Behavioral Responses and Design of Bequest Taxation

Maksym Khomenko Simon Schürz*

Abstract

This paper studies the optimal design of an intergenerational wealth tax, commonly represented by either inheritance or estate taxation. Depending on the tax design, old-age individuals can react with a number of responses, ranging from adjustments of wealth accumulation and inter-vivos gifts to changes in the distribution of inheritances among heirs. We leverage a unique and appropriate setup of Swedish inheritance taxation and rich administrative data. To understand individual responses to alternative tax schemes, e.g. adjustments in wealth accumulation, bequest distributions, and the resulting welfare changes, we estimate a comprehensive structural model of wealth accumulation and bequest decisions in old age. We find that comparable inheritance and estate taxes result in sizable, but similar distortions to wealth accumulation and bequest distributions. By limiting strategic avoidance to wealth adjustments, estate taxation outperforms inheritance taxes in terms of tax revenues. Our model enables policymakers to design an intergenerational wealth tax that balances distortions, progressiveness, tax revenue and tax incidence according to their social welfare functions.

Keywords: bequest taxation, bequest motives, lifecycle model, tax avoidance

JEL classification: H24, H26, D14, D15, D64

* *Maksym Khomenko*: Department of Economics, University of Gothenburg, maksym.khomenko@economics.gu.se; *Simon Schürz*: Department of Economics, University of Gothenburg, simon.schurz@economics.gu.se. We are thankful to Raj Chetty, Steven Durlauf, Mikael Lindahl, Lee Lockwood, Christopher Taber, Christopher Tonetti, Emmanuel Saez and participants of the HCEO Summer School for Social Inequality at the University of Chicago 2016, 2017 and seminars at the University of Gothenburg for valuable comments. Maksym Khomenko acknowledges financial support from the Nordic Tax Council.

1 Introduction

Inherited wealth plays a key role in the intergenerational persistence of wealth inequality across families. In an active policy debate, economists and policy-makers are exploring whether and how to promote social mobility by taxing estates or bequests.[1] While a large body of literature documents the intergenerational and across-family consequences of inherited wealth (Boserup *et al.*, 2016; Elinder *et al.*, 2016; Adermon *et al.*, 2018), the microeconomic behavior of old-age individuals, from here on referred to as *donors*, with respect to bequests and their response to taxation is poorly understood. This paper studies the outcomes of different intergenerational wealth tax designs, taking into account a variety of taxpayer reactions.

There are two types of intergenerational wealth transfer taxes commonly adopted around the world: estate and inheritance taxation. In the case of the estate tax, the base of taxation is the terminal wealth of a deceased individual, whereas inheritance tax is levied on the individual bequest that each heir receives. In this case, family structure affects the optimal decisions of donors to a larger degree. Therefore, fundamental differences in tax design may lead to different behavioral responses and welfare implications. When donors choose how to optimally transfer wealth to a heterogeneous set of heirs, they can have multi-dimensional responses to intergenerational wealth taxation. For instance, old-age individuals may react by altering the wealth accumulation or use inter-vivos gifts to fully use exemption levels or fall into lower marginal tax brackets (Joulfaian, 2006; Kopczuk & Lupton, 2007; Glogowsky, 2016). If the tax is levied on heirs, donors may decide to change the distribution of individual bequests to their offsprings. Adjusting the share of the terminal estate that an heir receives can position that individual bequest at a lower marginal tax rate.[2]

The trade-off that old-age individuals face under bequest taxation can be characterized as a *donor trilemma*. Essentially, old-age individuals gain utility from current consumption, the amount of total bequests and how the estate is split. An intergenerational wealth tax triggers a trade-off between these three factors as it reduces the total value of after-tax bequests to heirs. Depending on how donors address the trade-off, their responses affect tax revenues, the values, and distribution of transferred wealth and donor utility. Additionally, responses to specific tax

[1]Currently fifteen US states collect an estate tax in place and six states tax inheritances. Maryland and New Jersey have both systems. In Europe, bequest taxation is in place in Denmark, France, Spain, Germany and Finland (inheritance tax) and the UK (estate taxation). A large number of European countries, such as Sweden, Norway, Austria, Hungary, and Portugal as well as several US states have repealed bequest taxation in the last 20 years with ongoing debates about re-introduction.

[2]Additionally, in Sweden (1992-2004), bequests could be classified into lower tax brackets if heirs decided to cede part or all of the inheritance directly to their offsprings. We discuss this particular tax avoidance strategy in detail in Section 2C.

schemes interact with individual donor heterogeneity, such as family structure, age and initial wealth. The choice of tax design and taxpayer reactions to intergenerational wealth taxation are therefore inevitably related.

In this paper, we estimate a structural model of donor decisions to document how the complex nature of behavioral responses to bequest taxation plays out under a number of alternative policy designs. More precisely, using detailed data on bequests, wealth, family structure and characteristics of decedents and heirs in Sweden from 2001 to 2004, we estimate the dynamic model of donor decisions, which involves wealth accumulation, inter-vivos gifts, and end-of-life bequests. There are a number of reasons why this structural approach is appropriate and even necessary in this case. First, to define optimal policy design, it is required to understand the outcomes of policies that have not been observed in real life. By estimating policy-invariant individual preferences, the lack of sufficient variation in observed tax designs can be addressed. Second, the model uncovers the interplay between multiple responses to taxation, rather than pinpointing overall or partial elasticities to the tax, e.g. through bunching estimators (Saez, 2010; Glogowsky, 2016; Escobar *et al.*, 2019). Third, we can inform policy-makers by studying counterfactuals policy designs along several dimensions including wealth accumulation, bequests, and tax revenues.

The model allows obtaining the fundamental parameters that govern the donors' dynamic trade-off between consumption, the total amount of bequests and the split among heirs. Each period, an old-age individual decides on a fraction of wealth to be consumed or transferred to descendants as gifts. The remaining wealth is conserved for the next period, when, depending on whether the old-age individual survives, it is either bequeathed or subject to the same choices. The donor also anticipates that in the case of death, she/he will choose on a split of the terminal wealth among heirs. Finally, the utility from bequeathing is derived from the total after-tax bequest value and the way bequests are split among the potential heirs.

To understand how donors allocate bequests to heirs, it is required to identify policy-invariant bequest preferences. In other words, the model needs to capture the donor's motivation to give a certain split of the terminal wealth to a specific heir in the absence of any taxation. To do so, we exploit variation in family structure, wealth and the presence of a ceding rule in the Swedish inheritance tax. This rule allows heirs to transfer all or part of a bequest to direct offsprings on receipt. By doing so, an heir can minimize the tax bill because each recipient of a cede can again make full use of individual exemption levels.[3] We show that a subgroup of donors, whose heirs can all cede and potentially avoid the entire tax bill, do not distort their bequest distribution

[3]Ohlsson (2007) and Escobar *et al.* (2019) document this incentive for Swedish heirs and shows that it is a widely used practice.

away from their true preferences as a reaction to taxation.

Following the conventional approach in the literature, we estimate the structural model in a two-step procedure (French, 2005; Lockwood, 2012; Blundell *et al.*, 2016). First, policy-invariant donor preferences for splitting terminal wealth among heirs are recovered by exploiting variation in family structures and characteristics of donors and heirs in a subsample of decedents. The obtained parameters are then entering a second stage, in which we estimate a dynamic life-cycle model à la Blundell *et al.* (2016) and Lockwood (2018).

Using the resulting parameter estimates, counterfactual wealth paths and bequests are simulated under various tax schemes. We find that, compared to the absence of any policy, donors accumulate significantly lower levels of wealth in old age when intergenerational wealth transfers are taxed. When taxes are progressive, this effect is driven mainly by individuals at the upper tail of the wealth distribution. Consequently, terminal estates are lower and large masses in the bequest distribution are allocated below the kinks of the marginal tax rates. Estate and inheritance tax schedules perform similarly in terms of distortions when marginal tax rates are comparably progressive. Importantly, estate taxes lead to higher tax revenues because they limit donor responses to the adjustment of terminal wealth. Under inheritance taxation, donors can react to the policy multi-dimensionally by adjusting each individual bequest. Generally, we find that the more flexible the policy is in allowing the donor to react based on the family structure, the higher are behavioral responses that distort wealth and bequest distributions as well as tax revenues.

This paper contributes to the literature on bequest taxation in several ways. Broadly speaking, this paper is related to a number of papers studying the life-cycle behavior of old-age individuals that among others involve wealth accumulation and bequest decisions (French, 2005; Laitner *et al.*, 2018; Lockwood, 2012, 2018). Our contribution is to propose a novel model to analyze inheritance tax designs under multiple behavioral responses and several dimensions of policymaker objectives. The comprehensive structural model covers taxpayer reactions from adjusting the wealth accumulation (Slemrod & Kopczuk, 2000; Joulfaian, 2006; Kopczuk & Lupton, 2007), inter-vivos gifts (Joulfaian, 2005; Ohlsson, 2011) and strategic changes in individual bequests. It allows studying consequences of such responses on wealth holdings, bequest distributions, and government tax revenues. An important feature of the institutional setting of Sweden is its generous social security system, which includes elderly and health care. It allows us to recover behavioral patterns which are distorted by precautionary saving behavior to a very minor extent. Therefore, this institutional set-up is particularly suitable to study this question in comparison to, for instance, the US where precautionary saving motives must be an important determinant of the end-of-life wealth decisions (Lockwood, 2018).

To our knowledge, this paper is the first to estimate such a comprehensive model of bequeathing with detailed micro-level data. In particular, the structural empirical approach is crucial to overcome the complexity of the problem, to deal with limitations of data availability and to ensure the possibility of studying these policy counterfactuals of interest. In particular, it addresses the common problem of non-identification of policy-invariant bequest preferences due to the lack of micro-level data on bequests when no taxation is in place. By providing a micro-level analysis of old-age individual behavior, we complement macro evidence on life-cycle models (De Nardi, 2004; Piketty & Saez, 2013; De Nardi & Yang, 2014) and reduced form evidence on bequest distribution (Light & McGarry, 2004; Erixson & Ohlsson, 2014; Escobar *et al.*, 2019).

More generally, we contribute to the literature on intergenerational wealth taxation by empirically studying the equity-efficiency trade-off under taxpayer responses (Piketty & Saez, 2013). We also touch upon research on bequest motives (Barro, 1974; Becker & Tomes, 1979; Behrman *et al.*, 1982; Cox, 2003; Arrondel & Masson, 2006; Lockwood, 2012, 2018) by incorporating decedent's altruistic and equality preferences.

The remainder of the paper is structured as follows: Section 2 describes the data, the Swedish institutional background and provides a descriptive analysis of bequest distributions. Section 3 discusses the structural model and Section 4 presents the estimation strategy. Section 5 introduces the counterfactual analysis and reports results. Section 6 concludes.

