion of changes).

joint vs community property, treatment of charity, and treatment of transfers that skip generations. The purpose of this chapter is not to discuss all of these issues, although in Section 2, I will provide a short overview of history of the estate tax in the United States and international differences in transfer tax systems. Instead, my objective is to focus on the economics of transfer taxation and to offer a critical review 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 US will be present. However, the focus of the chapter is on taxation of transfers in general, with the US 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 is discussed elsewhere in this Handbook).

This is not the first survey of literature related to transfers and their taxation. Gale and Slemrod (2001) and Boadway, Chamberlain, and Emmerson (2010a) provide useful background to the issues surrounding the design of transfer taxation. Cremer and Pestieau (2006, chap. 16) discuss some theoretical contributions to the literature on taxation of bequests. Laitner (1997), Laferrère and Wolff (2006, chap. 13), and Arrondel and Masson (2006, chap. 14) 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) and Joulfaian (2011) provide excellent overviews of the history of estate tax legislations in the United States.

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

I will begin in the next section with a brief 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 discuss evidence on bequest motives and basic normative implications of different motives for thinking about taxation. Bequest motives are the key building block for theoretical analysis of taxation of transfers, but the empirical literature has not settled on a clear answer to the question about the nature of bequest motivations. I emphasize heterogeneity of two different kinds. First, search for *the* bequest motive is unlikely to be fruitful—saving plays dual role of protecting against lifetime risk and increasing 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 somewhat agnostic (general) about the nature of the bequest motive or it should explicitly account for heterogeneity.

In Section 4, I focus on the main theoretical framework for analyzing bequests taxation that builds on Mirrlees (1971) and Atkinson and Stiglitz (1976) approach to taxation of commodities in the presence of non-linear income taxation, adapted to transfer context by Kaplow (1998, 2001). This approach of course incorporates redistributive preferences for a policy maker. 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 besides 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 subsidizing 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: that is, they are irrelevant precisely where taxation of transfers is important in practice. Furthermore, I point out that theoretical 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 to review empirical evidence on 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 and 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 fit elsewhere. I elaborate further on the relationship of estate taxation and the shape of wealth distribution, discuss empirical work related to charitable bequests, and briefly comment on research on marital bequests, tax competition, and political economy of this type of taxation. The final section concludes.

The main message 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, and especially so when 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. Finally, potential negative externalities from wealth accumulation and concentration are yet to be seriously addressed.

2. OVERVIEW OF WEALTH AND ESTATE TAXATION IN PRACTICE

Taxation at death is administratively convenient: this is the time when assets change hands and need to be valued anyway, thereby increasing tax administration's ability to observe them. Wealth is arguably easier to observe than income, and taxes on wealth—most importantly on land but also on successions—have historically preceded taxation of income. Following the development of modern forms of income and consumption taxation, estate and wealth taxes are rarely a major source of revenue. However, they remain an unusually progressive component of the tax code. Limited revenue makes elimination of such taxes a realistic policy proposal and, indeed, a number of developed countries (for example, Canada, Sweden, and Australia and some states in the United States) repealed such taxes. Large financial stakes for a limited number of wealthy taxpayers make politics surrounding these issues contentious and subject to massive lobbying (Graetz & Shapiro, 2005).

Luckey (2008) and Joulfaian (2011) provide excellent and detailed overviews of the history of estate and inheritance taxation in the United States. A reader interested in state provisions should consult Bakija (2007) and sources referenced there. Gale and Slemrod (2001), Boadway et al. (2010b), and Scheve and Stasavage (2012) provide international comparison of transfer tax systems and rates for a few developed countries. An older discussion by Bird (1991) provides a comparison of inheritance taxes and annual wealth taxation for both developed and developing countries. Able reviews of economic aspects of inheritance tax systems exist for a number of other countries—for example, Boadway et al. (2010a, chap. 8) present detailed overview of inheritance taxation in the UK, Ohlsson (2011) does so for Sweden and Piketty (2011) describes the French system.

The United States instituted a number of short-lived inheritance or estates taxes[3] in the eighteenth and nineteenth century: the Death Stamp Tax was in place between 1789 and 1802, an inheritance tax during the Civil War, and an estate tax during the Spanish-American War.[4] A number of states had estate or inheritance taxes before the turn of the 20th century. The modern estate tax was enacted in 1916 with initial rates ranging from 1% to 10% and exemption of $50,000. The rates changed often before the top rate reached its peak of 77% by 1941, where it stayed until 1977. The estate tax was supplemented by a cumulative gift tax starting in 1932[5] in an effort to combat tax avoidance.[6]

The base of the tax was subject to numerous changes over the years. Tax treatment of marital transfers used to be a controversial issue in the early years of the estate tax, in

---

[3] This summary of the US provisions follows Luckey (2008) and Joulfaian (2011).

[4] See Scheve and Stasavage (2012) for exploration of the potential link between wars and transfer taxation.

[5] A short-lived and retroactively repealed gift tax was in place between 1924 and 1926.

[6] Nominal gift rates were set lower than estate tax rates apparently in deliberate effort to accelerate revenue flow (Joulfaian, 2011). This practice continues to date: even though tax rates are nominally the same, gift and estate taxes are calculated differently—gifts are taxed on a tax exclusive basis while estate tax applies to tax-inclusive base, resulting in higher effective marginal tax rates for estates than for gifts (see footnote 39).

part due to different marital property regimes in place in different states: in community property states half of property acquired during marriage belongs to each spouse and thus was automatically excluded from the estate, while in the remaining states such property originally could be excluded only to the extent that the surviving spouse could be shown to have contributed to its acquisition. This was perceived as inequitable and in 1942 the Congress attempted to address it by extending non-community property treatment to the community property states; this solution was considered complex and was replaced by marital deduction in 1948. Initially, half of adjusted gross estate was eligible for marital deduction (but community property was not); this was modified to allow for the marital deduction equal to the adjusted gross estate or $250,000 (whichever was greater) in 1977, and then unlimited marital deduction as of 1982.

Until 1976, gift and estate taxes were separate. The Tax Reform Act of 1976 introduced a uniform tax with a single tax schedule applying to the sum of cumulative lifetime taxable gifts and the estate; allowing, in particular, for a *unified* estate and gift tax credit (equivalent to a lifetime exemption). The same act also introduced a Generation Skipping Transfer Tax that imposes an additional layer of tax liability for certain transfers to grandchildren and related transactions. The effective exemption was increased at the same time from $60,000 (where it was since 1942) to $120,667 stopping the process of inflation-driven "bracket creep" that pushed the number of taxable tax returns to its all-time maximum of 7.65% of all deaths by 1976 (in contrast, the tax affected fewer than 2% of decedents in other periods). At the same time, the top tax rate was reduced to 70%. The additional rate reductions and increases in exemption level were enacted in 1981 and phased in over the number of following years, bringing the top rate to 55% and the exemption to $600,000 by 1987 where it stayed for more than a decade.

A number of other changes took place in the 1970s, 1980s, and 1990s. The 1976 reform attempted to introduce a carryover basis for capital gains. Capital gains that are not realized at the time of death are subject to the estate tax but escape capital gains taxation due to step-up provision that modifies the tax basis of the asset to its value at the time of death of the original owner. The carryover approach was intended to eliminate such resetting of the basis so that capital gains tax would still be due on full accumulation when a beneficiary ultimately sells the asset. This provision had, however, been delayed first and then repealed before it became effective. The same reform also changed treatment of gifts shortly before death by requiring that any gifts within three years of death be included in the estate (rather than relying on gifts in "contemplation of death"). During the 1980s, a number of changes benefiting closely-held businesses and farms were introduced, including alternate valuation rules, installment payments, and special rules applying to qualified family businesses.

In 2001, the estate tax was "repealed"—the Economic Growth and Tax Relief Reconciliation Act of 2001 enacted a 10-year schedule of reductions in rates and increases in exemption of the tax that was supposed to culminate in a repeal in 2010. The whole

legislation was scheduled to sunset at the end of 2010 so that, absent additional action, the repeal would last one year only, at which point the schedule would return to their 2001 shape. Few, if any, observers expected this sequence of events to play itself out and yet this is almost exactly what happened: for most of 2010 there was no estate tax on the books, for the first time since 1916. At the end of 2010, the Tax Relief, Unemployment Insurance Reauthorization, and Job Creation Act of 2010 reinstated the estate tax in a modified form. While this provision was retroactive, estates for decedents who died in 2010 had an option of electing out of the estate tax. Such decision was not without trade-offs—assets passed in that way will not benefit from the step-up in basis at death and instead will effectively be subject to carryover basis treatment for capital gains purposes—but it is expected that few estates will choose to pay the estate tax in 2010 (data is not available yet). The provisions enacted in 2010 were temporary and applied to 2011 and 2012. They were made permanent for 2013 and beyond as part of the Taxpayer Relief Act of 2012 passed on January 1st 2013.

Even prior to the"repeal," changes in the 2000s were substantial—the exemption increased from $675,000 in 2001 to $3.5 million in 2009 and the maximum rate fell from 55% to 45%. The reinstated tax features an even larger exemption ($5 million) and lower top rate (35%). The implications of the 2010 repeal are yet to be studied (or even observed given the lags involved in filing and publication of statistics). Over that period, the number of estate tax returns filed (the date-of-death numbers are not yet available for most recent years) declined from over 108,000 in 2001 to 15,000 in 2010; and the number of taxable returns fell from over 50,000 to 6700; however the reduction in net estate tax liability was much smaller than these numbers and tax reductions might suggest: it fell from $23 billion to $13 billion, primarily reflecting skewness of the wealth distribution.[7] Changes in rates that applied to estates, applied also to gifts. The sole exception was the 2010 repeal: the gift tax remained in place throughout that year. As of 2013, the marginal tax rate is set at 40% and the exemption, $5.25 million in 2013, will be indexed for inflation.

There are many differences in how transfer taxation could be implemented in practice. The United States federal tax is a tax on estates—the overall value of wealth left by a decedent. Somewhat similar structure is present in the UK and other English-speaking countries (or was present before the tax was repealed—as in the case of Canada). An alternative approach—in place in some US states (for example, Indiana, Maryland, Pennsylvania, New Jersey) and in many continental European countries—is to impose a tax on inheritances, that is to shift the base to the recipient. This has two noteworthy implications: the overall tax liability depends on how estate is distributed and tax payments may be made dependent on the relationship to the donee in a straightforward way. Indeed, it is common for inheritance tax systems to vary tax liability depending on the relationship, usually with

[7] IRS Statistics of Income Division, Tax Stats, available online at http://www.irs.gov/taxstats/indtaxstats/article/0,,id=210646,00.html, accessed on 3/20/2012.

preferences for close relatives.[8] Yet another possibility—rarely in place in practice—is to integrate transfer tax within income tax system so that inheritances increase income tax base.[9]

Taxation of transfers is not the only way of imposing a general tax on wealth. Mintz (1991) and Auerbach (2008) discuss many different forms such taxes can take. Obviously, many types of taxation of income from wealth (interests, capital gains, dividends, etc.) are common. Taxation of particular forms of wealth—property, land, certain durable goods—is also often in place. Taxation of capital income is discussed in a separate chapter of the Handbook and taxes imposed on particular goods raise many interesting issues related to these specific markets but are not the focus of this chapter. Some countries—for example, France and Norway—also impose a net worth tax (usually annual). Brown (1991) carefully analyzes administrative difficulties in implementing an annual net wealth tax and concludes that practical problems are intractable. Adam et al. (2011, chap. 14) assert that practical difficulties and actual experience of the implementation in countries that tried to do it have been discouraging. Defining base in a comprehensive manner is difficult—it requires costly and sometimes impractical valuation, especially in the case of business assets. Naturally, taxation of "human capital" in this way is not a realistic option. Deviating from the comprehensive basis, as is done in practice, allows for tax avoidance opportunities. A regular tax on wealth is closely related to a tax on capital income. As Adam et al. (2011, chap. 14) point out though, net worth tax predominantly taxes the normal rate of return rather than excess return. In particular, under wealth taxation exempting the normal rate of return is not a realistic option, thereby making it a less desirable form of taxation than direct taxes on capital income.[10] The main advantage of this form of taxation over capital income taxation is its ability to tax assets that do not generate income (in particular, due to tax avoidance). Transfer taxation, on which this chapter mostly focuses, retains this property and takes advantage of the administratively convenient (even if unpopular) timing—death—when assets change hands anyway so that they are more easily observable and when difficult valuation issues often would need to be addressed anyway, thereby easing tax compliance burden.