2 Data, Institutional Environment and Sample Selection

A Data sources

The study draws on a population-wide dataset on all bequests in Sweden between 2001 and 2004 provided by the Swedish Central Statistical Bureau (SCB). This so-called *Belinda* Population Database is a complete dataset of inheritances from 2001-2004 including an identifier for the deceased, the value of the terminal estate, the individual bequest, tax payments, and identifiers and characteristics of the heirs.[4] The bequest database is merged with detailed registry data to obtain background information on donors and heirs, such as labor market status, wealth, demographic characteristics, education, and income. The Swedish Multi-generational Registry is used to identify relationships between decedent and heirs and between heirs. To proxy expected conditional survival probabilities, we use life tables provided by SCB.

[4]See Elinder *et al.* (2014) for a detailed description of the database.

B Institutional details

Bequest taxation has had a long-standing tradition in Swedish tax policy. From 1885 to 2004 intergenerational wealth transfers were taxed at heir level (Henrekson *et al.*, 2014). After peaking in 1970, tax rates decreased steadily until the abolition, which was motivated by high administrative costs compared to small revenues and long-lasting opposition by entrepreneurial interest groups. While fiscally not important, the main purpose of the tax was the reduction of intergenerational transfers on the upper end of the wealth distribution. To adhere to the ability-to-pay principle of taxation, the scheme was designed in a progressive fashion (Kendrick, 1939).

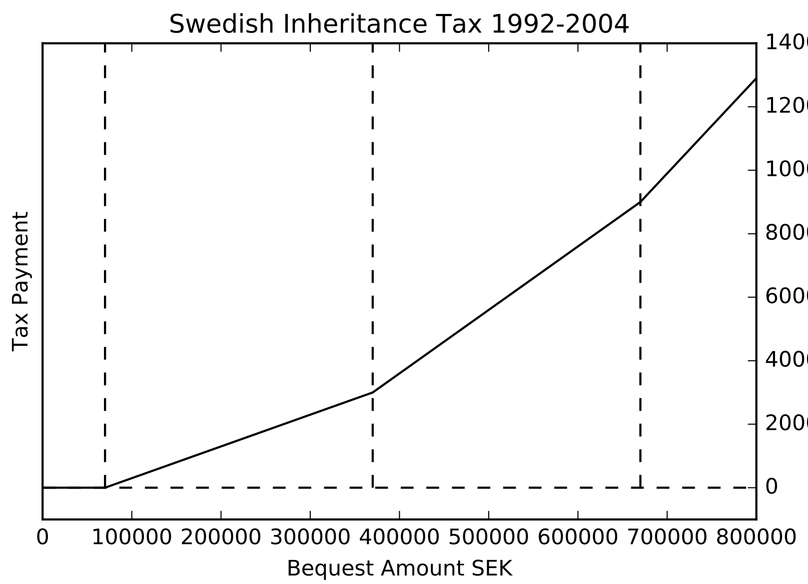
The inheritance and gift ordinance (*Lagen om arvs- och gåvoskatt*) stipulates the legislation for intergenerational wealth transfers. If an individual passes away, the decedent's estate is documented in the inventory estate report, which contains real and financial assets, private insurance, consumer durables, and debt. If the decedent is a surviving spouse it may include also part of the spouses' wealth (*giftorätt*).

Inheritance rules define the default succession order and distribution of bequests. If the decedent is survived by his/her spouse, she/he inherits the entire estate, except if the decedent has children from a different partner. The spouse has free disposal of the inheritance but cannot alter the bequest distribution set by the decedent. Such inheritances from previous decedents are separately marked in our data and we will consequently exclude spousal bequests from the analysis. Otherwise, the first non-empty parentelic group inherits equal splits of the final estate. Adoptive children are equal to biological offsprings before the law. Further groups are considered only if there are no heirs in the previous group. Any inheritance intended to a minor under 18 is directed towards the legal guardian of the child. Therefore we will focus on heirs older than 18 years.

A stipulated will can redefine the order and distribution of bequests with the limitation that a fraction, 50% of the inheritance without a will, is reserved for direct descendants (*Laglotter*). This puts institutional non-binding boundaries on how terminal wealth can be split among heirs. Clearly, wills are also set up for other purposes than unequal splits, for instance, the inclusion of further heirs or specific property transfer. A will is, therefore, a necessary, but not sufficient condition for unequal bequests. Gifts and insurance claims are included in the legislation. In our study period, they were generally taxed independently of their intergenerational character, but gifts 10 years prior to the death were taxed jointly with inheritances under the *summation rule*. This measure was introduced to counteract tax avoidance. Furthermore, gifts represent a way to transfer wealth to a set of heirs unequally in the absence of a testament. If an heir receives an inheritance, she/he can decide to cede part of or the total amount of the value to direct heirs, e.g. grandchildren of the decedent. Both the sender and the recipient of the cede can make full use of the individual inheritance tax exemption. This practice was widely used as a legal form of tax avoidance (Ohlsson, 2007; Escobar *et al.*, 2019).

The tax scheme in our study period is made up of three brackets. The details of tax brackets are displayed in Table 1. Inheritances to the first parentelic group (except spouses) over 70.000 SEK were taxed with 10%, followed by 20% and 30% rates for bequests exceeding 300.000 and 600.000 SEK. We focus on the tax scheme for most direct descendants, further schemes for relatives, friends and institutional recipients differ in exemption values, but not in marginal tax rates. Figure 1 depicts how the changes in marginal tax rates result in kinks in the overall tax scheme.

Figure 1: Kinks in the Swedish inheritance taxation for parentelic group 1



Notes: Tax scheme for the first parentelic group (spouses, children, grandchildren) with an exemption level of 70.000 per heir. Kinks of the marginal tax rate at SEK 370.000 and 670.000.

Table 1: Children, Spouses, Grandchildren: 1992-2004

Tax bracket, SEK	Tax (lump sum + tax rate)
0 - 300.000	0 + 10%
300.000 - 600.000	30.000 + 20%
> 600.000	90.000 + 30%
Basic Exemptions, SEK:	Amount
Children:	70.000
Gifts:	10.000

Notes: This tax scheme represents the marginal tax rates and exemption levels for the first parentelic group, including an annual inter-vivos gift exemption of SEK 10.000.

C Potential responses under Swedish inheritance taxation

The Swedish inheritance tax, in its 1992-2004, allows for the donor to respond to the tax in multiple ways. Old-age individuals can adjust terminal wealth, such that the total tax burden for heirs is reduced or the intended after-tax value of bequests is conserved. Changes in terminal wealth are achieved either by higher or lower levels of consumption or by wealth transfer by gifts, which have a yearly tax exemption of 10,000 SEK. The sign of the wealth adjustment is theoretically ambiguous and depends on the utility weights that are placed on consumption and bequeathing. The optimal response to a tax with a change in the level of wealth, however, requires high planning efforts and is costly due to uncertainty regarding the timing of death.

To give an example for a possible wealth response to taxation, consider a donor with wealth SEK 160.000, who wants to equally bequeath her two children. She can gradually decrease her wealth via gifts or consumption, such that the terminal wealth of 140.000 positions both individual bequests (SEK 70.000) below the first tax bracket.

Alternatively to changes in terminal wealth, an inheritance tax, which is based on the heir level, allows the donor to respond by adjusting individual bequests. By stipulating a will, a donor can change the relative shares to heirs in order to place one or several individual bequests at or under the thresholds of tax brackets. While this strategy may limit distortions in the wealth accumulation, it requires the donor to depart from the preferred split of bequests.

For example, a donor who wants to leave SEK 50.000 and 100.000 to her/his two children, may want to switch to a SEK 70.000/80.000 split in order to save SEK 2.000 in taxes.

Last but not least, ceding is a modification of the bequest distribution conducted by the heirs. Under the Swedish inheritance law, a bequest can be dissipated downstream within the family line. Ceding is therefore limited to heirs who have descendants of their own. In particular, if ceding can eliminate the entire tax bill, donors may not exhibit any response to taxation.

For example, an heir that receives a bequest SEK 200.000 may want to cede SEK 70.000 to each of her two children (the donor's grandchildren), such that, due to the individual exemption levels, all three individual bequests then fall below the first tax-bracket Table 2 summarizes all potential donor responses and alternatives for legal tax avoidance.

If the bequest taxation takes on the form of an estate tax, adjusting the terminal estate through wealth decumulation and inter-vivos gifts are the only available responses to old-age individuals. In this paper, we focus on wealth adjustment through consumption, gifts, and changes to individual bequests. As we document, these responses are widely used in Sweden 2001-2004 and are available to donors in most countries with bequest taxation. The ceding rule is very specific to the Swedish context and plays a key rule in identifying the interaction of the remaining responses to inheritance taxation. For simplicity and due to extremely low incidence and special tax exemptions, we exclude the strategic component of bequeathing to non-children heirs and assume a constant share per donor that is left to these types of heirs.

Table 2: Bequest taxation: Strategic responses and (legal) incentives for avoidance

	Estate Taxation	Inheritance Taxation
Wealth accumulation	✓	✓
Within family distribution	—	✓
Gifts	✓	✓
Cedes	—	(✓)

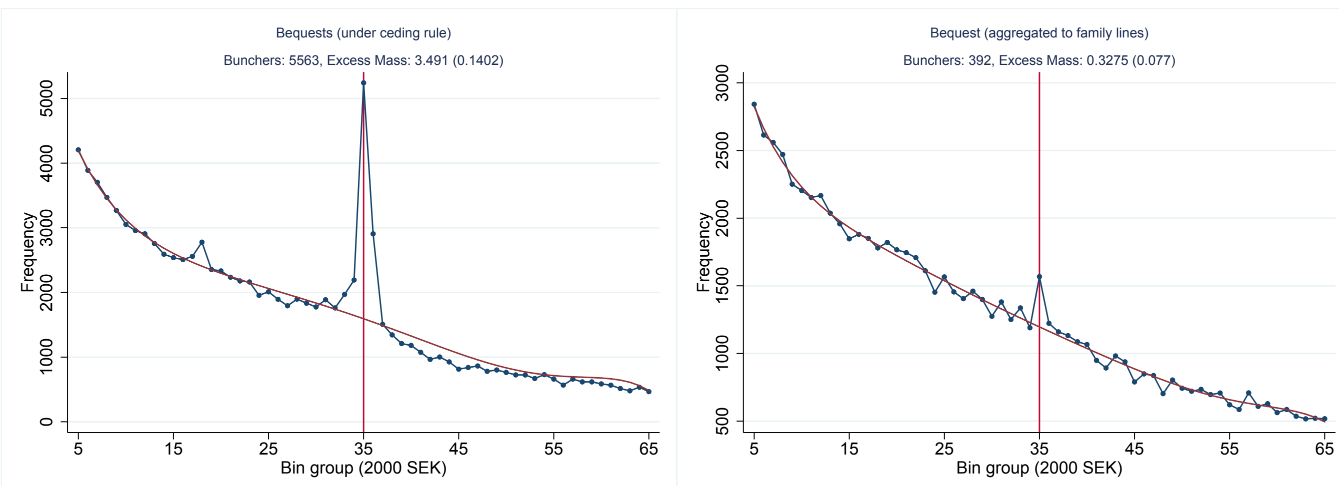
Notes: The availability of strategic responses of donors to the two main alternatives for bequest taxation, inheritance and estate taxes, depends on the definition of the tax-base. Taxation at the heir level allows for additional reactions by adjusting the distribution of bequests among heirs or, if the legislation provides this possibility, by ceding.

D Undistorted bequest distribution

A crucial step to study responses to bequest taxation is the identification of preferences that govern the donor's decision to allocate individual bequests to heirs, i.e. how to split the terminal wealth among potential heirs. Due to the ceding rule in Swedish institutional set-up, heirs could transfer part or all of their bequests to direct descendants upon receipt and thereby minimize the tax bill. This final bequest distribution in the data is significantly distorted by taxation and does therefore not allow to assess policy-invariant bequest preferences of donors. By aggregating bequests over family lines that originate at the children of the donor, we can eliminate any such distortions from the ceding rule. The resulting distribution reflects true bequest preferences for children's family lines, if i) all children of the donor had the possibility to cede, ii) these children were, on their own, able to eliminate the entire tax bill of the donor through ceding and iii) after aggregating over family lines, the sample distribution of bequests is smooth, in particular at the kinks of the tax schedule. While conditions i) and ii) ensure that there are no strategic responses of the donor with respect to a subset of heirs, the smooth distribution proves that donors' choices are not distorted by any other form of bequest adjustment.

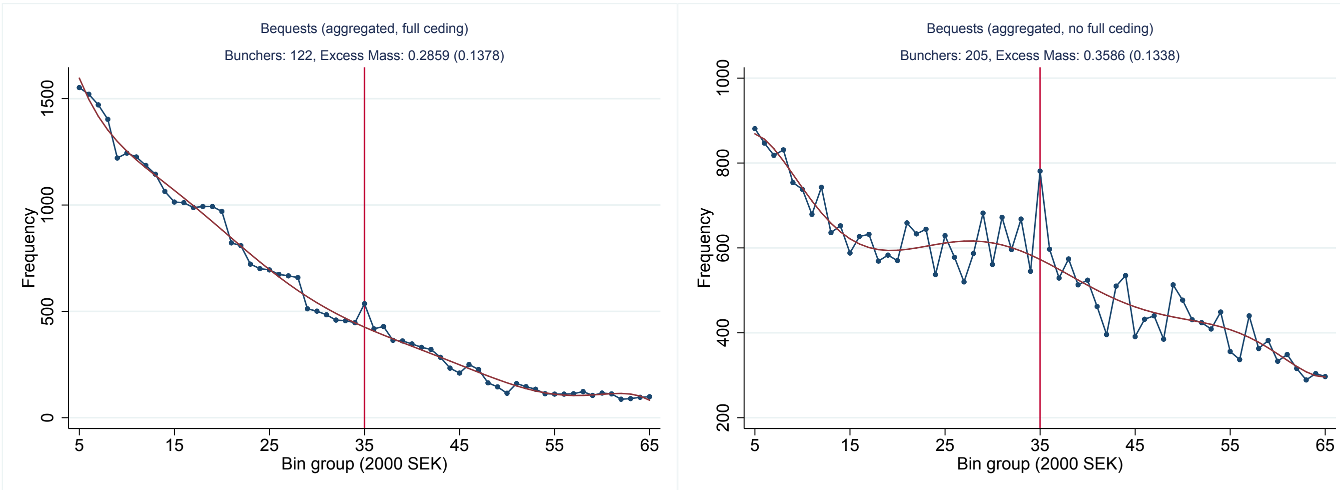
Figure 2 shows that before the aggregation, there is a massive bunching of individual bequests at the first tax kinks (70,000 SEK) of 3.5 times the average height of neighboring distribution bins. After aggregating, it becomes clear that ceding makes out 90,6% of this excess mass. According to the set of potential legal avoidance mechanisms, the remaining bunching at the kink could originate from wealth adjustment or distortions in the optimal bequest distribution of donors.

Figure 2: Bequest distributions: Raw data vs. aggregated over family lines



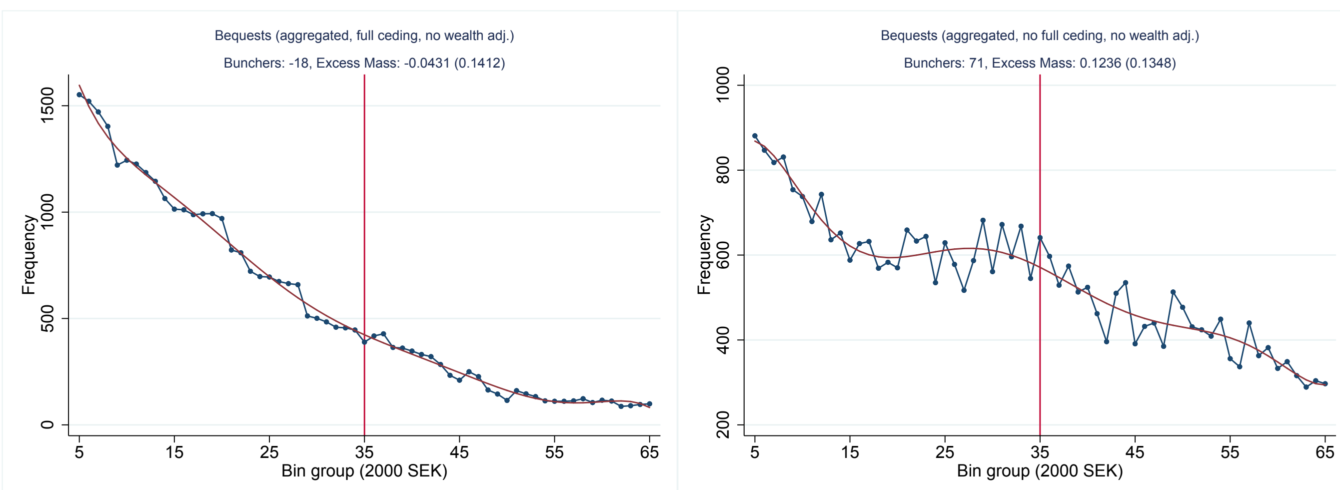
Notes: The graph on the left hand side plots the raw bequest distribution. Bunching at the first kink of the marginal tax rates is mostly attributable to the ceding rule. When aggregated over family lines (to reverse the ceding rule) a much smaller, yet significant excess mass at the kinks remains.

Figure 3: Aggregated over family lines and full ceding possible



Notes: The graph on the left hand side plots bequests aggregated over family lines for donors, whose children can fully eliminate the tax bill through the ceding rule. The complementary subgroup of donors, whose children cannot fully eliminate the tax bill through ceding is shown on the right-hand side.

Figure 4: Aggregated over family lines, full ceding possible and estate Adjustment



Notes: The graph on the left hand side plots bequests aggregated over family lines for donors, whose children can fully eliminate the tax bill through the ceding rule. It further excludes donors whose terminal wealth divided by the number of children results in bequests precisely on the tax kink ($\gtrsim$ SEK 500) in case of equal splitting. The complementary subgroup of donors, whose children cannot fully eliminate the tax bill through ceding and do not adjust wealth is shown on the right-hand side.

Figure 3 shows that donors, whose heirs do not fulfill conditions i) and ii) (right graph), are concentrated over-proportionally at the tax kink compared to the bequest distribution of heirs, who can eliminate the entire donor tax bill through ceding (left graph). The bunching excess mass of 0.358 (0.134) for heirs, who cannot fully cede exceeds the value of their ceding counterparts of 0.286 (0.138). Furthermore, the right bequest distribution exhibits non-smoothness at several intervals.

Finally, the left graph of Figure 4 provides evidence that any remaining bunching of donors, who can fully cede, is entirely explained by wealth adjustments. These are decedents, whose estates divided by the number of children falls within a close range of the tax kink. Since wealth adjustment will be captured in the donor's wealth accumulation part of our model, the bequest distribution aggregated over family lines is undistorted. The right graph of figure 4 shows that the distribution for heirs who cannot fully cede is not smooth, even after accounting for strategic wealth adjustments by donors.

In the following, we proceed by estimating *true* undistorted bequest preferences on a subsample, which excludes those donors, for which at least one child is unable to cede to a direct offspring. Notice that while we estimate undistorted, optimal preferences with this selected sample, the estimation of the dynamic problem including wealth accumulation and bequest shares as well as all counterfactual simulations are conducted on the full sample of aggregate family line bequests.

E Samples

After restricting the universe of decedents to surviving spouses over the age of 65 (to capture intentional bequests) with up to four children, the sample includes 61.044 donors. Summary statistics of decedents and heirs are presented in Table 3. Estates are on average 226.519 SEK while bequests received by a child's family line amount to 115.068 SEK. Around 3% of decedents distribute the inheritance unequally across the children's family lines (within family standard deviation greater than 1000 SEK). A fifth of decedents have stipulated a will and about 9% have transferred wealth to their heirs via inter-vivos gifts within ten years prior to their death.

The subsample used to recover bequest preferences from the undistorted distribution of inheritance is restricted to old-age individuals, of whom all heirs had the possibility to fully cede and eliminate the entire tax bill. We identify these heirs by matching them to their direct offsprings via the Swedish Multi-generational Register. Clearly, this subsample is selected and shows on average smaller estate and bequest values. This is due to two main reasons. First, rich donors would require the heirs to have a large number of descendants to enter this subsample and second, heirs with numerous direct offsprings may receive higher values of inter-vivos gifts. The full and ceding subsample are however well-balanced with respect to all other covariates.

Table 3: Summary statistics

	Full Sample		Full Ceding Sample (preference estimation)	
Heir Characteristics	Mean	Sd	Mean	Sd
Bequest	121087.8	(298984.7)	59822.0	(73320.8)
Share	0.506	(0.260)	0.561	(0.284)
Heir female (0/1)	1.495	(0.500)	1.517	(0.500)
Education of heirs	11.52	(2.162)	11.35	(2.096)
Married heir (0/1)	0.586	(0.493)	0.692	(0.462)
Age of heir	54.11	(9.577)	56.46	(7.936)
Income of heir	2024.9	(2746.6)	1922.0	(2585.1)
Observations	123659		58460	
Donor Characteristics				
Estate size	241106.6	(497310.1)	107260.2	(93262.0)
Will (0/1)	0.195	(0.396)	0.167	(0.373)
Inter-vivos transfers (0/1)	0.0942	(0.292)	0.0846	(0.278)
Donor female (0/1)	1.664	(0.472)	1.716	(0.451)
Education of donor	11.57	(3.826)	11.72	(4.051)
Age at death	84.10	(9.065)	85.91	(7.430)
Wealth at death	597325.9	(2076097.2)	296266.6	(362468.1)
Number of children	2.677	(0.747)	2.610	(0.725)
Unequal split of bequest	0.0315	(0.175)	0.0185	(0.135)
Observations	64707		33970	

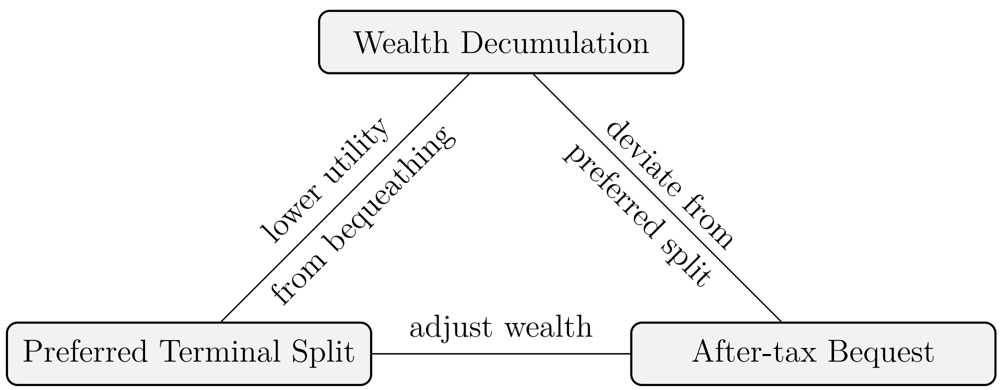
Notes: Education denotes years of schooling, coded consistently with Holmlund *et al.* (2011): 9 years for primary school, 9,5 for post-primary school, 11 for short high school, 12 for long high school, 14 for short university, 15.5 for long university and 19 for PhD university education.