---

[8] Estate tax systems do allow for some adjustments to the overall tax liability depending on the recipient through deductions—for example, in the US transfers to a spouse or to a charity are fully deductible. In principle, one could also imagine introducing variation in rates through tax credits for gifts to specific recipients; naturally though inheritance taxation makes such adjustments much less complex.

[9] See Batchelder (2009) for an extensive discussion of this possibility. In particular, she reports that a few countries (Denmark, Iceland, Lithuania, and Russia) include gifts in the income tax base. A short-lived (struck down as unconstitutional) 1894 the US income and inheritance tax was supposed to include inheritances as part of the income tax base.

[10] Using Adam et al. (2011, chap. 14) example, assume that the normal rate of return is 5%. A 20% capital income tax is equivalent to upfront 1% wealth tax. Both systems discourage saving, but under capital income tax excess return is taxed at 20%, while wealth tax exempts it. Shifting the net worth tax in this example to the second period makes only minor difference by taxing excess returns lightly (at 1% if keeping the present value of revenue constant). Exempting the normal rate of return under net worth tax would effectively wipe out tax liability, while doing so under capital income taxation is conceivable. Such an approach (rate-of-return allowance, RRA) has in fact been implemented in Norway.

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 with 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 estate tax rate by t, the budget constraint is $(1 - t)C + RB = (1 - t)y$ or, equivalently, $C + \frac{R}{1-t}B = y$. The estate tax increases relative price of bequests and stimulates negative substitution response and further reduction of bequests via income effect (unless bequests are an inferior good). Imposing a tax on the basis of the amount received by the beneficiary (as is done in the case of gift taxation in the US) 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 here 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 linear estate-tax formulation, $B = (1 - t)W/R$. The simplest wealth-in-utility formulation $u(C, W)$ is observationally 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. Hence, these two very similar formulations—joy-of-giving and wealth-in-utility—have very different implications when considering responses to and efficiency cost of taxation.

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; Francis, 2009; Reiter, 2004), the topic to which we will return in Section 7.1. Some possible justifications for considering wealth-in-utility are 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 (e.g. due to positional externalities as in Frank, 2008).

An alternative model of bequests that do not yield utility to the donor is the "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 & Warshawsky, 1990; Mitchell, Poterba, Warshawsky, & Brown, 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 (Kopczuk, 2003b). 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 underlying imperfections in the market for annuities. A taxpayer would clearly be better off by selling the right to (unintended) bequest conditional on dying at some time *t* (that has no value to him) and using proceeds for consumption in 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 T_i conditional on dying in period i, surviving from period i to period $i + 1$ implies a reduction of lifetime tax liability by $T_i - \frac{T_{i+1}}{1+r}$. 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.[11]

---

[11] 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.

3.2. Intergenerational Links

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 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 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

$$\gamma^P + R\gamma^K = C^P + RC^K, \tag{1}$$

where γ^P is income of parents and γ^K is income of children. In this formulation, bequests are equal to the unconsumed resources of the parent $\gamma^P - C^P$.[12] 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, Fumio, & Kotlikoff, 1992, 1997; Laitner & Ohlsson, 2001; Wilhelm, 1996) 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 (Light & McGarry, 2004; Menchik, 1980; Menchik & David, 1983).[13]

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

---

[12] 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.

[13] 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. Severinov (2006) shows that this model can generate transfer patterns that are very similar to those under standard altruism but does not imply Ricardian equivalence. Dunn and Phillips (1997) and McGarry (1999) provide evidence that (presumably harder to observe) inter vivos gifts are compensatory while bequests are split equally.

$\max_{B,C^P} v^P(C^P) + \rho u^K(y^K + B)$ 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 relationship that they establish is not invariant to changes in taxation).

An alternative approach to bequests treats them as 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, Shleifer, & Summers, 1985; Cox, 1987; Perozek, 1998).[14] As with altruism, evidence in support of the exchange motive is mixed, see Arrondel and Masson (2006, chap. 14) and Laferrère and Wolff (2006, chap. 13) 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, Skinner, & Zeldes, 2002, 2004); a person may also have a mix of altruistic and 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% of people with children consider bequests important relative to 21% of the 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 (joy-of-giving) bequests and tests for the presence of a bequest motive by assuming that people with kids have one and those without them do not. Given similarity of wealth profiles for those with and without children, he rejects that a bequest motive is present. Noting evidence in Laitner and Juster (1996) suggesting that children are unlikely to be a deterministic indicator of the presence of a bequest motive, Kopczuk and Lupton (2007) extend this approach to exploit longitudinal information and allow kids to be just one of the many potential (but non-deterministic) indicators of the presence of the motive in a switching

---

[14] 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 & Waldman, 1991; Coate, 1995).

regression framework with unknown sample separation. They conclude that bequest motive is present (in fact, their parameters indicate that 3/4 of the population has one), but only weakly related to having children. Ameriks, Caplin, Laufer, and Van Nieuwerburgh (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 a luxury good. The results indicate that intentional bequest motive is not operational for much of the distribution where accidental bequest dominates. However, both type of motives become quantitatively important at higher levels of wealth.

Using estate tax data for filers in 1977, 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, chap. 14) advocate a mix of altruistic and strategic motives. The literature on the determinants of savings and wealth distribution grappled with this question as well (it will be discussed in more detail in Section 7.1). A strand of this literature assumes away the presence of a bequest motive (Hubbard, Skinner, & Zeldes, 1994, 1995; Scholz, Seshadri, & Khitatrakun, 2006), but it has problems explaining the very top of the wealth distribution. Adding an explicit bequest motive helps (Cagetti & De Nardi, 2008; De Nardi, 2004; 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 (though, as mentioned before, the choice between the two is not innocuous when considering implications of tax policy).

3.3. Normative Issues

The lack of consensus about the nature of bequest motives 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.[15]

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

---

[15] Kopczuk (2001) and Cremer and Pestieau (2006, chap. 16) analyze bequest taxation using models that have different types of bequest motives as special cases.

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 justification of 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 (Kaplow, 1995, 1998). A dollar of bequests provides utility to both parents and children. From the social point of view, the benefit of a bequest is given by $u_B^P + \rho u_B^K$, but when maximizing her own utility the parent is only taking into account her own marginal benefit u_B^P and ignores the ρu_B^K 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 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 $\beta = 0$, 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.[16]

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

---

[16] 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.

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 just as preferences for any other good. 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 f 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 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 need to be differentiated. This is not the case with bequest externality: the externality is interpersonal and, with sufficient heterogeneity, marginal social welfare 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, externality due to giving 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 type of a bequest motive influences the presence and nature of this externality. Still, a normative choice may 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 its complete elimination.

4. REDISTRIBUTION

In the previous section, I abstracted from redistributive motives. Taxation of estates in practice is about closely connected with redistribution. For example, according to the Piketty and Saez (2007) assessment of the overall progressivity of the US tax code in 1970,[17] 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 taxes that apply predominantly to those with high net worth and have 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) and 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

---

[17] The study assigns (cross-sectional) realized tax liability to groups defined by their gross income reported on tax returns. In the case of the estate tax, they assign the current tax liability for a given group of decedents (e.g., top 0.01% ranked by the size of estate) to the corresponding group of living taxpayers (top 0.01%, ranked by gross income). There are many limitations to this procedure—in particular, it assumes that top income and top estate taxpayers are the same group, it assumes that the current law will continue and it makes no attempt to adjust for the expected differences in terminal wealth of the decedents and that of current high income taxpayers.

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 $u(v(C, B), L)$.[18] 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 the price of bequests does not provide any additional information about wages beyond that already contained in income. Hence, such a distortion is redundant.

There are many limitations of this exercise of course, some of which we will cover below (see also Kaplow, 2001, for an extensive discussion), 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 (Diamond & Spinnewijn, 2010; Golosov, Tsyvinski, & Weinzierl, 2010; Saez, 2002b) and shows that uniform tax on capital income may be desirable even under the weak separability assumption if higher ability individuals have a lower discount rate,[19,20] while the tax on savings of just high ability individuals may be optimal under

---

[18] Laroque (2005) and Kaplow (2006) show that commodity taxation is redundant even when income tax is not optimally selected. Kaplow (2008) builds on the Atkinson-Stiglitz framework to analyze a wide range of tax policy issues.

[19] Banks and Diamond (2010, chap. 6) 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 empirical support for it.

[20] 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, Pestieau, and Rochet (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.

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) may be written without loss of generality as

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

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 at the optimum, the point previously made by Kaplow (2001).

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

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

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

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

with $\underline{V}$ indexing the problem. When $\underline{V}$ 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 $\int u_p(C^P) - v(L) + \rho u_K(C^K) + v u_K(C^K)$ as in Eq. (5) where (with some abuse of the notation) v 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

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

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 v on the constraint (7), the planner's objective is

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

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 $\left(\frac{W'_2}{W'_1}v + \rho\right) R u'_K = u'_P$ 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 if 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 (Cremer, Gahvari, & Ladoux, 1998; Kopczuk, 2003a; Sandmo, 1975) is straightforward. The problem with an externality is equivalent to the one without any externalities, but with (1) the price of a dirty good adjusted (via the marginal tax rate τ) to internalize the otherwise ignored social cost of increasing $\int D$, and (2) 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 through the tax on the dirty good.[21]

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 marginal 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.[22]

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 of individuals $\tilde{D}$, so that it's given by $g\left(\int_{\tilde{D}} D\right)$, 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 implications. Furthermore, multiple externalities need not be a problem if each of them can be targeted independently (or, if they are linked in a way that allows restricted instruments to work as well as independent targeting).[23]

---

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

[22] 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 $\int \frac{\partial u}{\partial G} = \mu \left(1 - \int T' \frac{\partial y}{\partial G}\right)$ 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), 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. Consequently, the optimal way of collecting revenue on the margin relies on a mix of lump-sum tax and distortionary taxation.

[23] 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.

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 $v W_1 \left(u_K(C^K) \right)$ 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 $-R \frac{v W_1' u_K'}{\mu}$ (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 such a case, 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 $t^E = -R \frac{v W_1' u_K'}{\mu}$ 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):

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

(obviously, R, μ, and v 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:

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

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 that implies 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 Eq. (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. Effectively, the role of the estate tax is to facilitate redistribution across generations within the same dynasty, 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 that 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 is 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 and hence non-negligible.

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 v on the next generation can be expressed as $u(C^P, L) + v(C^K) + vu(C^K)$: 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) and Piketty and Saez (2012a,2012b) consider a different extension of the Mirrlees model by allowing for both bequest decisions and heterogeneity in inheritance received by those who make bequests themselves and imposing the steady-state restriction that the distribution of bequests received should be the same as that of bequests left. Kopczuk (2001) makes the simplifying assumption of perfect correlation in skills across generation, which implies that on the individual level bequest received and left should be the same.

There are important limitations of such an approach. In particular, it ignores the issue of whether an economy can converge to the "optimal" allocation from an arbitrary initial point. Very importantly, it also ignores welfare implications of the transition by placing no weight on welfare of generations during the transition. For many questions such as, for example, analyzing consumption vs income tax or pay-as-you-go Social Security systems, the impact on the initial generation and transitional issues are of foremost importance and the approach along these lines is not suitable. The focus on the long-term welfare is inherent to this type of steady-state analysis, and simplifications that it introduces help in clarifying implications of the dual role of inheritance (it is both given and received). Piketty and Saez (2012a) argue that this kind of approach is more suitable for normative analysis. In part, this is so by making it possible to abstract from the ability of government to go after the initial wealth that plays an important role in the Judd–Chamley framework. Focusing on the steady state makes it also more straightforward to express the results in terms of empirically observable parameters.

The main result of Kopczuk (2001) may 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 $X + wL - T(B, wL) = C + RB$.[24] 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 may be expressed as $u(D + X) + v(L) + g(B)$ subject to $wL - T(B, wL) = D + RB$. In this (equivalent to

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

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: its strength depends on the level of consumption. Absent externality, labor income tax would be sufficient here, because the Atkinson–Stiglitz separability assumption holds. 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, X also interacts with incentive constraints.[25] It turns out that it is important: the optimal formula for estate taxation includes three additive components: an aggregate term,[26] 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 two recent papers, Piketty and Saez (2012a, 2012b) follow the approach of focusing on the welfare of a steady state generation to characterizing the optimal Mirleesian policy and further 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

---

[25] 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, the externality from giving 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: $-\frac{RE[u'(X+D)|X]}{\mu}$ and does not affect incentive constraints either.

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

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. In Piketty and Saez (2012b), they consider multiple extensions that show robustness of their conclusions to relaxing the focus on steady state.