3 Wealth accumulation and bequest model

We now introduce a dynamic life-cycle model of a retired individual who plans the wealth path and end-of-life bequests in the event of death in the spirit of De Nardi *et al.* (2010) and Lockwood (2018). In an environment with a strong welfare state, old-age individuals jointly optimize utility from consumption and bequeathing, i.e. warm glow. The latter is a function of both the total value of bequest and the split among heirs. The introduction of a tax on intergenerational wealth transfers forces the donor to a trade-off between the two sources of utility. In the absence of any response, the tax reduces after-tax bequests. Depending on the tax design, the donor can change the wealth accumulation process or the way terminal wealth is split among heirs to re-optimize her utility.

Figure 5 shows this 'trilemma' of an old-age individual when the bequests are subject to taxation. If the individual decides to keep the wealth path and bequest distribution unchanged, the tax will reduce utility from post-tax bequests value. If the preferences for a specific wealth path and the after-tax value of bequests dominate the desire to split terminal wealth in a specific way, then a deviation from the initially preferred split delivers a tax-relief. Finally, if the preferred split and after-tax bequests ought to be undistorted, wealth adjustments or inter-vivos gifts are the tools to maximize utility.

Figure 5: The "Bequest Trilemma" under bequest taxation



Notes: The nodes define the three objectives of the donor, when making decisions in old-age on wealth accumulation and bequeathing. The edges show the individual's reaction in order to keep the objectives in the two connected nodes constant when a bequest tax is levied on inter-generational transfers. As the preferred terminal split does not play a role under estate taxation, the trilemma is transformed into a dilemma under such a tax scheme.

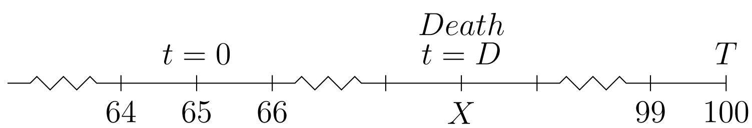
Consequently, individual behavior is summarized by three groups of parameters. The first group of parameters describes the trade-off outlined above and represents the main object of interest in this paper. More precisely, these parameters represent the weight an old-age individual places on after-tax bequests and a bequest split among heirs relative to consumption. The second group of parameters is solely related to the donor preferences over how bequests are distributed to children, conditionally on donor and heir characteristics. The third group includes parameters of a standard life-cycle model without labor choice, namely a discounting rate and risk aversion.

Essentially there are three building blocks to the model: A dynamic problem of wealth accumulation in old age, utility from bequeathing and policy-invariant preference for giving individual bequests to the potential set of heirs.

A Life-cycle problem of old-age individuals

Consider an old-age individual living from the time of retirement at 65 ($t = 0$) and up to a age 100.

Figure 6: Dynamic Model Time-Line



Each period individuals maximize expected discounted utility of consumption and bequests by choosing the consumption path $\{c\}_{t=1}^{T}$ and the vector of individual bequest split $s = \{s_1 ... s_j\}$ to children $1...J$ if the donor dies in the following period. This implies that the model assumes that individuals can adjust the allocation every new period costlessly. By consumption in this model, we mean either actual consumption or inter-vivos gifts to children using the yearly tax exemption of SEK 10.000 or higher.[5] The value of holding individual wealth W at $t+1$ depends on the probability of surviving $1 - D$ and the utility of bequeathing B in case of death D. The value function of this dynamic programming problem at time t consists of the utility from consumption and gifts and expected value function at time $t + 1$, discounted by a factor δ.

$$V_{it}(W_{it}) = u(c_{it}) + \delta \left[\overbrace{D_{it} B(s, W_{it+1})}^{\text{Bequest Utility}} + \overbrace{(1 - D_{it}) E_{it}[V_{it+1}(W_{it+1})]}^{\text{Utility of Survival}} \right] \tag{1}$$

We assume a CRRA utility function of consumption with risk parameter η:

$$u(c) = \frac{(c)^{1-\eta} - 1}{1 - \eta} \tag{2}$$

Old-age individuals in the model hold rational beliefs about survival probabilities in line with mean death probabilities conditionally on age and gender (D).[6] The only state variable directly

[5]Any inter-vivos gifts above the exemption within 10 years prior to death is taxed as an inheritance and should, therefore, be inelastic to changes in the tax rate. Uncertainty over future wealth shocks makes transfers above the basic annual exemption for gifts costly for the bequest donor.

[6]We ignore the marginal endogeneity of survival with respect to taxes, as documented in Kopczuk & Slemrod (2003); Eliason & Ohlsson (2013).

affected by the individual's choice at time t is the next period wealth W_{t+1}, which is either kept as wealth or transformed to the terminal estate if the individual dies. Wealth evolves an auto-regressive process with normally distributed random shocks v. The donor's consumption possibility is restricted to her wealth, which implies that borrowing is not allowed in the model: $c_t \in [0, W_t + y_t]$, where y_t denotes yearly income. The model regards decisions in old age, for which the income data show fairly stable income, mainly consisting of pension income and social benefits. Therefore donors are assumed to expect to receive their mean observed income in future periods. The formulation of wealth evolution leaves the possibility of a negative realization of wealth. We allow for a minimalistic uncertainty structure since we do not observe large wealth fluctuations in the data.

$$W_{i,t+1} = W_{it} + (y_{it} - c_{it}) + v_{it} \tag{3}$$

$$v_{it} = \mathbb{N}\left(0, \sigma_W^2\right) \tag{4}$$

B Bequest utility

Next, consider the bequest function B, which determines the interplay between wealth accumulation and bequeathing. We adopt the general functional form for bequest utility from Lockwood (2018), which generalizes the approaches used in the literature. It allows capturing a range of components of bequest preferences. For instance, it implies that bequest motives kick in after a consumption threshold c_b. Under this value, individuals do not leave bequests. If $c_b > 0$, bequests are a luxury good, over which individuals are less risk-averse than over consumption. λ_1 denotes the marginal propensity to bequest out of left-over wealth after consuming at least c_b.

This functional form is combined with a term with weight λ_2, which represents the donor's preference on how to split bequests among the potential heirs. Essentially, the model punishes the donor for deviating from preferred individual bequest shares $s^* = \{s_1^*, ..., s_j^*\}$. This vector s^* represents the policy-invariant preferences for bequeathing to specific heirs with a specific share of the terminal estate. Expressing individual bequests as shares of the terminal wealth implies $\sum_{j=0}^{J} s_j^* = 1$. As discussed above, the donor might deviate and choose a different vector $s = \{s_1, ..., s_J\}$ if it reduces the tax bill by placing individual bequests at lower marginal tax rates.

$$B\left(s, s^{*}, W\right)=\overbrace{\left(\frac{\lambda_1}{1-\lambda_1}\right)^{\eta} \cdot \frac{\left(\frac{\lambda_1}{1-\lambda_1} \cdot c_b+\sum_{j=1}^{J}(1-\tau_j) \cdot s_j \cdot W\right)^{1-\eta}}{1-\eta}}^{\text{utility from after-tax bequests}} \tag{5}$$
$$+\lambda_2 \cdot \overbrace{\left(g(s_1^{*}, ..., s_J^{*})-g(s_1, ..., s_J)\right)}^{\text{Disutility from deviating}}$$

Parameters λ_1 and λ_2 represent the weight of this pair of bequest utilities with respect to the utility from consumption and inter-vivos gifts (weight normalized to 1). As λ_2 is sensitive to the number of children, we allow it to vary with this characteristic.

C Bequest preferences

The finalize the model, the preferred split of the terminal estate s^{*} is specified as the bequest shares that a donor would choose in the absence of any intergenerational wealth taxation. We assume that each donor has a utility function over the bequest shares that her children $j \in 1, ..., J$ receive.

$$g=\sum_{j=1}^{J} \exp(\phi_j) \cdot \log(s_j) \tag{6}$$

The preference parameter ϕ_j determines how large the bequest to child j is relative to the siblings. However, ϕ_j does not include any concerns for inequality, which are captured by switching costs of leaving the default option of equal shares. Preference parameter ϕ_j is parametrized to be a linear function of observable heir characteristics $\phi_j=\alpha X_j$, where X_j contains age, gender and income. We denote the vector of preferred unequal shares as:

$$\hat{s}=\{\hat{s}_1, ..., \hat{s}_J\} \quad \text{with} \quad \hat{s}_j=\frac{\exp(\phi_j)}{\sum_{k=1}^{J} \exp(\phi_k)} \tag{7}$$

Motivated by the prevalence of exactly equal allocations in the data (the default split in the institutional set-up in the absence of the will), we assume that donors face fixed costs ψ of deviating from a vector of equal shares $1/J$. These fixed switching costs reflect both a preference for equality across children and material costs of deviation, such as writing a will. We assume a normal distribution with variance σ_ψ^2 and a linear parameterization of the mean: $\psi_i \sim N(\beta Z_i, \sigma_\psi^2)$. The set of donor characteristics Z_i includes a constant, years of schooling, age at death, estate value and dummies that control for the number of children of the decedent.

If the difference between utilities derived from the vector that maximizes g and equal shares $\frac{1}{J}$ is smaller than these switching costs (ψ), the individual remains with the default of the equal share allocation.

$$s^* = \begin{cases} \{\hat{s}_1, ..., \hat{s}_J\} & \text{if} \quad g\left(\hat{s}_1, ..., \hat{s}_J\right) - g\left(\frac{1}{J}, ..., \frac{1}{J}\right) > \psi \\ \{\frac{1}{J}, ..., \frac{1}{J}\} & \text{otherwise} \end{cases} \tag{8}$$

D Solution method

The model is solved using backward induction from a terminal age 100. The model has two choice variables: consumption and bequest shares. We discretize bequest shares depending on the number of children using steps of 5% in addition to equal split shares. For example, for a family with three children, equal split shares are 33%, which would not be covered by a 5% grid.[7] The discretization yields a number of bequest allocations depending on the number of children. For instance, a donor with two children has the following choices: (25%, 75%), (30%, 70%), ... , (50%, 50%), ... , (70%, 30%), (75%, 25%). Note that allocations, in this case, are limited to an individual minimum of 25% per child because of the legal restriction that each heir is eligible to a minimum of 50% of the default allocation.

Introducing both a discrete (shares) and a continuous (consumption) choice variable does not suffer the solution problem commonly associated with secondary kinks in the value function (Fella, 2014; Blundell *et al.*, 2016; Iskhakov *et al.*, 2017). The reason is that the time-invariant bequest utility is only a function of terminal wealth and bequest shares and can be pre-computed on a wealth grid. These pre-computed values are then used to solve the dynamic model with linear interpolation if the realized value of wealth is not on the grid point. Hence, the dynamic problem involves only one continuous choice variable conditionally on precomputed utilities of bequests for various wealth levels.

In the estimation, the state variable wealth is discretized and inter- or extrapolation is used when the state variable value is not on the grid. To integrate over wealth shocks, we employ Hermite quadratures. The solution method is roughly the same as in, e.g., French (2005); Lockwood (2018). The main difference is related to the structure of the bequest function. It includes additional parameters and introduces a richer structure of the donor bequest decision required to study the responses to bequest taxes.

[7] This discretization could be avoided with continuous shares. However, besides avoiding computational issues, it is not unlikely that individuals think in terms of such discrete shares, which is observed in the data.

4 Estimation

To estimate the model, we use the Simulated Method of Moments (SMM). It extends the minimum distance estimator to cases when a closed form solution of the problem cannot be obtained. The estimation of life-cycle style models is usually separated into two steps. In the first stage, all the parameters that can be identified without solving the model are estimated. In the second stage, these parameters are fed into the estimation of the remaining parameters, which requires solving the dynamic model. Such a two-step procedure allows reducing the computational costs of repeatedly solving the dynamic model to search for a large set of parameters.

In our model estimation, we use this two-step approach to first estimate bequest preference parameters that define the preferred vector of bequest shares. To identify these parameters, we leverage the *ceding rule*, which is a special institutional feature of Swedish inheritance tax. It provides a unique opportunity to identify these policy-invariant preferences without a need to observe bequests in the absence of taxation. The rule enables heirs to transfers part or all of the received bequest directly to own descendants, i.e. grandchildren of the deceased. Ceding is therefore only available to those heirs who have children of their own. As discussed in Sections 2D and 2E we use a subset of decedents for whom all children have the possibility to fully cede and avoid all taxes. As a result, bequest shares observed for old-age individuals in this sample represent their true preferences and allows us to recover the underlying parameters α, β and σ_ψ.

The estimates of step one are used in the second stage, i.e. the dynamic model. Here we recover the remaining parameters for the wealth accumulation and the bequest utility function $\lambda_1, \lambda_2, \eta, c_b, \sigma_W$. We use a discount factor δ from the literature since it is not well-identified separately from other parameters of interest. The remainder of this section describes the estimation of the first and second stage parameters.

A Bequest preferences - stage one

We start by estimating preferences for bequest shares. From equation (8), the probability of donor i to deviate from the equal allocation is defined as:

$$P_i = F\left(\frac{g_i\left(\hat{s}_1, ..., \hat{s}_J\right) - g_i\left(\frac{1}{J}, ..., \frac{1}{J}\right) - \psi_i}{\sigma_\psi}\right) \tag{9}$$

where $F(\cdot)$ denotes the normal cumulative density function. It describes the probability that individual's switching costs are lower than utility gains of leaving the equal share default. Consequently, the expected allocation of bequest shares to children is given by the weighted sum of unequal and equal share vectors.

$$\{s_1^*, ..., s_J^*\} = P_i \cdot \{\hat{s}_1, ..., \hat{s}_J\} + (1 - P_i) \cdot \left\{\frac{1}{J}, ..., \frac{1}{J}\right\} \tag{10}$$

There are two important clarifications regarding donor and heir characteristics used in this part of the model. Firstly, the matrix X does not contain a constant. The reason is that bequest preferences are modeled as shares, which implies that only characteristics of heirs in relation to each other matter. Secondly, Z contains dummies for the number of children to adjust each donor's switching cost to the family structure. We set a constant term in Z to 1 since it is also not identified separately.[8]

This first stage model implies that the vector of heir-level coefficients and the parameters of the distribution of switching costs are estimated jointly: $\varphi = \{\alpha, \beta, \sigma_\psi\}$. We use Generalized Method of Moments to find parameters that match the moments of the observed bequest distribution for the *ceding* subsample.

$$\varphi^* = \arg\max(m(\varphi) - \hat{m}(\varphi))' W (m(\varphi) - \hat{m}(\varphi)) \tag{11}$$

where $m(\varphi)$ are the observed moments of the data that we match, $\hat{m}(\varphi)$ denotes the corresponding moments generated by the model, W - optimal GMM weighting matrix. We match three groups of moments. Firstly, we match a percentage of non-equal splits by a number of kids and age bins $\{< 75, [75; 85), \geq 85\}$. Secondly, since the preferences over heirs are primarily informed by those who actually give unequal bequests, we match the distribution of shares for those heirs who deviated from the default.[9]

B Dynamic model - stage two

Upon estimating the first stage parameters, we proceed to recover the parameters of the life-cycle model in the presence of the Swedish inheritance tax scheme of 1992-2004. More precisely, the parameters that give utility weights to bequests, the parameter of standard deviation of the wealth process and a risk preference parameter: $\xi = \{\lambda_1, \lambda_2^{2kids}, \lambda_2^{3kids}, \lambda_2^{4kids}, c_b, \sigma_W, \eta\}$.

To identify the parameters of interest, we use a discount factor $\delta = 0.96$ in line with De Nardi (2004) and De Nardi & Yang (2014).[10] Probabilities of survival conditional on age and gender are recovered from the life tables provided by the Swedish Statistical Office (SCB). Due to

[8] A constant parameter of switching costs is not identified separately from the level of parameters in X.

[9] Equal splits also inform parameters since preferences over heirs also affect how large fixed costs should be to keep a default choice.

[10] These authors use this calibration for both Swedish and US data.

the evaluation of estate by the tax authority at below-market prices, each individual knows with certainty that the terminal wealth is reduced by a constant percentage before bequests are distributed.[11] The Simulated Method of Moments (SMM) estimator minimizes the distance between data and model-generated moments:

$$\xi^* = \arg\max(m(\xi) - \hat{m}(\xi))' W(m(\xi) - \hat{m}(\xi)) \tag{12}$$

where W is GMM optimal weighting matrix. To minimize the criterion function, we firstly run a global stochastic Covariance Matrix Adaptation Evolution Strategy (CMA-ES) optimizer from various starting values (Hansen, 2006). After convergence, we start the local derivative-free simplex optimizer from the best parameters generated by the global optimizer to refine the solution.[12]

5 Results

A Parameter Estimates

Table 4 reports the parameter estimates for the undistorted bequest preferences of decedents (estimation stage one). Donors exhibit preferences that reinforce existing differences in income. This behavior is consistent with a wide range of theories, such as evolutionary bequeathing, exchange motives or even the altruistic theory. The positive coefficient towards female partly reflects the size of the heirs family line.

[11] Actual asset evaluation (as a percentage of market value) depends on the asset type and therefore on the composition of wealth: 75% for real estate and stocks traded on main Swedish exchange lists and foreign exchanges, 100% for cash, inventories and debt, 30% for stocks on minor Swedish exchange lists and NASDAQ and 30% of book value for firms. Apartments' evaluation percentage depends on the net wealth of the housing society. For simplicity, we fix the evaluation percentage at the observed value in the data. Additionally, surviving spouses may hold some estate from the deceased spouse. While they are allowed to consume this wealth, it cannot be bequeathed and does therefore not affect any strategic behavior to avoid taxes.

[12] In fact, the concept of convergence of global optimizers does not exist. Therefore, by convergence, we mean the best value obtained from a stochastic pattern-based search algorithm from a fixed number of iterations.

Table 4: First Stage Model Parameter Estimates

	Coefficient	Std. Errors
Heir Characteristics (ϕ)		
Gender (0/1)	1.044	(0.213)
Age	0.001	(0.004)
Income	0.014	(0.022)
Distribution of Donors' Deviation Costs		
Constant	*1*	—
Gender (0/1)	-0.523	(0.008)
# Children = 3 (0/1)	0.0003	(0.329)
# Children = 4 (0/1)	-0.362	(0.109)
σ	1.383	(1.232)

Notes: Second column in brackets presents GMM asymptotic standard errors.