In an attempt to isolate the key mechanism behind optimality of estate taxes in these frameworks, Kopczuk (2013) considers a model with two generations and the joy-of-giving bequest motive. He explicitly derives the optimal tax formula. He shows that beyond correction of the externality from giving, the optimal formula includes an additional term that reflects fiscal externality due to the income effect of bequests on effort of children. He argues that this is reflecting the key trade-off in this context: between encouraging generosity and reducing inequality in the subsequent generations. The sign of the optimal tax in this case is theoretically ambiguous due to this trade-off but the rates are likely to be positive toward the top of the distribution where the externality from giving is no longer of interest.

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 allow 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.[27] 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-sum tax on the second generation results in higher aggregate labor supply, the 100% tax need not be optimal. Note that the bequests are 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 non-linear income tax context.

---

[27] This result is shown assuming logarithmic preferences.

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.[28] 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 & 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, Chari, and Kehoe (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:

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

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

---

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

In the ergodic steady state, the distribution of consumption in each period is the same. Suppose then that we integrate Eq. (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 $u'(C_i) = E[u'(C_{i+1})]$ implies $u'(C_{i+1}) = u'(C_i) + \varepsilon_i$ where $\varepsilon_{i+1} \perp u'(C_i)$ and $E[\varepsilon_{i+1}^2] > 0$, so that $var(u'(C_{i+1})) > var(u'(C_i))$ which violates the ergodicity assumption. As Atkeson and Lucas (1992) demonstrate, efficiency in fact requires immiseration so that inequality is ever increasing and $E[u'(C_i)]$ grows over time. By allowing for liquidity constraints, Aiyagari (1995) breaks the Euler equation for some individuals:

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

with strict 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 motive 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 (2002a) introduces heterogeneity in the Ramsey model in 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 US 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.[29] 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 tax treatment of initial resources, government commitment and its ability to save play an important role.

The new dynamic public finance literature initiated by Golosov, Kocherlakota, and Tsyvinski (2003) seeks to remedy the first concern by embedding dynamic capital income taxation questions in the Mirrleesian framework. It 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

---

[29] 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.

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 & 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 (e)

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

and that the overall resource constraint is given by $(1 + r)C_0 + E[C_1(\theta)] = E[\theta e(\theta)]$, where the rate of return is given by r and neither effort e nor θ (state of the world, interpreted as skills) are observable (though their product is). The person consumes in the two periods and provides effort in the second period, however the return to effort is not observable. The objective of policy is to provide insurance in the second period, while accounting for disincentive effects.

The new dynamic public finance adopts the mechanism design approach to characterizing Pareto efficient allocations by proposing schedules $C_1(\theta)$ and $\gamma(\theta) = \theta e(\theta)$. Suppose that an individual chooses to report its type as θ'. A change in reported type from θ' to θ'' results in a change in utility from consumption by $u(C_1(\theta'')) - u(C_1(\theta'))$. The key result can be derived by considering the following variation to the optimal policy. 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, $u(\tilde{C}_1(\theta)) \stackrel{\text{df}}{=} u(C_1(\theta)) + \varepsilon$, 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: the gain in utility from consumption due to a change in report is unaffected and hence there is no reason to change the original level of effort ($\gamma(\theta')/\theta$, where θ is the true type).

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. Expressed explicitly in consumption terms, the necessary 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

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

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 Eq. (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 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.[30] 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 (note though that this is a different mechanism that considered by Kopczuk (2013), where uncertainty is not present but instead incentives of parents and children are not aligned)—it seems, though it 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 5.4.

---

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

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, Eq. (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, 2009). 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. Uncertainty adds richness of interactions, 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 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, chap. 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 capital or wealth distortions that either explicitly or on average fall with the current income/productivity.[31] As the result, this line of work

---

[31] 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.

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 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 of the optimal allocation 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 Eq. (10) thereby preserving their conclusions about negativity and "progressivity."

5. BEHAVIORAL RESPONSES TO TRANSFER TAXATION

Empirical work on effects of estate taxation on taxpayers' decisions is marred with practical 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,[32] makes it difficult to

---

[32] 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.

credibly establish the causal 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. Surveys that oversample high net worth population (such as Survey of Consumer Finance), probate records and administrative datasets are potential 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 a 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, Slemrod, & Giertz, 2012) focused on the "sufficient statistic" for behavioral response—the responsiveness of taxable income. A 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 margins of response such as, for example, tax avoidance. 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 effect 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 a change in the tax rate on the size of estate: $T'(E) \cdot dE + dt \cdot (E - \bar{E})$. In either case, it is the level and responsiveness of the overall estate rather than the composition of the 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 Feldstein's argument requires understanding the effect on *overall* tax liability. Even within the simplest framework, it calls for estimating the effect not just on estates but also on gifts, as well as that on another main source of tax revenue—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 natural 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. Response of transfers also naturally has implications for behavior and tax liabilities of 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, 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 & 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 magnitude of distortions that are caused by such 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 an annual wealth tax, t^k is a tax imposed on return to savings, r is in the interest rate and S_k is net new saving as of the end of period k. I am going to assume that $A_i \geq 0$ 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 Eq. (16), yields

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

Assuming first no saving, $S_i = 0$ for all i, the tax on estate is equivalent to an annual wealth tax given by $1 - t^w = (1 - t)^{1/(k+1)}$. 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 (resulting in the same bequest holding behavior constant) 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 (as measured by equivalent tax rates) 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, but it does not appropriately describe the distortion induced by the tax. As one example, 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 reliance on marital and other types of deductions, valuation discounts, interaction with state taxation and deferred payment schedule among other things.[33] We will delay talking about such complications until reviewing related empirical papers.

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

Poterba (2000b) 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 $(1 - m_i)(1 + (1 - t^k)r) + m_i(1 - t)(1 + (1 - t^k)r) = (1 - m_i t)(1 + (1 - t^k)r)$: on the margin, the estate tax is equivalent to a wealth tax at the rate of $m_i t$. Equivalently, the rate of return is $1 + (1 - t^k)r - m_i t(1 + (1 - t^k)r)$. Closer inspection of this formula reveals that the effect of the estate tax on the marginal rate of return increases with mortality rate, but

---

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

otherwise is not very sensitive to the rate of return because $1 + (1 - t^k)r \approx 1$. 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 the purpose of leaving a bequest. In that case, as the first pass, mortality risk does not matter. To see it, consider period N that is sufficiently distant to guarantee that a taxpayer dies by then. The future value of a marginal dollar of current 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 before 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 assumptions are 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 many margins that may 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 US (and other countries') tax law. Marginal tax rates applying to gifts and estates are usually different. Denote the gift tax rate by t^G and continue to denote the estate tax rate as t. For simplicity of the exposition, I assume here that they are applied to the same base—the total amount of the transfer. 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 should select the mode of transfer that corresponds to the lower marginal tax rate. In practice, usually $t^G < t$,[34] 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.[35] Denoting by G is the value of the transfer considered, the value of tax at period n is $(1 + r)^n t^G G$ or $(1 + r)^n t G$ so that the comparison is unchanged.

However, in practice, the tax schedule is often non-linear. Using T and T^G to represent the two tax schedules, the correct comparison of net ex-post tax liabilities is between $T\left(B^E\right) + (1 + r)^n T^G \left(B^G + G\right)$ and $T\left(B^E + (1 + r)^n G\right) + (1 + r)^n T^G \left(B^G\right)$ where B^E and B^G are bases for gift and estate tax, respectively. A 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. The effect will depend on the relationship between the two schedules, but in practice it is likely to make gift taxation further preferred when

---

[34] In the US, gifts are taxed on tax exclusive basis that implies lower marginal tax rate; see footnote 39.

[35] This assumption is not necessarily correct—it is possible that the donor may have access to different investment opportunities than the donee.

the two schedules are closely related and the nominal rate of return is high or the horizon is long. In particular, in the US case, B^E and B^G are (imperfectly) tied to each other via unified credit (lifetime exemption). Ignoring for simplicity the difference in how gift and estate tax is calculated and writing the tax as a single schedule, on the margin the taxpayer should be comparing $(1 + r)^n T' \left(B^E + B^G + G \right)$ in the case when an inter vivos gift is made to $(1 + r)^n T' \left(B^E + B^G + (1 + r)^n G \right)$ when it is delayed until death, providing the strong case for pre-paying the estate tax via gift taxation as long as the marginal tax rate increases with the base without even appealing to the difference in statutory marginal tax rate schedules. The key point here is that early gifts remove future appreciation from the tax base and this effect is beneficial in the presence of progressivity. The effect is especially strong when assets are expected to appreciate (even if just in nominal terms) or horizon is long.

Another wrinkle influencing the choice between gifts and bequests relates to tax treatment of capital gains. 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 US, 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 tax basis of an 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 benefits of deferral and 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 Joulfaian (2005a) for further discussion and illustrative calculations. Another aspect that has not been accounted for here are differences in marginal tax rates of donors and donees that create additional opportunities for tax arbitrage (Agell & Persson, 2000; Stiglitz, 1985). 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.[36] They first pursue aggregate and micro-based analysis using aggregate time-series variation in top marginal tax rate and marginal tax rates evaluated at 40 or 100 times average wealth as instruments for the individual marginal tax rate at death. This is not a particularly convincing identification strategy and, unsurprisingly, the results are not particularly robust. They subsequently attempt to exploit cross-sectional variation.

---

[36] 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.

To do so, they note that it is expected taxation over individual lifetime that should matter and they propose using the imputed marginal 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. This strategy yields 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. In principle, this could be 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 good proxy for 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 and endogeneity of location decisions. Joulfaian (2006b) pursues a systematic attempt to exploit time variation to explain the size of estates. Rather than using marginal tax rates 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 (2000b): 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 fully 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.[37] The baseline specifications in each case attempt to shed a light on the response to incentives over the lifetime (rather than marginal tax rates 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 the response, the topic to which we will return below.[38]

---

[37] Both Holtz–Eakin and Marples (2001) and Joulfaian (2006b) 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).

[38] In a very recent paper, Seim (2012) finds evidence of significant responses of reported wealth to marginal tax rate incentives under the Swedish annual wealth tax that he attributes to tax evasion.

5.3. Inter Vivos Giving

In contrast to the work on responsiveness of wealth and estates, the literature has made more significant strides in estimating the effect of taxation on giving while alive. The US and many other countries tax gifts and estates in ways that create opportunities for tax avoidance (see Nordblom & Ohlsson, 2006, for a theoretical analysis). In the US, estate and gift taxes have operated completely independently since 1932 (when the gift tax was introduced) until 1976. Since 1977, 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 applies 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.[39] Gift taxation allows for annual exemptions[40] 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.[41] 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 at 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 to well below pre-1976 levels for another decade or so. This is further supported by more formal aggregate time-series econometric specifications that convert 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, Lemke, and Scholz (2004) provide 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

---

[39] 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').

[40] $13,000 in 2011, for each donee separately.

[41] 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.

never expected to be above the exemption, it should discourage gifts for those who are phased out of the tax reach and it could possibly increase gifts for very high net worth individuals via wealth effect.[42] 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 others while gifts for the top category (insignificantly) increase. Page (2003) relies on cross-sectional variation in state estate tax rates and shows relationship between marginal tax rates and the size 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 the planning 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 (Joulfaian & McGarry, 2004; McGarry, 2000, 2001; Poterba, 2001) focused on a different outcome: reliance of taxpayers on the annual gift tax exclusion. This is estate tax planning 101: taxpayers are allowed to transfer tax free up to $13,000 (in 2011) per donee annually, 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.[43] Poterba (2001) concludes that the results imply that taxpayers fail to minimize tax liability.[44]

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

---

[42] 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.

[43] 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.

[44] This is in 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).

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. These patterns are consistent with previously mentioned results of Kopczuk (2007) who concluded that avoidance is particularly pronounced shortly before death and, when coupled with robustly increasing wealth profiles, it reveals that tax minimization is not by itself a complete description of the objective of estate tax payers. As the result, evidence on gift-giving is consistent with the notion stressed previously in this chapter that some motive going beyond consumption value for holding onto wealth until late in life is necessary. I will return to this theme in Section 6.1.

5.4. Labor Supply of Recipients

A number of papers have considered the effect of receiving inheritance on labor supply of recipients. 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 $C = \gamma + B(t) + wL$, where $B(t)$ is the bequest received given donor's tax rate t and γ represents income from other exogenous sources. Labor supply may be written as $L(w, \gamma + B(t))$, making it clear that when B is taken as given by the donee, the receipt of inheritance should generate income effect response of labor supply. Under the standard assumption that leisure is a normal good, receiving inheritance should then lead to a reduction in labor supply.