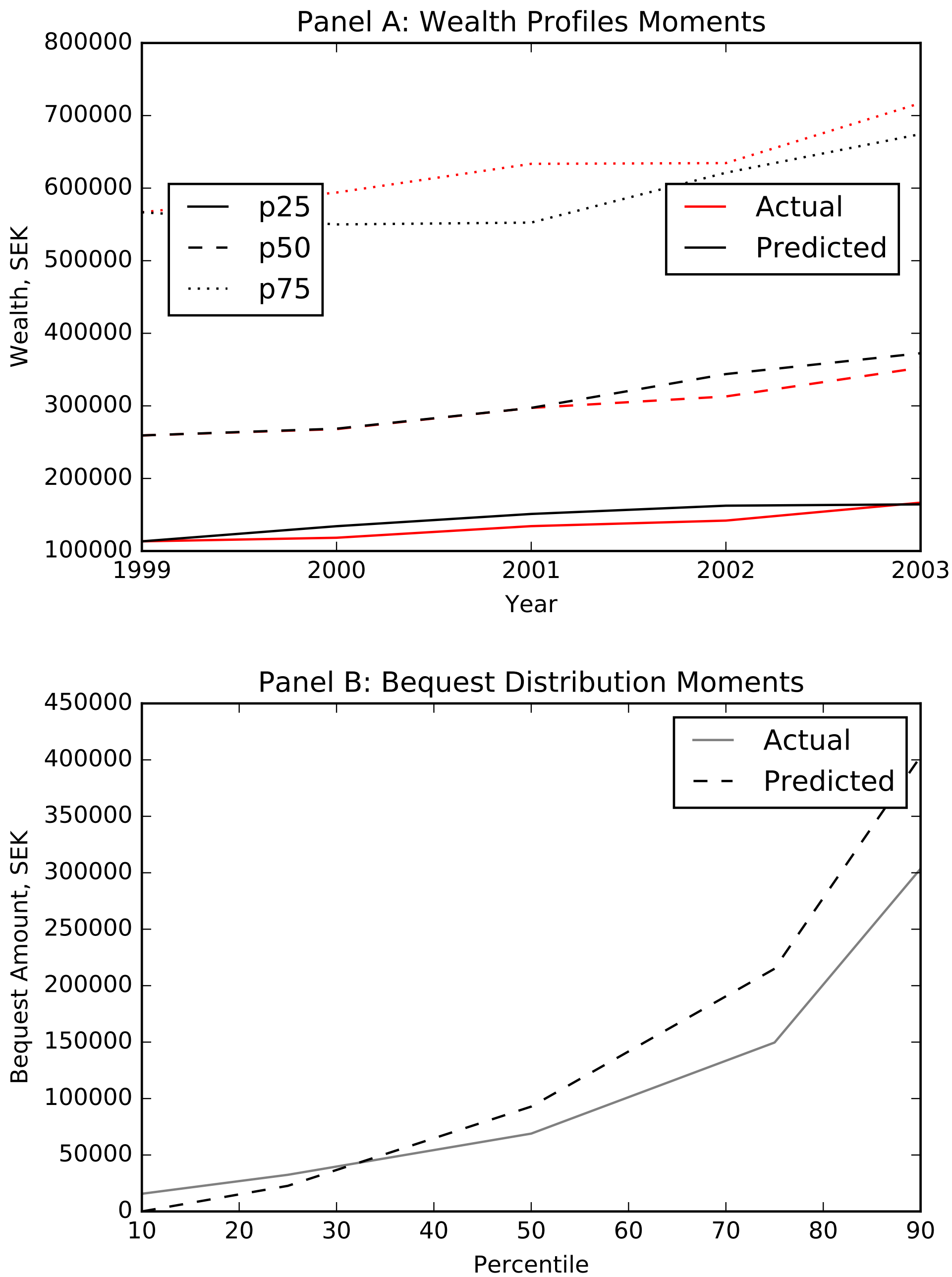
For the switching costs away from the default rule of equal shares, Table 4 shows plausible coefficients, that relate to either the monetary and cognitive costs of deviating or to a preference for equality. Male and older donors are more likely to stipulate a will to deviate from the default. Finally, Table 13 confirms that due to the flexible form of the utility function for bequest shares, the model fits targeted moments well.

Table 5: Main Model (Second Stage): Parameter Estimates

		Coef.	Std. Err
Propensity to bequeath	λ_1	0.99	(0.001)
	λ_2^{2kids}	< 0.001	(< 0.001)
Weight of the disutility term from deviation	λ_2^{3kids}	< 0.001	(< 0.001)
	λ_2^{4kids}	< 0.001	(< 0.001)
Std. Dev. of wealth shocks	σ	947.96	(14.43)
Consumption threshold for bequest motives	c_b	10765.49	(60.83)
CRRA risk aversion parameter	η	3.78	(0.27)

Notes: Second column in brackets presents GMM asymptotic standard errors explained in more details in Appendix C.

Figure 7: Model Fit



Notes: Actual and predicted moments of the wealth and bequest distributions. Predicted values result from simulating the wealth accumulation model with bequest utilities using the estimated parameters.

For the second stage parameter estimates, Table 5 reports the weights on bequest utility.

In line with Lockwood (2018), we find a large propensity to bequeath and a high risk aversion parameter (3.78). The threshold of minimum consumption at which bequest motives kick in is low at SEK 10.765 and partly reflects the inter-vivos gift tax exemption, which, if used, is part of the consumption. The variance of wealth shocks is small at 884. The model fit is shown in Figure 7. Overall, the model fits the moments reasonably well.

B Counterfactuals

To analyze behavioral responses to bequest taxation, counterfactual wealth paths and bequest distributions are simulated for alternative tax designs. It means that we fully simulate donor's decisions and responses to any given tax using our estimates of preference parameters and individual characteristics. The resulting wealth paths and bequest distributions for each alternative tax scheme represent the bases of our further analysis.

0. No taxation of intergenerational wealth transfers (baseline)
1. The 1992-2004 Swedish inheritance tax scheme, excluding the ceding rule
2. Estate tax with marginal tax rates, exemption levels and tax rate kinks comparable to (1.)
3. Estate tax with fixed exemption levels and kinks. Marginal tax rates comparable to (1.)

The no-tax case (0.) is the baseline counterfactual. It describes the wealth path and bequest distribution in the absence of taxation. Alternatively one can think of it as the potential tax base of any policy if there were no behavioral responses of any kind. In this case, donors would make their decisions as if there was no tax in place. Conveniently, we can anchor various alternative tax designs at this zero-tax counterfactual to make them comparable. In fact, any policy that collects equivalent tax revenues when applied to the no-tax (no response) counterfactual wealth and bequest distributions, is comparable from the government's budget point of view. Table 6 shows the tax schedules of comparable counterfactual policies. The marginal tax-rates result from anchoring at the no-tax counterfactual.

Table 6: Tax Schedules of the Policy Counterfactuals

1. Inheritance Tax	
Tax bracket, SEK	Tax (lump sum + tax rate)
0 - 300.000	0 + 10%
300.000 - 600.000	30.000 + 20%
> 600.000	90.000 + 30%
Basic Exemptions: SEK 70.000 per child	
2. Estate Tax with Flexible Kinks	
0 - 300.000*# children	0 + 10%
300.000*# children - 600.000*# children	30.000*# children + 20%
> 600.000*# children	90.000*# children + 30%
Basic Exemptions: SEK 70.000 per child	
3. Estate Tax with Fixed Kinks	
0 - 600.000	0 + 9,76%
600.000 - 1.200.000	58.560 + 19,52%
> 1.200.000	175.680 + 29,28%
Basic Exemptions: SEK 140.000	

Notes: Table presents tax schedules used in counterfactual policy simulations. The first counterfactual is no tax and is not listed in the Table. The upper tax schedule presents the existing inheritance tax. The remaining two counterfactual tax schemes mimic similar tax scheme to inheritance tax with fixed and flexible kinks.

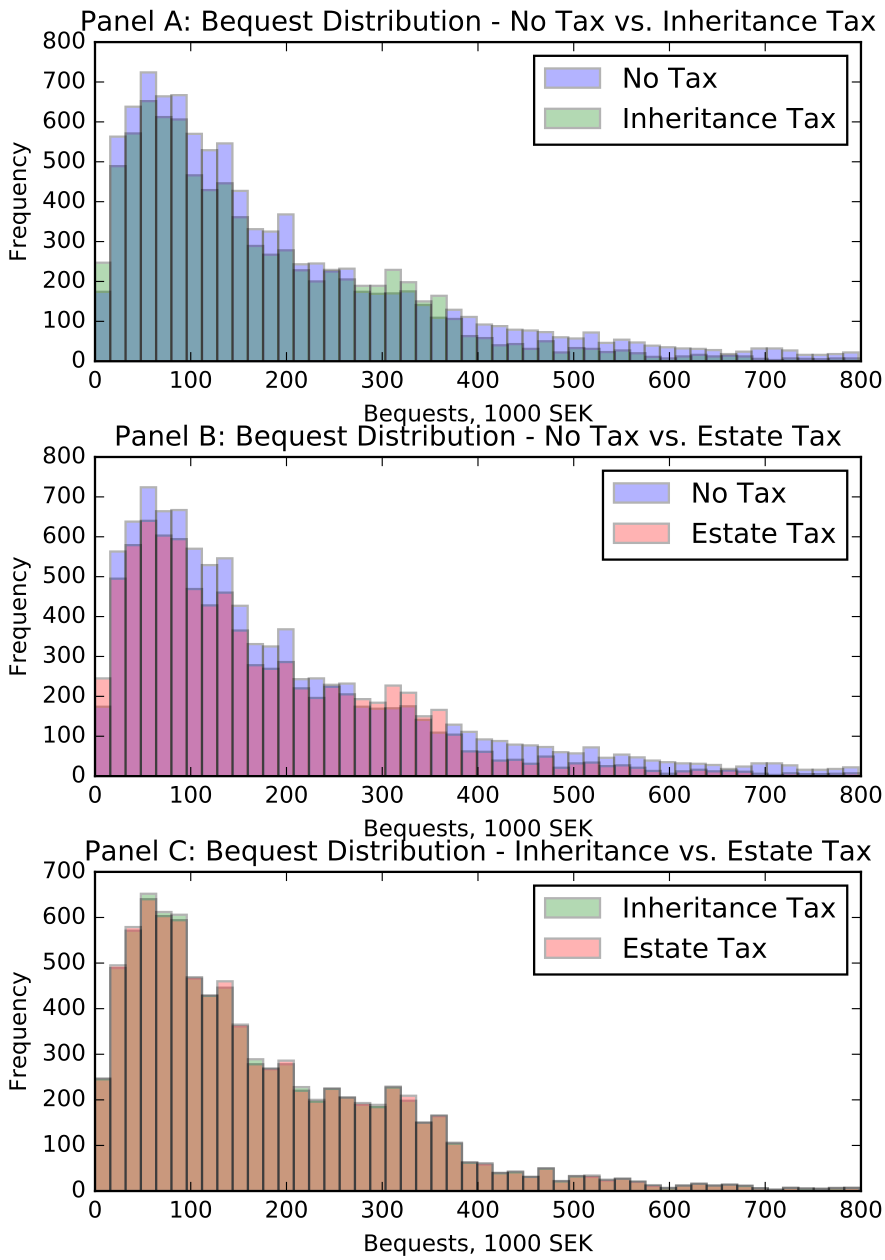
C Taxes and the bequest distribution

Responses to bequest taxation lead to overall changes in the distribution of terminal bequests. Figure 8 shows the bequest distribution of the Swedish inheritance tax (1.) and a comparable estate tax (2.) in contrast to the no-tax counterfactual (0.). Panels A and B show that any type of bequest taxation reduces the total mass in the bequest distribution. The taxes provide an incentive to bequeath less in general, and at lower marginal tax rates in particular. The latter incentive is particularly strong in the upper tail of the bequest distribution. The tax kinks in both tax schedules at SEK 70.000, 370.000 and 670.000 create visible excess masses of bequests, that are absent in the no tax case. These bequest adjustments highlight the elasticity of bequest to taxation caused by the behavioral responses of donors. Panel C shows that the two alternative bequest taxes (1. and 2.) fare similar compared to each other. Slight differences are due to the additional potential strategy of donors to avoid taxes by changing individual bequest rather than total estate.

While inheritance and estate tax with comparable schedules affect the bequest distribution very similarly, an estate tax with fixed exemption levels and tax rate kinks (3.) may generate a very different distribution. One example of this widely used tax design, which is not flexible to family structure, is shown in Figure 9. It suggests that policymakers need to be aware of the incentives that family-specific exemptions provide for donors to avoid paying taxes. In particular, such an estate tax design puts higher burden on large and wealthy families.

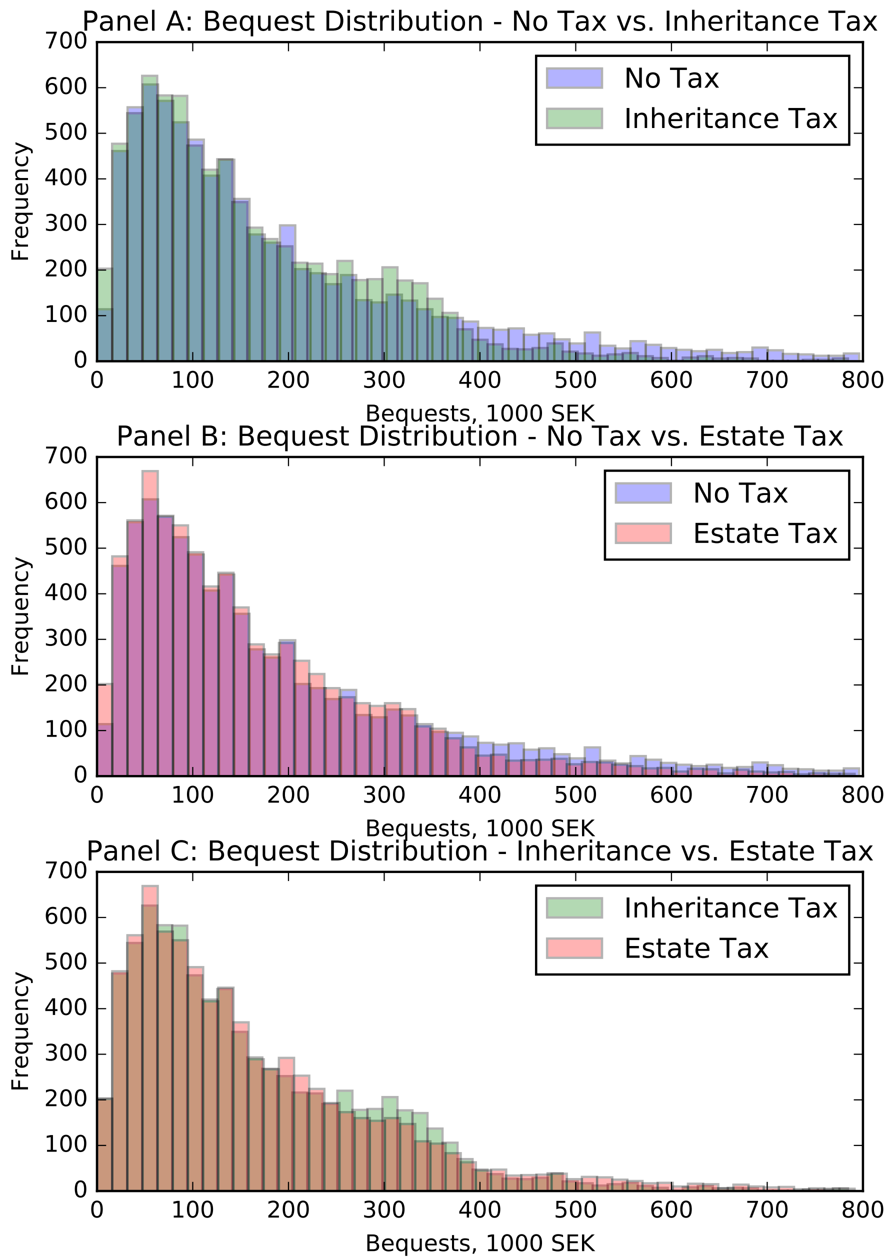
Notice that a tax on intergenerational wealth transfers can create incentives for donors to transfer more wealth via inter-vivos gifts since a SEK 10.000 annual exemption is applied to them. While the model allows for this strategic response, it is ignorant about the exact levels and recipients of such transfers. At the moment, the bequest distributions in this paper only reflect transfers out of the terminal wealth.

Figure 8: Effect of Taxation on bequest distributions: Counterfactual policies 1. and 2.



Notes: Bins of the bequest distribution for the inheritance and estate tax counterfactuals (1. and 2.) in comparison to the no-tax counterfactual.

Figure 9: Effect of taxation on bequest distributions: Counterfactual policies 1. and 3.



Notes: Bins of the bequest distribution for the inheritance and estate tax counterfactuals (1. and 3.) in comparison to the no-tax counterfactual (Panel A and B). Panel C compares the distributions under the two taxes to each other.

D Effect on the Wealth Accumulation Process in Old Age

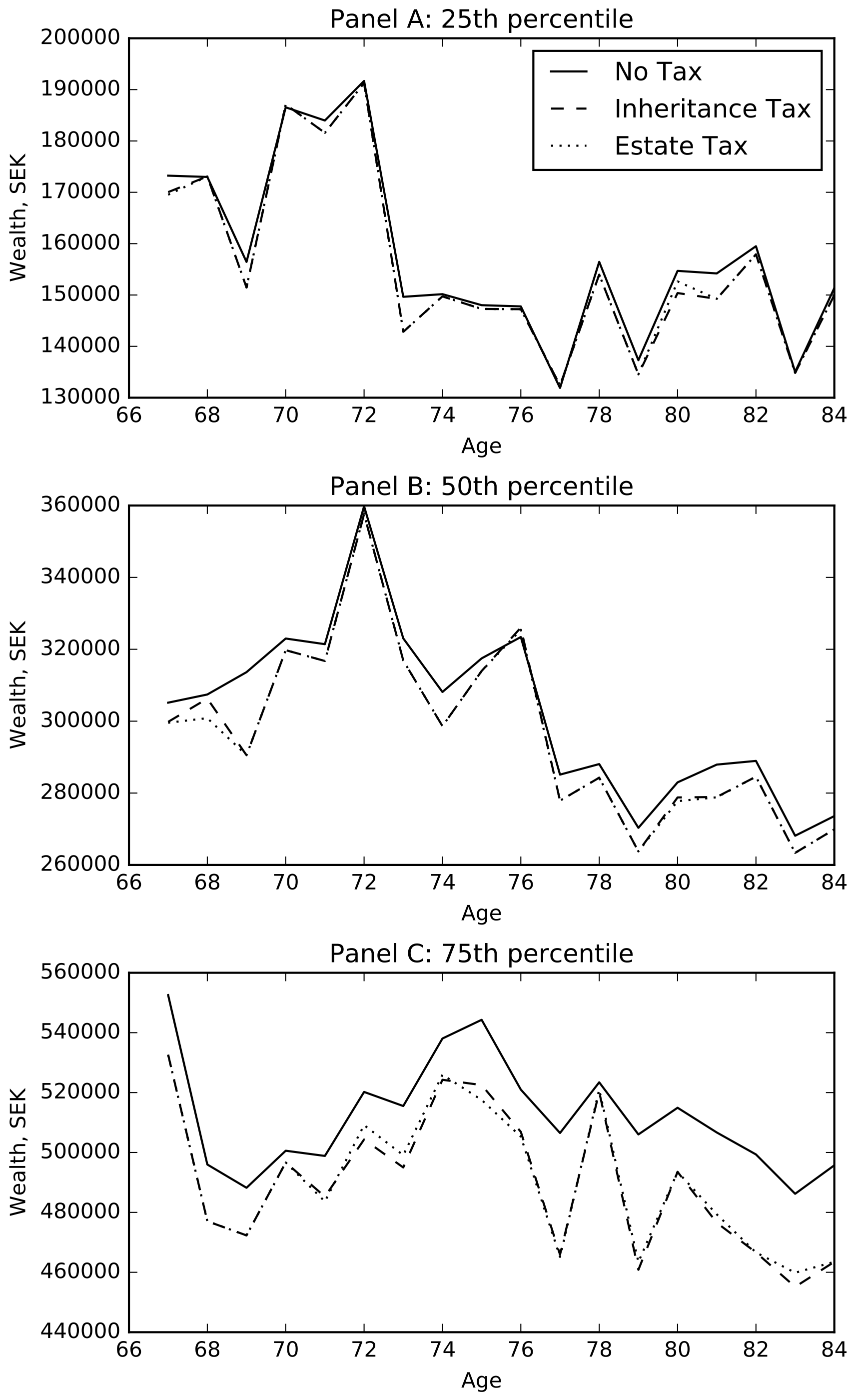
A fundamental concern with respect to bequest taxation is whether behavioral responses to the tax lead to distortions in wealth accumulation. Such behavior can lead to a wide range of second-order effects, e.g. change in the overall saving rate and inequality across families and generations. We find relatively large responses to both inheritance and estate taxation in the wealth accumulation of old-age individuals compared to the no-tax counterfactuals. Figure 10 plots the wealth levels per age for the 25th, 50th and 75th percentile of the wealth distribution for the comparable counterfactual policies 1) and 2).[13] Due to the progressiveness of both taxes, wealthier donors are adjusting their wealth path to a larger degree. Due to the similarity of policies 1) and 2) in terms of tax schedule, inheritance and estate tax affect the wealth accumulation process almost equally. These changes in the wealth path in old age translate directly into the distribution of terminal estates in Figure 12.

Again, if a more common estate tax design with fixed progressiveness is chosen, the behavioral responses may differ significantly, as shown in Figure 11. Non-individual exemption levels and tax rate kinks reduce the incentive for many families to engage in tax avoidance via wealth accumulation and therefore reduces the wealth distortion. This is due to the fact that for many families, lower tax-brackets are out of reach. The caveat of such policy is the tax incidence is shifted particularly towards large families.

The effects of taxation on terminal wealth are in line with the literature. The elasticity of terminal estate to tax of 0.22 generated by our model compares well to the 0.1-0.2 elasticities estimated by Slemrod & Kopczuk (2000), Joulfaian (2006) and Glogowsky (2016) as well as to the 0.09-0.27 elasticity of wealth to the Swedish wealth taxation documented by Seim (2017).