An increase in the estate tax should (other things constant) generate a response in labor supply proportional to the effect that the tax has on the size of inheritance. In this simple framework, knowledge of the effect of taxation on the size of inheritance ($B'(t)$) and an estimate of the income effect on labor supply coming from elsewhere would be sufficient for evaluating the effect of taxation on labor supply of the donees. As we have seen though, estimating the effect of taxation on inheritance is not trivial (and credible direct estimates of income elasticity 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 recipients' 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 tax implications because bequests are correlated with individual characteristics. Tax induced variation remains necessary 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 and labor supply of the donee are jointly determined. In particular, labor supply may be affected by characteristics of the donor beyond the size of the bequest itself. The natural way of considering this issue is to do so in the context of a dynamic framework along the lines considered in the Samaritan's dilemma (Bruce & Waldman, 1991; Coate, 1995) results that break Becker's "rotten kid" theorem (Becker, 1974; Bergstrom, 2008; Bruce & Waldman, 1990).[45] Casting it in an empirical framework, labor supply may be written as $L(w, w^P, \gamma+B(t, w, w^P))$, 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 a 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 characteristics and circumstances of the recipient. In particular, potential 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.[46]

Holtz-Eakin, Joulfaian, and Rosen (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 one with important economic consequences.[47] Holtz-Eakin et al. (1993) used information from income tax returns of

---

[45] The "rotten kid" results clarify conditions under which a benevolent household head can use transfers to guarantee that self-interested household members will act to maximize family income/utility. See Bergstrom (2008) for discussion. In particular, when parents cannot commit not to help their children in the future ("Samaritan's dilemma"), it creates a strategic incentive for children to take actions that will increase such transfers (such as overconsumption/underinvesting) and may lead to an inefficient allocation for the family as a whole.

[46] 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.

[47] A conceptually different but fascinating channel that wealthy parents invest in preference for leisure for their children has been analyzed theoretically by Doepke and Zilibotti (2008).

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, Coile, and Weisbenner (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 hours response—this is important, because evidence suggests that labor supply responses are much stronger on the extensive margin in general and on retirement in particular. Importantly, HRS includes questions 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 actual size compares to prior expectations. They confirm the finding of negative participation effect and further show that the effect is stronger for unexpected inheritances.[48]

Elinder, Erixson, and Ohlsson (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).[49] 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 small consumption increases following receipt of inheritance and Joulfaian (2006a) who documents that wealth responds 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, Rubin, & Sacerdote, 2001; Kuhn, Kooreman, Soetevent, & Kapteyn, 2011) and stock market wealth (e.g., Coile & Levine, 2006; Poterba, 2000a), although the results do not paint a fully consistent picture. As discussed before, their applicability in the context of bequests requires the strong assumption of inheritance being equivalent to exogenous income so that direct studies of the impact of inheritance remain of independent interest (and they can be also viewed as identifying the effect of an unearned income shock). The literature on the effect of wealth shocks has provided evidence suggestive of negative effects on labor supply of donors, 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 carefully studied.

---

[48] 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).

[49] 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.

5.5. Entrepreneurship, Family Firms, and Inherited Control

Going beyond simple labor supply responses, Holtz–Eakin, Joulfaian, and Rosen (1994a, 1994b) 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. 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, positive effects on entrepreneurship may not have immediate revenue consequences so that estimates of short–term revenue impact may not be informative about longer term effects. Evidence on the link between inheritance and lifting liquidity constraints is mixed. Tsoutsoura (2011) uses 2002 repeal of inheritance taxation in Greece to show the positive effect of the tax on investment in transferred firms and provides some evidence consistent with the importance of financial constraints in driving this effect. 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 & 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 (Bloom & Van Reenen, 2007; Pérez–González, 2006; Villalonga & Amit, 2006).[50] Evidence on whether 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 reduction 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 the likelihood that the business will have to be sold. Holtz–Eakin, Phillips, and Rosen (2001) study life insurance purchases of entrepreneurs and conclude that they do not take full advantage of opportunities to protect their firms from being sold in order to meet the estate tax liability.

---

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

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, lifetime transfers—and those that are solely intended to reduce tax liability with no real consequences.

As usual in tax-related contexts, drawing a line between "real" and "avoidance" responses is difficult. Consider for example an extreme type of response that has been discussed in the literature: Kopczuk and Slemrod (2003a) 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 (2003a), than 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.

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 a trade-off between reducing tax liability and losing 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 cross-sectional upward sloping wealth profile extending to people in the 1990s). However, since the estate tax return form used to include (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—15–20% drop with an illness lasting "months to years." After evaluating alternative explanations, he concludes that the most likely one is tax avoidance. In particular, composition of assets and deductions reported on tax returns changes in a way indicating tax avoidance and a very strong response is present for "lifetime transfers"—gifts that need to be included on the estate tax return because they were made shortly before or in anticipation of death (and hence are "unsuccessful" in obtaining preferential gift tax treatment), but that are likely fingerprints of tax avoidance.[51] Taken together, it suggests that despite continuing accumulation, wealthy taxpayers underinvest in tax planning until the onset of a terminal illness, but they still reveal by their actions at the end of life that they value tax reductions. Because tax planning is much more effective when done early, it implies 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 notion of benefits from controlling wealth is a natural candidate here.

The desire to retain control is also consistent with previously discussed evidence about strong responsiveness of gifts to tax incentives, even though the level of giving is grossly insufficient for the purpose of minimizing tax liability. Such behavior could be naturally explained by simultaneous desire to reduce taxation coupled with some 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.[52]

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 very high net worth individuals

[51] Such transfers are reported on Schedule G of the estate tax form. They may correspond to attempts at tax avoidance that were not successful because of premature death but they may also correspond to successful tax avoidance strategies. For example, many trust instruments involve a transfer from the taxpayer to the trust. A popular example is an irrevocable life insurance trust that is intended to exclude the proceeds of a policy from the estate (incidentally, popularity of gifts of life insurance might possibly be partially explained by the desire to postpone giving control over assets to children). Private annuities discussed by Cooper (1979) may involve a transfer if not fairly priced. Disposing of stocks by a majority shareholder at or before death, in order to reduce holdings to a minority position and therefore qualify for a minority discount may involve a direct transfer. A transfer of an asset to a family limited partnership in exchange for a minority interest (with associated minority discount, see Schmalbeck, 2001 p. 133, for an example) and retained right to interest or use would be included on Schedule G. Proceeds of buy-out agreements to be executed at death popular at that time may also have been reported on Schedule G. A non-estate tax reason for Schedule G transfers may be an attempt to avoid probate.

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

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 onto wealth appears necessary for successfully explaining the upper tail of wealth distribution and direct microeconomic evidence on this topic remains limited.

6.2. Tax Avoidance and Evasion

There were some attempts to estimate the overall extent of tax avoidance and evasion in this context. Wolff (1996) and Poterba (2000b) proposed an approach that is based on comparing estate tax returns to wealth of the living population (in practice, the data from Survey of Consumer Finances). In order to make this comparison, wealth of the living is *weighted by mortality risk*, and the difference between such mortality risk weighted wealth and observed estates is interpreted as reflecting the extent of tax evasion (and forms of tax avoidance that would give rise to the difference in wealth during lifetime and at death). This procedure needs to take a stand on the appropriate mortality rates to use (mortality experience of high net worth individuals is unlikely to be well proxied by mortality rates for the general public) and cannot account for adjustments shortly before death discussed before. As the result, as Eller, Erard, and Ho (2001) elaborate, different assumptions about mortality assumptions lead to estimates varying from 70% of tax loss to the very small amount. Furthermore, this procedure is sensitive to assumptions about mortality differences between married and single individuals and about the distribution of charitable bequests.

An alternative approach is to rely on audits. Audit-based studies estimate the extent of non-compliance at 8–13% of the overall tax liability (Eller & Johnson, 1999; Erard, 1999) but they cannot identify legal or unchallenged types of responses. Audits of estate tax returns used to be fairly common—Eller et al. (2001) report that they applied to 19% of estates overall and almost 50% of returns with gross estate over $5 million—so that the scope for easily detectable and obviously illegal tax evasion is arguably not large. Instead, responses likely take the form of plausibly legal but often legally uncertain strategies. Somewhat surprisingly, Eller et al. (2001) find that almost 20% of estates had their tax liability *reduced*. At the same time, in 60% of cases the assessed tax increased indicating an important role for enforcement. Changes primarily involved revaluation of assets, with non-compliance spread out over most categories of assets. Mortgages and notes and insurance featured the largest percentage revaluation, adjustments to closely held stock were most important in aggregate and there were only small adjustments for stocks and mutual funds. These patterns are consistent with the presence of opportunities for tax evasion and tax avoidance motivated by legal uncertainty surrounding valuation of assets that result in aggressive tax planning.

A small number of papers looked at particular types of legal avoidance-related responses. Johnson, Mikow, and Eller (2001) and Schmalbeck (2001) discuss the most important

avenues for tax avoidance. Valuation of assets is one of the key issues. Estates have an option of valuing assets either as of time of death or using alternative valuation as of six months after death. For assets that are easily marketable, IRS regulations specify that they should be valued using market prices in the case of stocks or using comparative sales in case of real estate. Assets such as pieces of art should be valued by experts. Valuation tables taking into account life expectancy and market rates of returns exist for valuing annuities and partial interest—some tax avoidance strategies are designed by arbitraging the tables and personal circumstances (e.g., when mortality risk is known to deviate from that assumed by the IRS).

Valuing closely-held businesses is notoriously difficult. This is aggravated by existence of additional rules applying in this context. Special use valuation applies to particular types of businesses and family farms that, if they qualify, can be valued at actual rather than market use. Minority and marketability discounts are a particularly important avenue for reducing tax liability. They allow for a reduction in taxable value of assets if the tax-payer does not hold a controlling interest or when there is no easily available market for the particular asset. Johnson et al. (2001) found that approximately 6% of returns used minority or lack-of-marketability discounts that were on average 10% of the gross estate (conditional on use). Poterba and Weisbenner (2003) compare asset information from the Survey of Consumer Finances, weighted by mortality rates, to the asset composition on estate tax returns reported by the IRS. They find patterns consistent with significant use of minority discounts for non-marketable assets. Their results raise the possibility that the opportunities for valuation discounts present in the estate tax system induce important inter-asset distortions.

6.3. Unrealized Capital Gains

One of the important features of the US capital gains tax is the step-up in basis at death. Consider an asset with the basis of p that by the time of death of the owner is worth $p \cdot (1 + r)$. If sold just before taxpayer's death, proceeds would be subject to the capital gains tax t^G and then to the estate tax τ, resulting in the overall bequest of $p \cdot \left(1 + r \cdot (1 - t^G)\right) \cdot (1 - \tau)$. If held until death without realizing the gain, the basis of the asset is reset to its value at the time of death and the overall value of a transfer is $p \cdot (1 + r) \cdot (1 - \tau)$. In particular, this is the liquidation value of the bequest if the recipient/estate chooses to realize the gain immediately after taxpayer's death.

This feature of the tax system gives rise to a strong incentive to hold capital gains until death. The distortion to the holding period is present even in the absence of step-up in basis, reflecting benefits from deferral of taxation on realized capital gains,[53] but the step-up in basis introduces a particularly strong form of the associated lock-in effect.

---

[53] See Auerbach (1991) and Auerbach and Bradford (2004) for a theoretical proposals of a realization-based tax system that would eliminate holding period distortions.

The literature on capital gains realizations (for example, Auerbach, Burman, & Siegel, 2000; Burman & Randolph, 1994; Burman, 1999; Dai, Maydew, Shackelford, & Zhang, 2008) focused on distortions to the holding period and related tax avoidance strategies.

A small number of papers analyzed the interaction of estate tax and capital gains taxation. Poterba (2001) shows that taxpayers with larger unrealized capital gains are less likely to make inter vivos gifts, thereby providing indirect evidence of the lock-in effect being present in practice. On the other hand, Auten and Joulfaian (2001) present evidence that higher estate tax weakens the magnitude of the lock-in effect by encouraging capital gains realizations earlier in life. One possible channel is via the estate tax encouraging consumption or charitable bequests; another theoretical possibility is that in the presence of the estate tax, the effective capital gains marginal tax rates associated with rebalancing of taxpayer's portfolio are smaller than otherwise because they reduce taxable estate and hence tax liability. Finally, taxpayers may want to realize capital gains early as part of their tax avoidance strategy.

Poterba and Weisbenner (2001) use the SCF data to analyze distributional implications of replacing the estate tax by constructive realization of capital gains at death. While revenue estimates are dated, because of changes in rates and exemptions since then, the paper documents significant heterogeneity in the importance of unrealized gains and hence distributional implications of such a policy switch. Indeed, the 2010 "repeal" highlighted some of these issues in practice. The elimination of the estate tax was associated with basis carryover (rather than constructive realization that Poterba and Weisbenner (2001), assumed).[54] Reduction in tax liability generated by differences in marginal tax rates under capital gains and estate taxation was not uniform, because (among other reasons) the distribution of unrealized capital gains is not uniform. Naturally, many taxpayers dying with capital gains are not subject to the estate tax but they still do benefit from step-up. Hence, the replacement of carryover provision could actually increase their tax burden relative to the estate tax. This has been mitigated by allowing for up to $1.3 million of assets to continue to benefit from the step-up. Still, given the 2009 exemption of $3.5 million and availability of deductions for marital bequests (among other things) that change would have resulted in an increase in tax burden for some otherwise non-taxable taxpayers.[55] While the retroactive repeal made this issue moot for 2010, distributional implications of the relative estate tax vs capital gains treatment will continue to be an important issue in considering future reform proposals.

[54] Under the constructive realization regime, capital gains would be subject to the tax at the time of death with the step-up basis for the recipient reset to the value at death. Under carryover basis, the recipient assumes the basis of the original owner and the tax is not due until the asset is sold. Hence, under the carryover basis capital gains are taxed later and realization may continue to be distorted due to the presence of locked-in gains that accrued to the original owner.

[55] Many commentators and practitioners also worry about administrative complexity of implementing a carryover regime that requires recipients to keep track of the basis of the original owner—potentially for generations. Presumed complexity was one the main reasons why the 1976 provision to introduce carryover basis has never gone into effect.

7. OTHER TOPICS

7.1. Implications for Wealth Distribution and Intergenerational Transmission of Inequality

A large literature has focused on the role of transfers and, sometimes, their taxation in understanding wealth accumulation. Prominently, Kotlikoff and Summers (1981) and Modigliani (1988) reached dramatically different conclusions about the importance of intergenerational transfers in overall wealth accumulation; methodological issues and findings of the resulting literature are summarized in Davies and Shorrocks (2000) who conclude that inheritances are responsible for approximately 35–45% of current wealth. More interestingly, the literature—recently reviewed by Cagetti and De Nardi (2008)—has firmly concluded that while realistic life-cycle framework may account for much of wealth accumulation (for example, Hubbard et al., 1994, 1995; Scholz et al., 2006), understanding the full distribution of wealth requires incorporating some form of a bequest motive (De Nardi, 2004; Dynan et al., 2002, 2004; Gale & Scholz, 1994). A small number of papers analyzed implications of changes in estate taxation for the wealth distribution. The long-term implications are ambiguous in general. Stiglitz (1978) highlights general equilibrium implications of the estate tax: an increase in the estate tax leads to a reduction in capital accumulation and an increase in return to capital; depending on the elasticity of substitution between labor and capital, it may then result in overall increase in the share of capital income. Since capital income is more unequally distributed than labor income, this may result in overall increase in inequality of income and consumption. Using dynamic models augmented to account for, respectively, idiosyncratic labor endowment risk and entrepreneurial risk and calibrated to match selected moments of income and wealth distribution, Castaneda, Diaz-Gimenez, and Rios-Rull (2003) and Cagetti and De Nardi (2009) conclude that long-term implications of repealing the estate tax for wealth inequality are small, although distributional implications depend on the source of revenue used to replace the estate tax and efficiency improves. Benhabib, Bisin, and Zhu (2011) instead allow for capital income risk and, in their policy experiment, find that implications of changes in estate taxation for inequality of wealth are quantitatively large. They argue that the strong effect of estate taxation on wealth inequality in their framework is driven by its effect on capital income dispersion.

Papers in this strand of the literature are routinely calibrated to (static) moments of income and wealth distribution. However, such analysis has not yet incorporated empirical evidence on behavioral responses to estate taxation or bequest motives, hence quantitative implications are likely to be model-dependent. It also treats tax systems in a very stylized manner. Consequently, the results that are supposed to pertain to changes in estate taxation have to be treated very cautiously. In particular, as mentioned when discussing bequest motives, observational implications of the joy-of-giving and wealth-in-utility models are identical except for the consequences of changes in inheritance

taxation. Hence, in a model calibrated to the empirical moments of wealth and income distribution for a particular estate tax structure (the standard procedure in this literature), implications of changing the estate tax are driven by assumptions about bequest motives that are completely arbitrary: because the commonly assumed[56] joy-of-giving motive and wealth-in-utility can generate the same distribution of wealth, calibration procedure pins down the implicit strength of behavioral response to taxation using the assumption about the form of a bequest motive rather than by appealing to any observable quantities. As the result, predictions about the impact of changes in estate taxation in this kind of framework are to a large (but unknown) extent driven by modeling assumptions. One could imagine calibrating a model of this kind to the empirical elasticities of bequests to marginal tax rates, perhaps simply by allowing for a mix of the joy-of-giving and wealth-in-utility motivations in order to gain an extra degree of freedom, but it has not been done yet (and neither are the implied micro elasticities explicitly reported in these papers). This approach has, however, the advantage of considering general equilibrium and long-term implications.

Calibration of either initial or steady state of the economy to a distribution at a point in time is convenient but likely to miss first-order features of the actual practical experience. Recent work of Piketty, Postel-Vinay, and Rosenthal (2003), Kopczuk and Saez (2004a), and Roine and Waldenström (2009) documents long-term evolution of wealth concentration.[57] The key point is that wealth concentration has significantly evolved over time. Piketty (2011) shows that the overall annual flow of bequests in France exceeded 20% in the 19th century, fell to 5% by the 1950s and increased to 15% by the early 21st century. These changes highlight that modeling long-term steady-state distribution of wealth, as much of literature has done, is likely to miss first-order facts about historical and future experience. Piketty (2011) shows that these patterns can be accounted by changes in the relationship of growth rate and private rate of return and, hence, as that relationship changes, the role of inheritance in generating inequality changes as well. Piketty and Saez (2012a) build on this insight to study the role of capital and inheritance taxation in a framework that allows for both labor and inheritance inequality to influence welfare objectives and find a role for inheritance taxation that varies with the overall share of inheritances.

The US evidence on changes in wealth concentration is somewhat inconclusive. In an influential paper, Piketty and Saez (2003) documented dramatic changes in *income* shares at the very top of the distribution (see Atkinson, Piketty & Saez, 2011, for a recent review of this literature) since the 1970s. These changes were to a large extent labor income phenomena. A number of papers attempted to document the corresponding trends in wealth concentration but found no or only small increase in wealth accruing to the very

---

[56] Altruistic motive is not sufficient to generate a thick upper tail of the wealth distribution.

[57] This line of research generally relies on estate multiplier technique, building on older work of Mallet (1908), Atkinson and Harrison (1978), and Lampman (1962). See Piketty (2011) for additional references and appendices in Kopczuk and Saez (2004b) for the discussion of the limitations of this approach.

richest (see Kennickell (2006) and Scholz (2003) using Survey of Consumer Finances data and Kopczuk and Saez (2004a) using estate multiplier technique). Edlund and Kopczuk (2009) show that the composition of the top wealth holders has dramatically changed over the last 30 years, shifting from inheritors to self-made—these patterns are visible explicitly in the Forbes list of the richest Americans and can be indirectly corroborated based on the gender composition of top estate taxpayers relying on the assumption that self-made wealth is more male-biased than inheritances. They argue that this change in composition—new wealth building up and old (presumably, in large part dating back to the early 20th century) wealth declining—are consistent with both income inequality increases and relative stability of wealth concentration.

7.2. Charity

Deduction for charitable contributions is the second largest deduction (after marital deduction) used by estate taxpayers and the largest one for unmarried decedents. By exempting charitable bequests from taxation, the estate tax provides price incentives to contribute to charity rather than to other beneficiaries. The presence of the estate tax also has wealth effect that should reduce overall charitable bequests (assuming that they are a normal good). A strand of literature tried to understand how these incentives affect charitable contributions; in particular, attempting to separate price and wealth effects and attempting to evaluate the overall impact of taxation on the flow of charity. Cross-sectional evidence (Auten & Joulfaian, 1996; Boskin, 1976; Joulfaian, 1991,1998; McNees, 1973) has problems separately identifying the direct of wealth from the effect of tax price because wealth and tax price are in practice correlated. Joulfaian (2000) uses cross-sectional state-level variation that improves on earlier identification strategies but continues to be subject to concerns about endogeneity of state tax policy and taxpayers' mobility. Barthold and Plotnick (1984) use data from probate records in Connecticut in the 1930s and 1940s. Brunetti (2005) studied effect of the repeal of the California estate tax in 1982 using 1980–1982 probate records from San Francisco. Kopczuk and Slemrod (2003b) rely instead of time series evidence using current and lagged marginal tax rates. Bakija, Gale and Slemrod (2003) rely on repeated cross-sectional tabulations based on IRS estate tax return data and state-level variation in tax rates. Joulfaian (2005b) uses microdata for estate taxpayers in 1976 and 1982 and exploits variation generated by changes in marginal tax rates and exemption that took place in between.

The literature generally finds that charitable bequests are very sensitive to both their tax price and to wealth, with the first effect dominating so that eliminating the estate tax would likely lead to reduction in charitable bequests. A number of papers additionally recognized that incentives to give at death are also affected by other taxes. For example, a taxpayer making lifetime gifts could additionally benefit from income tax deduction. Alternatively, giving in life and at death may be complements, so that high income tax rates may increase charitable bequests. Bakija, Gale, and Slemrod (2005) and Joulfaian

(2001) account for lifetime incentives and find that they matter but that bequests remain very sensitive to tax rates at death.

A channel that the literature acknowledges (e.g., Bernheim, 1987; Kopczuk & Slemrod, 2003b) but that has not been extensively studied empirically is the interaction of charitable bequests with tax avoidance. Indeed, some of the prominent tax avoidance techniques have charitable contribution components or, conditional on making legitimate charitable gift, there may be additional tax opportunities depending on how the gift is structured (Schmalbeck, 2001). This is an interesting area for future work.

7.3. Other Issues

An important consideration in tax planning is the treatment of marital bequests. Before 2011, taxpayers with a surviving spouse had an incentive to limit the bequest to the surviving spouse in order to take advantage of the exemption. More generally, there is an incentive (quantitatively important, at least at low estate levels) to go through the progressive tax schedule twice by splitting the bequest between the spouse and other potential beneficiaries. Empirically, heavy reliance on the marital deduction is very common but, as recognized by estate tax planners, it raises difficult issues related to the conflict of interest between spouses—preferences about the ultimate disposition of bequest may differ and hence the transfer to the spouse may result in the ultimate disposition of bequests that is inconsistent with preferences of the first-to-die spouse. While this could potentially be evidence in support of a unitary household model, Johnson et al. (2001) document widespread use and Kopczuk and Slemrod (2003b) analyze implications of the so-called Qualified Terminable Interest Property (QTIP) trusts that can resolve this tension: introduced in 1982, such trusts are transferred to the widow or widower who gains access to earnings but at her death the principal is transferred to the beneficiaries indicated by the first-to-die spouse's will. Hence, they qualify for a marital deduction but allow for addressing the conflict of interest.

A few papers considered the issues related to jurisdictional-level estate taxes. Bakija and Slemrod (2004) and Conway and Rork (2006) focus on impact of state-level taxes on elderly mobility using longitudinal variation in state tax rates. Bakija and Slemrod (2004) find significant but modest effects, while Conway and Rork (2006) find no impact and instead argue for the reverse causality with mobility affecting tax rate setting. Conway and Rork (2004) find some evidence of tax competition between US estates in this dimension, while Brülhart and Parchet (2011) do not find competitive effects for the Swiss cantons.

Several papers attempted to provide explanations for introduction and elimination of estate taxes. Bertocchi (2011) attempted to link evolution of estate taxation to the dynamics of wealth accumulation induced by the process of industrialization. Scheve and Stasavage (2012) focus on the role of wars. Graetz and Shapiro (2005) discuss politics surrounding estate tax changes in the United States in the 2000s.

8. SUMMARY AND CONCLUSIONS

In this chapter, I reviewed theoretical and empirical literature on taxation of intergenerational transfers. As signaled in the introduction, the conclusions are mixed and pointing to the need for further research. The normative analysis has overemphasized the importance of externalities from giving, while not adequately accounting for the importance of inherited wealth. When taking an *ex ante* perspective, inherited wealth is endogenous and does not constitute a reason for redistribution per se. While this is consistent with the logic of that model, it does not reflect the possibility that policy makers may be instead focused on redistribution within a generation. The steady-state approaches (Kopczuk, 2001; Piketty & Saez, 2012a,2012b) find a positive role for capital taxation but it imposes restrictive assumptions.

The key empirical observation that should feed into theoretical work is the presence of heterogeneity. The literature on bequest motives has failed to identify the single motive and instead points to both mixed motives present at the same time for a given person and to heterogeneity in preferences in the population. In particular, the first-order fact in understanding behavior of the very wealthy is the importance of control over wealth. The presence of such a motive is consistent with relative scarcity of giving during lifetime coupled with significant tax avoidance and increasing age-wealth profiles. Some direct preference for wealth (rather than consumption or welfare of children) is necessary to account for the extent of wealth concentration.

The empirical work has focused on many dimensions of distortions due to the estate tax and generally has found evidence that a number of decisions are sensitive to the tax rates—estates at death, tax avoidance, inter vivos giving, charitable contributions. It has also identified interesting and adverse effects on labor supply of donors and performance of family businesses. At the same time, the identification issues in this context are difficult and rarely satisfactorily addressed. In particular, the estate tax incentives operate over a long period of time and relating such incentives to wealth accumulation (which is perhaps the most interesting dimension of responsiveness) is notoriously difficult. The literature has been more successful in analyzing decisions where focusing on short-term incentives is appropriate, such as the inter vivos giving.

Tax avoidance is believed to be very important in practice but precise econometric evidence is scarce. In particular, issues related to separating response of wealth accumulation from tax avoidance (both of which affect reported estates), the interaction of charity and tax avoidance, impact of cross-asset differences in effective tax rates, the role of family structure and conflicts of interest, use of family-related partnerships, valuation issues for closely-held business and other types of assets have all received only limited attention from economists.

Finally, the topic that has not yet been incorporated in the formal models of intergenerational transfers are potential negative externalities from wealth concentration. Examples

are positional externalities (Frank, 2008), the impact of corporate power on the political process (Morck, Wolfenzon, & Yeung, 2005) and previously discussed evidence of inheritance of control resulting in mis-allocation of entrepreneurial skills. Recently, a number of papers started considering implications of negative externalities associated with income inequality—primarily rent-seeking—for optimal taxation (Rothschild & Scheuer, 2011; Piketty, Saez, & Stantcheva, 2011). Incorporating such considerations in the analysis of intergenerational transfers is long overdue.

ACKNOWLEDGMENTS

REFERENCES

Abel, A. B., & Warshawsky, M. (1988). Specification of the joy of giving: Insights from altruism. *Review of Economics and Statistics, 70*(1), 145–149.

Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., et al. (2011). Taxes on wealth transfers. In *Tax by design. The mirrlees review* (pp. 347–367). Oxford University Press.

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

Agell, J., Persson, M. (2000). Tax arbitrage and labor supply. *Journal of Public Economics, 78*(1–2), 3–24.

Aiyagari, S. R. (1995). Optimal capital income taxation with incomplete markets, borrowing constraints and constant discounting. *Quarterly Journal of Economics, 103*(6), 1158–1175.

Albanesi, S., & Sleet, C. (2006). Dynamic optimal taxation with private information. *Review of Economic Studies, 73*(1), 1–30.

Altonji, J. G., Hayashi, F., & Kotlikoff, L. J. (1992). Is the extended family altruistically linked? Direct tests using micro data. *American Economic Review, 82*(5), 1177–1198.

Altonji, J. G., Hayashi, F., & Kotlikoff, L. J. (1997). Parental altruism and inter vivos transfers: Theory and evidence. *Journal of Political Economy, 105*(6), 1121–66.

Ameriks, J., Caplin, A., Laufer, S., & Van Nieuwerburgh, S. (2011). The joy of giving or assisted living? Using strategic surveys to separate public care aversion from bequest motives. *Journal of Finance, 66*(2), 519–561.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *Economic Journal, 100*(401), 464–477.

Arrondel, L., & Masson, A. (2006). Altruism, exchange or indirect reciprocity: What do the data on family transfers show? In S. -C. Kolm, & J. M. Ythier (Eds.), *Handbook on the economics of giving, reciprocity and altruism, Vol. 2* (pp. 971–1053). Elsevier.

Arrondel, L., & Laferrère, A. (2001). Taxation and wealth transmission in France. *Journal of Public Economics, 79*(1), 3–33. ISPE 1998: Bequests and Wealth Taxation.

Atkeson, A., & Lucas, R. E., Jr. (1992). On efficient distribution with private information. *Review of Economic Studies, 59*(3), 427–453.

Atkeson, Andrew, Chari, V. V., & Kehoe, Patrick J. (1999). Taxing capital income: A bad idea. *Federal Reserve Bank of Minneapolis Quarterly Review, 23*(3), 3–17.

Atkinson, A. B., & Harrison, A. J. (1978). *Distribution of personal wealth in Britain.* Cambridge: Cambridge University Press.

Atkinson, A. B., & Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics, 6*(1–2), 55–75.

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature, 49*(1), 3–71.

Auerbach, A. J. (1991). Retrospective capital gains taxation. *American Economic Review, 81*(1), 167–178.

Auerbach, A. J. (2008). Taxation of wealth. In S. N. Durlauf, & L. E. Blume (Eds.), *The new palgrave dictionary of economics* (second ed.). Palgrave Macmillan.

Auerbach, A. J. (2009). The choice between income and consumption taxes: A primer. In A. J. Auerbach, & D. N. Shavior (Eds.), *Institutional foundations of public finance: Economic and legal perspectives* (pp. 13–46). Harvard University Press.

Auerbach, A. J., & Bradford, D. F. (2004). Generalized cash flow taxation. *Journal of Public Economics, 88*(5), 957–80.

Auerbach, A. J., Burman, L. E., & Siegel, J. M. (2000). Capital gains taxation and tax avoidance: New evidence from panel data. In J. Slemrod (Ed.), *Does atlas shrug? The economic consequences of taxing the rich*. New York: Harvard University Press and Russell Sage Foundation.

Auten, G., Joulfaian, D. (1996). Charitable contributions and intergenerational transfers. *Journal of Public Economics, 59*(1), 55–68.

Auten, G., & Joulfaian, D. (2001). Bequest taxes and capital gains realizations. *Journal of Public Economics, 81*(2), 213–229.

Bakija, J. (2007). *Documentation for a federal and state inheritance and estate tax calculator*. Williams College: Mimeo.

Bakija, J., & Slemrod, J. (2004). Do the rich flee from high tax states? Evidence from federal estate tax returns. *Working paper* 10645, National Bureau of Economic Research.

Bakija, J., Gale, W. G., & Slemrod, J. B. (2003). Charitable bequests and taxes on inheritances and estates: Aggregate evidence from across states and time. *American Economic Review Papers and Proceedings, 93*(2), 366–370.

Bakija, J., Gale, W., & Slemrod, J. (2005). *New evidence on the effect of taxes on charitable bequests*. Williams College, Brookings and University of Michigan: Mimeo.

Banks, J., & Diamond, P. A. (2010). The Base for direct taxation. In J. A. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. D. Myles, & J. Poterba (Eds.), *Dimensions of tax design: The mirrlees review*. Oxford University Press.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy, 82*(6), 1095–1117.

Barthold, T., & Plotnick, R. (1984). Estate taxation and other determinants of charitable bequests. *ntj, 37*(2), 225–37.

Batchelder, L. L. (2009). What should society expect from heirs? A proposal for a comprexensive inheritance tax. *Tax Law Review, 63*(1), 1–111.

Becker, E. (1973). *The denial of death*. New York: The Free Press.

Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy, 82*(6), 1063–1093.

Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics, 117*(3), 871–915.

Benhabib, J., Bisin, A., & Zhu, S. (2011). The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica, 79*(1), 123–157.

Bergstrom, T. C. (2008). Rotten kid theorem. In S. N. Durlauf & L. E. Blume (Eds.), *The new palgrave dictionary of economics*. Basingstoke: Palgrave Macmillan.

Bernheim, B. D. (1987). Does the estate tax raise revenue? In L. H. Summers (Ed.), *Tax policy and the economy, Vol. 1* (pp. 113–138). Chicago: National Bureau of Economic Research; Cambridge, Mass: MIT Press.

Bernheim, B. D., & Severinov, S. (2003). Bequests as signals: An explanation for the equal division puzzle. *Journal of Political Economy, 111*(4), 733–764.

Bernheim, B. D., Shleifer, A., & Summers, L. H. (1985). The strategic bequest motive. *Journal of Political Economy, 93*(6), 1045–76.

Bernheim, B. D., Lemke, R. J., & Scholz, J. K. (2004). Do estate and gift taxes affect the timing of private transfers? *Journal of Public Economics, 88*(12), 2617–34.

Bertocchi, G. (2011). The vanishing bequest tax: The comparative evolution of bequest taxation in historical perspective. *Economics & Politics, 23,* 107–131.

Bird, R. M. (1991). The taxation of personal wealth in international perspective. *Canadian Public Policy andsang Analyse de Politiques, 17*(3), 322–334.

Bloom, N., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics, 122*(4), 1351–1408.

Blumkin, T., & Sadka, E. (2004). Estate taxation with intended and accidental bequests. *Journal of Public Economics, 88*(1–22), 1–21.

Boadway, R., Chamberlain, E., & Emmerson, C. (2010a). Taxation of wealth and wealth transfers. In J. A. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. D. Myles, & J. Poterba (Eds.), *Dimensions of tax design: The mirrlees review*. Oxford University Press.

Boadway, R., Chamberlain, E., & Emmerson, C. (2010b). Taxation of wealth and wealth transfers. Online Appendix, <http://www.ifs.org.uk/mirrleesreview/reports/wealthtransfersapps.pdf>

Boskin, M. (1976). Estate taxation and charitable bequests. *Journal of Public Economics, 5*(1–2), 27–56.

Brown, J. R., Coile, C. C., & Weisbenner, S. J. (2010). The Effect of inheritance receipt on retirement. *Review of Economics and Statistics, 92*(2), 425–434.

Brown, R. D. (1991). A primer on the implementation of wealth taxes. *Canadian Public Policy/ Analyse de Politiques, 17*(3), 335–50.

Bruce, N., & Waldman, M. (1990). The rotten-kid theorem meets the Samaritan's dilemma. *Quarterly Journal of Economics, 105*(1), 155–165.

Bruce, N., & Waldman, M. (1991). Transfers in kind: Why they can be efficient and nonpaternalistic. *American Economic Review, 81*(5), 1345–1351.

Brülhart, M., & Parchet, R. (2011). *Alleged tax competition: The mysterious death of bequest tax in Switzerland*. University of Lausanne: Mimeo.

Brunetti, M. J. (2005). The estate tax and charitable bequests: Elasticity estimates using probate records. *National Tax Journal, 58*(2), 165–188.

Brunetti, M. J. (2006). The estate tax and the demise of the family business. *Journal of Public Economics, 90*(10–11), 1975–93.

Burman, L. E. (1999). *The labirynth of capital gains tax policy*. Washington, D.C.: Brookings Institution Press.

Burman, L. E., & Randolph, W. C. (1994). Measuring permanent responses to capital-gains tax changes in panel data. *American Economic Review, 84*(4), 794–809.

Cagetti, M., & De Nardi, M. (2008). Wealth inequality: Data and models. *Macroeconomic Dynamics, 12*, 285–313.

Cagetti, M., & De Nardi, M. (2009). Estate taxation, entrepreneurship, and wealth. *American Economic Review, 99*(1), 85–111.

Carrillo, J. D., & Mariotti, T. (2000). Strategic ignorance as a self-disciplining device. *Review of Economic Studies, 67*, 529–544.

Carroll, C. D. (2000). Why do the rich save so much? In J. Slemrod (Ed.), *Does atlas shrug? The economic consequences of taxing the rich*. New York: Harvard University Press and Russell Sage Foundation.

Castaneda, A., Diaz-Gimenez, J., & Rios-Rull, J.-V. (2003). Accounting for the U.S. earnings and wealth inequality. *Journal of Political Economy, 111*(4), 818–857.

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica, 54*(3), 607–22.

Chari, V. V., & Kehoe, P. J. (1999). Optimal fiscal and monetary policy. In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics* (pp. 1671–1745). Number 1C: Elsevier.

Charles, K. K., & Hurst, E. (2003). The correlation of wealth acrross generations. *Journal of Political Economy, 6*(111), 1155–1182.

Chetty, R. (2009). Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance. *American Economic Journal: Economic Policy, 1* (2), 31–52.

Coate, S. (1995). Altruism, the Samaritan's dilemma, and government transfer policy. *American Economic Review, 85*(1), 46–57.

Coile, C. C., & Levine, P. B. (2006). Bulls, bears, and retirement behavior. *Industrial and Labor Relations Review, 59*(3), 408–429.

Conway, K. S., & Rork, J. C. (2004). Diagnosis murder: The death of state death taxes. *Economic Inquiry, 42*(4), 537–559.

Conway, K. S., & Rork, J. C. (2006). State 'death' taxes and elderly migration—the chicken or the egg? *National Tax Journal, 59*(1), 97–128.

Cooper, G. (1979). *A voluntary tax? New perspectives on sophisticated tax avoidance*. Studies of Government Finance, Washington D.C.: The Brookings Institution.

Cox, D. (1987). Motives for private income transfers. *Journal of Political Economy, 95*(3), 508–546.

Cremer, H., & Pestieau, P. (2001). Non-linear taxation of bequests, equal sharing rules and the tradeoff between intar- and inter-family inequalities. *Journal of Public Economics, 79*(1), 35–54.

Cremer, H., & Pestieau, P. (2006). Wealth transfer taxation: A survey of the theoretical literature. In S.-C. Kolm & J. M. Ythier (Eds.), *Handbook on the economics of giving, reciprocity and altruism, Vol. 2* (pp. 1107–1134). Elsevier.

Cremer, H., Gahvari, F., & Ladoux, N. (1998). Externalities and optimal taxation. *Journal of Public Economics, 70*(3), 343–364.

Cremer, H., Pestieau, P., & Rochet, J. -C. (2003). Capital income taxation when inherited wealth is not observable. *Journal of Public Economics, 87*(11), 2475–2490.

Dai, Z., Maydew, E. L., Shackelford, D. A., & Zhang, H. H. (2008). Capital gains taxes and asset prices: Capitalization or lock-in? *Journal of Finance, 63*(2), 709–742.

Davies, J. B., & Shorrocks, A. F. (2000). The distribution of wealth. In A. B. Atkinson & F. Bourguignon (Eds.), *Handbook of income distribution*. Amsterdam, New York: Elsevier.

De Nardi, M. (2004). Wealth inequality and intergenerational links. *Review of Economic Studies, 71*(3), 743–768.

Diamond, P. (1998). Optimal income taxation: An example with u-shaped pattern of optimal tax rates. *American Economic Review, 88*(1), 83–95.

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

Diamond, P. A., & Mirrlees, J. A. (1978). A model of social insurance with variable retirement. *Journal of Public Economics, 10*(3), 295–336.

Diamond, P., & Spinnewijn, J. (2010). *Capital income taxes with heterogeneous discount rates*. MIT and LSE: Mimeo.

Doepke, M., & Zilibotti, F. (2008). Occupational choice and the spirit of capitalism. *Quarterly Journal of Economics, 123*(2), 747–793.

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

Dynan, K. E., Skinner, J., & Zeldes, S. P. (2002). The importance of bequests and life-cycle saving in capital accumulation: A new answer. *American Economic Review, 92*(2), 274–78.

Dynan, K. E., Skinner, J., & Zeldes, S. P. (2004). Do the rich save more? *Journal of Political Economy, 112*(2), 397–444.

Eaton, J., & Rosen, H. S. (1980). Taxation, human capital and uncertainty. *American Economic Review, 70*(4), 705–715.

Edlund, L., & Kopczuk, W. (2009). Women, wealth and mobility. *American Economic Review, 99*(1), 146–78.

Eliason, M., & Ohlsson, H. (2008). Living to save taxes. *Economics Letters, 100*(3), 340–343.

Elinder, M., Erixson, O., & Ohlsson, H. (2011). Carnegie visits nobel: Do inheritances affect labor and capital income? *Working paper* 2011:5. Department of Economics: Uppsala University.

Eller, M., & Johnson, B. W., (1999). Using a sample of federal estate returns to examine the effects of audit revaluation on pre-audit estimates. In: *Proceedings of the 1999 meeting of the American statistical association, section on government statistics*.

Eller, M., Erard, B., & Ho, C. -C. (2001). The magnitude and determinants of federal estate tax noncompli-ance. In W. G. Gale, J. R. Hines, Jr., & J. Slemrod (Eds.), *Rethinking estate and gift taxation* (pp. 375–410) Brookings Institution Press.

Erard, B. (1999). *Estate tax underreporting gap study: A report prepared for the internal revenue service economic analysis and modeling group*. TIRNO-98-P-00406, Internal Revenue Service.

Farhi, E., & Werning, I. (2007). Inequality and social discounting. *Journal of Political Economy, 115*(3), 365–402.

Farhi, E., & Werning, I. (2010). Progressive estate taxation. *Quarterly Journal of Economics, 125*(2), 635–73.

Feldstein, M. S. (1995). The effect of marginal tax rates on taxable income: A panel study of the 1986 tax reform act. *Journal of Political Economy, 103*(3), 551–572.

Feldstein, M. S. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics, 4*(81), 674–680.

Francis, J. L. (2009). Wealth and the capitalist spirit. *Journal of Macroeconomics, 31*(3), 394–408.

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

Friedman, B. M., & Warshawsky, M. (1990). The cost of annuities: Implications for saving behavior and bequests. *Quarterly Journal of Economics, 105*(1), 135–54.

Fullerton, D. (1997). Environmental levies and distortionary taxes: Comment. *American Economic Review, 87*(1), 245–51.

Gale, W. G., Slemrod, J. (2001). Rethinking the estate and gift tax: Overview. In W. G. Gale, J. R. Hines, Jr., & J. Slemrod (Eds.), *Rethinking estate and gift taxation*. Brookings Institution Press.

Gale, W. G., & Scholz, J. K. (1994). Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives, 8*(4), 145–160.

Gans, J. S., & Leigh, A. (2006). Did the death of Australian inheritance taxes affect deaths? *B.E. Journal of Economic Analysis and Policy: Topics in Economic Analysis and Policy, 6*(1), 1–7.

Golosov, M., Tsyvinski, A., & Weinzierl, M. (2010). *Preference heterogeneity and optimal commodity taxation*. Yale and Harvard: Mimeo.

Golosov, M., & Tsyvinski, A. (2006). Designing optimal disability insurance: A case for asset testing. *Journal of Political Economy, 114*(2), 257–279.

Golosov, M., Kocherlakota, N., & Tsyvinski, A. (2003). Optimal indirect and capital taxation. *Review of Economic Studies, 70*(244), 569–587.

Gordon, R., & Kopczuk, W. (2010). *The choice of personal income tax base*. UC-San Diego and Columbia, preliminary draft: Mimeo. <http://www.columbia.edu/~wk2110/bin/choiceBase.pdf>.

Graetz, M. J., & Shapiro, I. (2005). *Death by a thousand cuts: The fight over taxing inherited wealth* (1st ed.). Princeton University Press.

Green, J., & Sheshinsky, E. (1976). Direct versus indirect remedies for externalities. *Journal of Political Economy, 84*(4), 797–808.

Grossmann, V., & Strulik, H. (2010). Should continued family firms face lower taxes than other estates? *Journal of Public Economics, 94*(1–2), 87–101.

Holtz-Eakin, D., & Marples, D. (2001). Distortion costs of taxing wealth accumulation: Income versus estate taxes. *Working paper* 8261, National Bureau of Economic Research.

Holtz-Eakin, D., Joulfaian, D., & Rosen, H. S. (1993). The carnegie conjecture: Some empirical evidence. *Quarterly Journal of Economics, 108*(2), 413–435.

Holtz-Eakin, D., Joulfaian, D., & Rosen, H. S. (1994a). Entrepreneurial decisions and liquidity constraints. *RAND Journal of Economics, 25*(2), 334–347.

Holtz-Eakin, D., Joulfaian, D., & Rosen, H. S. (1994b). Sticking it out: Entrepreneurial survival and liquidity constraints. *Journal of Political Economy, 102*(1), 53–75.

Holtz-Eakin, D., Phillips, J. W., & Rosen, H. S. (2001). Estate taxes, life insurance, and small business. *Review of Economics and Statistics, 83*(1), 52–63.

Hubbard, R. G., Skinner, J., & Zeldes, S. P. (1994). The importance of precautionary motives in explaining individual and aggregate saving. *Carnegie-Rochester Conference Series on Public Policy, 40*, 59–125.

Hubbard, R. G., Skinner, J., & Zeldes, S. (1995). Precautionary saving and social insurance. *Journal of Political Economy, 103*(2), 360–399.

Hurd, M. D. (1987). Savings of the elderly and desired bequests. *American Economic Review, 77*(3), 298–312.

Hurd, M. D. (1989). Mortality risk and bequests. *Econometrica, 57*(4), 779–813.

Hurst, E., Lusardi, A. (2004). Liquidity constraints, household wealth, and entrepreneurship. *Journal of Political Economy, 112* (2), 319–347.

Imbens, G. W., Rubin, D. B., & Sacerdote, B. I. (2001). Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players. *American Economic Review, 91*(4), 778–794.

Johnson, B. W. (Ed.). (1994). Compendium of federal estate tax and personal wealth studies. Department of Treasury, *Internal Revenue Service, Pub., 1773*(4–94), 1994.

Johnson, B. W., Mikow, J. M., & Eller, M. B. (2001). Elements of federal estate taxation. In W. G. Gale, J. R. Hines, Jr., & J. Slemrod (Eds.), *Rethinking estate and gift taxation*. Brookings Institution Press.

Joulfaian, D. (1991). Charitable bequests and estate taxes. *National Tax Journal, 44*(2), 169–180.

Joulfaian, D. (1998). *Charitable bequests and estate taxes: Another look at the evidence*. Office of Tax Analysis, US Department of Treasury: Mimeo.

Joulfaian, D. (2000). Estate taxes and charitable bequests by the wealthy. *Working paper* 7663, National Bureau of Economic Research.

Joulfaian, D. (2001). Charitable giving in life and death. In W. G. Gale, J. R. Hines, Jr., & J. Slemrod (Eds.), *Rethinking estate and gift taxation* (pp. 350–69). Brookings Institution Press.

Joulfaian, D. (2004). Gift taxes and lifetime transfers: Time series evidence. *Journal of Public Economics, 88*(9–10), 1917–1929.

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

Joulfaian, D. (2005b). Estate taxes and charitable bequests. Evidence from two tax regimes. *Office of tax policy analysis working paper 92*, US Department of Treasury 2005.

Joulfaian, D. (2006a). *Inheritance and saving.* WP 12569, National Bureau of Economic Research October 2006.

Joulfaian, D. (2006b). The behavioral response of wealth accumulation to estate taxation: Time series evidence. *National Tax Journal, 59*(2), 253–68.

Joulfaian, D. (2011). *The federal estate tax: History, law, and economics.* Office of Tax Analysis, US Department of Treasury. Available at SSRN: <http://ssrn.com/abstract=1579829>.

Joulfaian, D., & McGarry, K. (2004). Estate and gift tax incentives and inter vivos giving. *National Tax Journal, 57*(2 (part 2)), 429–444.

Joulfaian, D., & Wilhelm, M. O. (1994). Inheritance and labor supply. *Journal of Human Resources, 29*(4), 1205–1234.

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics, 28*(1), 59–83.

Kaplow, L. (1995). A note on subsidizing gifts. *Journal of Public Economics, 58*(3), 496–77.

Kaplow, L. (1998). Tax policy and gifts. *American Economic Review Papers and Proceedings, 88*(2), 283–88.

Kaplow, L. (2001). A framework for assessing estate and gift taxation. In W. G. Gale, J. R. Hines, Jr., & J. Slemrod (Eds.), *Rethinking estate and gift taxation.* Brookings Institution Press.

Kaplow, L. (2006). On the undesirability of commodity taxation even when income taxation is not optimal. *Journal of Public Economics, 90*(6–7), 1235–50.

Kaplow, L. (2008). *The theory of taxation and public economics.* Princeton and Oxford: Princeton University Press.

Kaplow, L. (2009). On the taxation of private transfers. *Tax Law Review, 63*(1), 159–188.

Kennickell, A. B. (2006). Currents and undercurrents: Changes in the distribution of wealth, 1989–2004. *Discussion paper 2006–13*, Federal Reserve Board.

Kocherlakota, N. R. (2005). Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. *Econometrica, 73*(5), 1587–1622.

Kocherlakota, N. R. (2010). *The new dynamic public finance.* Princeton and Oxford: Princeton University Press.

Kopczuk, W. (2001). *Optimal estate taxation in the steady state.* University of Michigan: Mimeo.

Kopczuk, W. (2003a). A note on optimal taxation in the presence of externalities. *Economics Letters, 80*(1), 81–86.

Kopczuk, W. (2003b). The trick is to live: Is the estate tax social security for the rich? *Journal of Political Economy, 111*(6), 1318–1341.

Kopczuk, W. (2007). Bequest and tax planning: Evidence from estate tax returns. *Quarterly Journal of Economics, 122*(4), 1801–1854.

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

Kopczuk, W. (2013). Incentive effects of inheritances and optimal estate taxation. *American Economic Review Papers and Proceedings, 103*(3).

Kopczuk, W., & Saez, E. (2004a). Top wealth shares in the united states, 1916–2000: Evidence from estate tax returns. *National Tax Journal, 57* (2 (part 2)), 445–488.

Kopczuk, W., & Saez, E. (2004b). Top wealth shares in the united states, 1916–2000: Evidence from estate tax returns. *Working paper* 10399, National Bureau of Economic Research 2004.

Kopczuk, W., & Slemrod, J. (2001). The impact of the estate tax on the wealth accumulation and avoidance behavior of donors. In W. G. Gale, J. R. Hines, Jr., & J. Slemrod (Eds.), *Rethinking estate and gift taxation* (pp. 299–343). Brookings Institution Press.

Kopczuk, W., & Slemrod, J. (2003a). Dying to save taxes: Evidence from estate tax returns on the death elasticity. *Review of Economics and Statistics, 85*(2), 256–265.

Kopczuk, W., & Slemrod, J. (2003b). Tax consequences on wealth accumulation and transfers of the rich. In A. H. Munnell & A. Sundén (Eds.), *Death and dollars: The role of gifts and bequests in America* (pp. 213–249). Brookings Institution Press.

Kopczuk, W., & Slemrod, J. (2005). Denial of death and economic behavior. *Advances in Theoretical Economics, 5*(1), Article 5. <http://www.bepress.com/bejte/advances/vol5/iss1/art5>

Kopczuk, W., & Lupton, J. (2007). To leave or not to leave: An empirical investigation of the distribution of bequest motives. *Review of Economic Studies, 74*(1), 207–235.

Kotlikoff, L. J., & Summers, L. H. (1981). The role of intergenerational transfers in aggregate capital accumulation. *Journal of Political Economy, 89*(4), 706–732.

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

Kuhn, P., Kooreman, P., Soetevent, A., & Kapteyn, A. (2011). The effects of lottery prizes on winners and their neighbors: Evidence from the Dutch postcode lottery. *American Economic Review, 101*(5), 2226–2247.

Laferrère, A., & Wolff, F. -C. (2006). Microeconomic models of family transfers. In: S. -C. Kolm & J. M. Ythier (Eds.), *Handbook on the economics of giving, reciprocity and altruism, Vol. 2* (pp. 889–969). Elsevier.

Laitner, J. (1997). Intergenerational and interhousehold economic links. In M. K. Rosenzweig & O. Stark (Eds.), *Handbook of population and family economics, Vol. 1A*. Amsterdam, New York: Elsevier/North Holland.

Laitner, J., & Juster, F. T. (1996). New evidence on altruism: A study of TIAA-CREF retirees. *American Economic Review, 86*(4), 893–908.

Laitner, J., & Ohlsson, H. (2001). Bequest motives: A comparison of Sweden and the United States. *Journal of Public Economics, 79*(1), 205–236. ISPE 1998: Bequests and Wealth Taxation.

Lampman, R. J. (1962). *The share of top wealth-holders in national wealth, 1922–56*, Princeton, NJ: Princeton University Press.

Laroque, G. (2005). Indirect taxation is superfluous under separability and taste homogeneity: A simple proof. *Economics Letters, 87*(1), 141–44.

Light, A., McGarry, K. (2004). Why parents play favorites: Explanations for unequal bequests. *American Economic Review, 94*(5), 1669–81.

Luckey, J. R. (2008). A history of federal estate, gift and generation-skipping taxes, *CRS Report for Congress 95–444*, Congressional Research Service January 2008.

Mallet, B. (1908). A method of estimating capital wealth from estate duty statistics. *Journal of the Royal Statistical Society, 71*, 65–101.

McCaffery, E. J. (1994). The uneasy case for wealth transfer taxation. *Yale Law Journal, 104*(2), 283–365.

McCubbin, J. G. (1994). Improving wealth estimates derived from estate tax data. In Johnson (Ed.), 363–369.

McGarry, K. (1999). Inter vivos transfers and intended bequests. *Journal of Public Economics, 73*(3), 321–51.

McGarry, K. (2000). Inter vivos transfers or bequests? Estate taxes and the timing of parental giving. *Tax Policy and the Economy, 14*, 93–121.

McGarry, K. (2001). The cost of equality: Unequal bequests and tax avoidance. *Journal of Public Economics, 79*(1), 179–204.

McNees, S. (1973). Deductibility of charitable bequests. *National Tax Journal, 26*(1), 79–98.

Menchik, P. L. (1980). Primogeniture, equal sharing, and the U.S. distribution of wealth. *Quarterly Journal of Economics, 94*(2), 299–316.

Menchik, P. L., & David, M. (1983). Income, distribution, lifetime savings, and bequests. *American Economic Review, 73*(4), 672–690.

Micheletto, L. (2008). Redistribution and optimal mixed taxation in the presence of consumption externalities. *Journal of Public Economics, 92*(10–11), 2262–2274.

Milesi-Ferretti, G. M., & Roubini, N. (1998). On the taxation of human and physical capital in models of endogeneous growth. *Journal of Public Economics, 70*(2), 237–54.

Mintz, J. M. (1991). The role of wealth taxation in the overall tax system. *Canadian Public Policy / Analyse de Politiques, 17*(3), 248–63.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies, 38*(114), 175–208.

Mitchell, O. S., Poterba, J. M., Warshawsky, M. J., & Brown, J. R. (1999). New evidence on the money's worth of individual annuities. *American Economic Review, 89*(5), 1299–1317.

Modigliani, F. (1988). The role of intergenerational transfers and life cycle saving in the accumulation of wealth. *Journal of Economic Perspectives, 2*(2), 15–40.

Morck, R., Wolfenzon, D., & Yeung, B. (2005). Corporate governance, economic entrenchment, and growth. *Journal of Economic Literature, 43*(3), 655–720.

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

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

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

Pérez-González, F. (2006). Inherited control and firm performance. *American Economic Review, 96*(5), 1559–1588.

Perozek, M. G. (1998). Comment: A reexamination of the strategic bequest motive. *Journal of Political Economy, 106*(2), 423–445.

Phelan, C. (2006). Opportunity and social mobility. *Review of Economic Studies, 72*(2), 487–504.

Pigou, A. C. (1920). *Economics of welfare*. London: Macmillan and Co.

Piketty, T. (2001). Income inequality in France 1901–98. *DP 2876*, CEPR 2001.

Piketty, T. (2011). On the long-run evolution of inheritance: France 1820–2050. *Quarterly Journal of Economics, 126*(3), 1071–1131.

Piketty, T., & Saez, E. (2003). Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics, 118,* 1–39.

Piketty, T., & Saez, E. (2007). How progressive is the U.S. federal tax system? A historical and international perspective. *Journal of Economic Perspectives, 21*(1), 3–24.

Piketty, T., & Saez, E. (2012a). A theory of capital income taxation. *Working paper* 17989, NBER April 2012.

Piketty, T., & Saez, E. (2012b). *A theory of optimal inheritance taxation*. Paris School of Economics and UC Berkeley: Mimeo.

Piketty, T., Saez, E., & Stantcheva, S. (2011). Optimal taxation of labor incomes: A tale of three elasticities. *Working paper* 17616, National Bureau of Economic Research November 2011.

Piketty, T., Postel-Vinay, G., & Rosenthal, J. -L. (2003). *Wealth concentration in a developing economy: Paris and France, 1807–1994.* EHESS and UCLA: Mimeo.

Poterba, J. M. (2000a). Stock market wealth and consumption. *Journal of Economic Perspectives, 14*(2), 99–118.

Poterba, J. M. (2000b). The estate tax and after-tax investment returns. In J. Slemrod (Ed.), *Does atlas shrug? The economic consequences of taxing the rich* (pp. 329–349). New York: Harvard University Press and Russell Sage Foundation.

Poterba, J. M. (2001). Estate and gift taxes and incentives for inter vivos giving in the U.S. *Journal of Public Economics, 79*(1), 237–64.

Poterba, J. M., & Weisbenner, S. (2001). The distributional burden of taxing estates and unrealized capital gains at the time of death. In W. G. Gale, J. R. Hines, Jr., & J. Slemrod (Eds.), *Rethinking estate and gift taxation*. Brookings Institution Press.

Poterba, J. M., & Weisbenner, S. (2003). Inter-asset differences in effective estate-tax burdens. *American Economic Review, 93*(2), 360–365.

Reiter, M. (2004). *Do the rich save too much? How to explain the top tail of the wealth distribution*. Universitat Pompeu Fabra: Mimeo.

Rogerson, W. P. (1985). Repeated moral hazard. *Econometrica, 53*(1), 69–76.

Roine, J., & Waldenström, D. (2009). Wealth concentration over the path of development: Sweden, 1873–2006. *Scandinavian Journal of Economics, 111*(1), 151–187.

Rothschild, C., & Scheuer, F. (2011). Optimal taxation with rent-seeking. *Working paper* 17035, National Bureau of Economic Research May 2011.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies, 68*(1), 205–29.

Saez, E. (2002a). Optimal progressive capital income taxes in the infinite horizon model. *Working paper* 9046, National Bureau of Economic Research July 2002.

Saez, E. (2002b). The desirability of commodity taxation under non-linear income taxation and heteroge-
  neous tastes. *Journal of Public Economics, 83*(2), 217–320.

Saez, E., Slemrod, J. B., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax
  rates: A critical review. *Journal of Economic Literature, 50*(1), 3–50.

Salanié, B. (2011). *The economics of taxation* (2nd ed.). Cambrige, MA and London, England: The MIT Press.

Sandmo, A. (1975). Optimal taxation in the presence of externalities. *Swedish Journal of Economics, 77*(1),
  86–98.

Sandmo, A. (1998). Redistribution and the marginal cost of public funds. *Journal of Public Economics, 70*(3),
  365–382.

Scheve, K., & Stasavage, D. (2012). Democracy, war and wealth: Evidence from two centuries of inheritance
  taxation. *American Political Science Review, 106*(1), 81–102. Yale and NYU, mimeo.

Schmalbeck, R. (2001). Avoiding federal wealth transfer taxes. In W. G. Gale, J. R. Hines, Jr., & J. Slemrod
  (Eds.), *Rethinking estate and gift taxation.* Brookings Institution Press.

Scholz, J. K. (2003). *Wealth inequality and the wealth of cohorts.* University of Wisconsin: Mimeo.

Scholz, J. K., Seshadri, A., & Khitatrakun, S. (2006). Are americans saving optimally for retirement? *Journal
  of Political Economy, 114*(4), 607–643.

Seim, D. (2012). *Smart enough to evade*? On the Incidence of a Wealth Tax, Mimeo.

Severinov, S. (2006). Bequests as signals: Implications for fiscal policy. *Journal of Public Economics, 90*(10–11),
  1995–2008.

Slemrod, J. (1998). Methodological issues in measuring and interpreting taxable income elasticities. *National
  Tax Journal, 51*(4), 773–788.

Slemrod, J., & Kopczuk, W. (2002). The optimal elasticity of taxable income. *Journal of Public Economics, 84*(1),
  91–112.

Stiglitz, J. E. (1978). Notes on estate taxes, redistribution, and the concept of balanced growth path incidence.
  Journal of Political Economy, 86(2), S137–S150.

Stiglitz, J. E. (1985). The general theory of tax avoidance. *National Tax Journal, 38*(3), 325–37.

Tsoutsoura, M. (2011). *The effect of succession taxes on family firm investment: Evidence from a natural experiment.*
  Chicago Booth: Mimeo.

Villalonga, B., & Amit, R. (2006). How do family ownership, control and management affect firm value?
  Journal of Financial Economics, 80(2), 385–417.

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

Werning, I. (2011). *Nonlinear capital taxation.* MIT: Mimeo.

Wilhelm, M. O. (1996). Bequest behavior and the effect of heirs' earnings: Testing the altruistic model of
  bequests. *American Economic Review, 86*(4), 874–892.

Wolff, E. N. (1996). Discussant's comments on Douglas Holtz-Eakin, The uneasy case for abolishing the
  estate tax. *Tax Law Review, 51*(3), 517–22.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic
  Studies, 32*(2), 137–150.