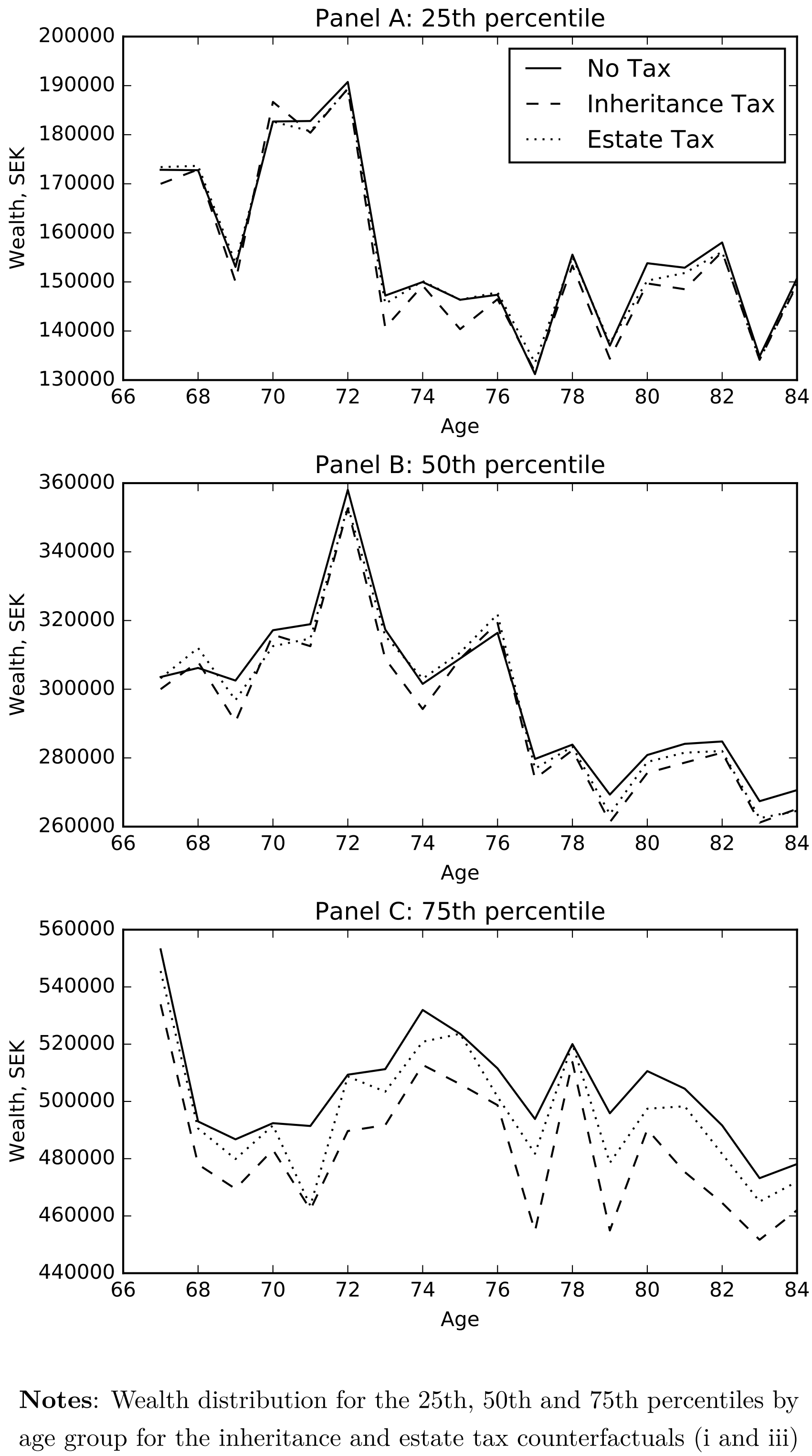
[13]Wealth levels per age group are subject to the cohort composition effect, as they are pooled over different calendar years.

Figure 10: Effect of taxation on wealth accumulation: Counterfactual policies 1. and 2.



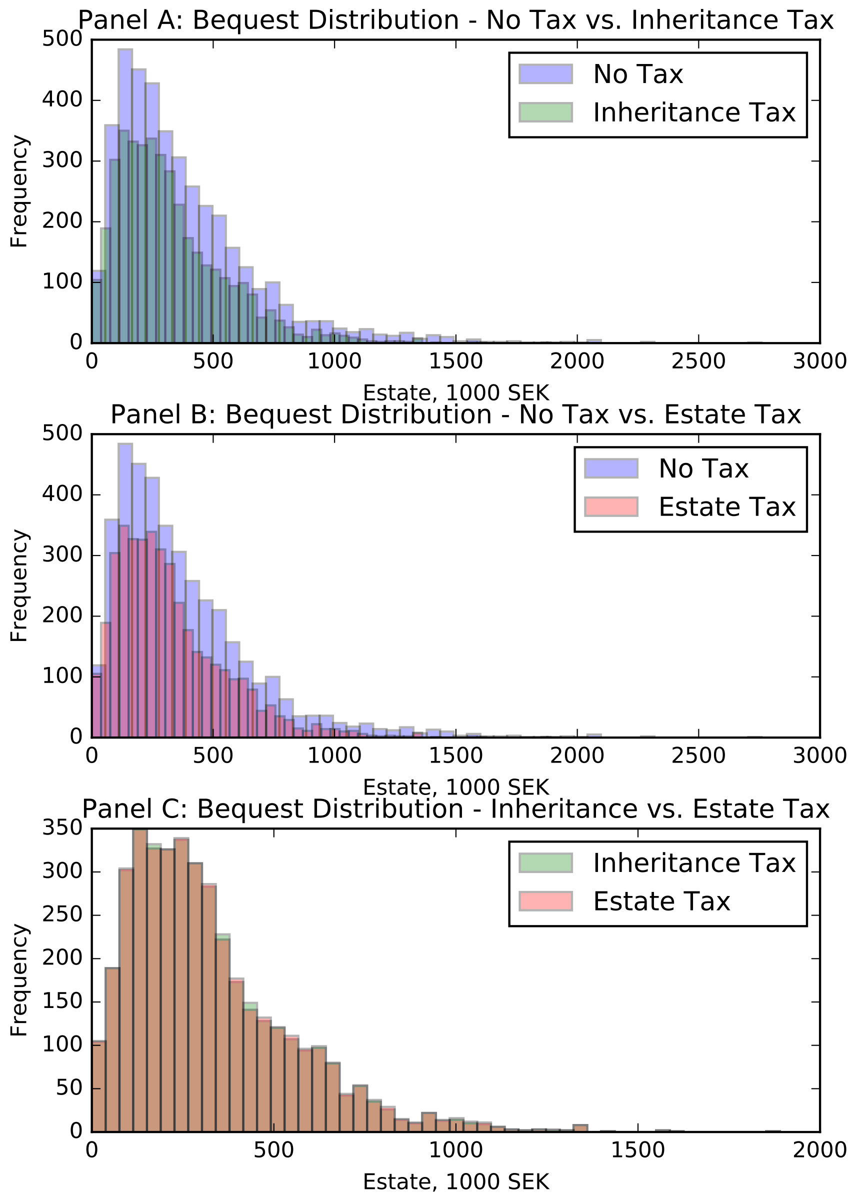
Notes: Wealth distribution for the 25th, 50th and 75th percentiles by age group for the inheritance and estate tax counterfactuals (i and ii) in comparison to the no-tax counterfactual (Panel A and B). Panel C compares the distributions under the two taxes to each other.

Figure 11: Effect of taxation on wealth accumulation: Counterfactual policies 1. and 3.



Notes: Wealth distribution for the 25th, 50th and 75th percentiles by age group for the inheritance and estate tax counterfactuals (i and iii) in comparison to the no-tax counterfactual.

Figure 12: Effect of taxation on terminal wealth: Counterfactual policies 1. and 2.



Notes: Distribution of terminal estates for the inheritance and estate tax counterfactuals (i and ii) in comparison to the no-tax counterfactual (Panel A and B). Panel C compares the distributions under the two taxes to each other.

E Response Decomposition and Tax Revenue

The counterfactuals allow decomposing the responses to bequest taxation. As documented below, wealth reactions to the tax are almost equal, if inheritance and estate tax have comparable schedules. In addition, we find that for the case of inheritance taxation, 1.4% of individuals change their bequest distribution among children to minimize the tax bill. Given the low incidence of unequal bequests in the Swedish context (2-5%), this represents a non-trivial fraction of the donors.

Finally, in light of high administrative costs of bequest taxation, policymakers care about the impact that responses to taxation have on the generated tax revenue. As the estate and inheritance tax in our counterfactuals are designed to have equal tax revenue in the absence of any behavioral responses, it is possible to measure the loss for the government revenues from taxpayers' reactions. We find a significant loss in revenue of 34.4% for inheritance (i) and 29.3 % for estate taxations. While the bulk of this loss is explained by the adjustment in wealth accumulation, in particular of wealthy donors, the difference between the tax designs has two origins. First, under inheritance taxation donors can use an additional strategy to avoid taxation, i.e. changing the bequest distribution. Second, strong preferences for unequal bequests under inheritance tax may lead some donors to not fully use the individual exemptions levels, which gives additional incentive to adjust the wealth path.

6 Conclusion

The bequest taxation is often at the center of policy debates because of being a tool to correct distributional inefficiencies propagated over generations with bequests. In addition, this tax is often viewed as a tax that does not cause undesirable distortion. For example, the Economist writes: *"In fact, people who are against tax in general ought to be less hostile to inheritance taxes than other sorts. However disliked they are, they are some of the least distorting" Economist (2017).*

Although this statement is theoretically appealing, this paper shows that bequest taxation implies a range of behavioral responses that should be taken into account by policymakers who are aimed at minimizing tax distortions, while simultaneously collecting tax revenues. More precisely, progressiveness and exemption levels are the predominant tools to control the incidence of the tax on particular groups in the population. Using a comprehensive structural model that captures main behavioral responses of old-age individuals, we can compare the impact of various tax designs on the wealth accumulation, bequest distributions and tax revenue, which are main policy outcomes of interest. Our results show that comparable inheritance and estate tax schedules in terms of progressiveness have similar but important effects of individual behavior. At the same time, due to the additional margins of strategic behavior under inheritance taxation, estate taxes lead to higher overall tax revenues. Therefore, this paper emphasizes that at the cost of relaxing the control over tax incidence, estate taxation can be designed to further minimize distortions.

Our model is comprehensive and flexible enough to allow for the simulation of counterfactuals for the universe of alternative bequest tax designs. Various tax structures enable policymakers to balance distortions, progressiveness, tax revenues, and tax incidence according to their social welfare functions. The context in which we apply the structural model is ideal, as it abstracts to a large degree from precautionary savings, due to the generous social welfare system in Sweden. The results are therefore likely to be applicable to other contexts, which feature additional incentives to accumulate wealth at old age. Due to the strong preferences for equal bequests within the family in our sample, differences between estate and inheritance taxation are relatively modest. They may be amplified in contexts like the US, where up to 20% (Light & McGarry, 2004), ten times the percentage of Sweden, leave unequal terminal bequests to their children and have therefore potentially higher responses through changes in the bequest distribution.

The findings in this paper also open up several paths for future research. One important question is considering inter-generational wealth taxation in a general equilibrium framework that takes into account wealth taxation and returns to capital. The large wealth effects in our model emphasize that the design of the inheritance tax might have important spillovers on the wealth taxation and capital markets more broadly. Another important area is spousal tax-planning in old ages since distortions from the tax and responses to it are often realized at the household rather than individual level. In addition, the relevance of behavioral responses to bequest taxation for inequality should be investigated. Taxes on inter-generational transfers are almost always designed to redistribute wealth from the upper to the lower tail of the distribution. If, however, strategic responses benefit specific subgroups, the effectiveness of the tax might be harmed. Finally, although we touched upon the question of within-family inequality, further works might be needed to study the aggregate implications of changes in within-family wealth allocation.

References

Adermon, Adrian, Lindahl, Mikael, & Waldenström, Daniel. 2018. Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations. *The Economic Journal*, **128**(612), 482–513.

Arrondel, Luc, & Masson, André. 2006. Chapter 14 Altruism, Exchange or Indirect Reciprocity: What Do the Data on Family Transfers Show? *Pages 971 – 1053 of:* Kolm, Serge-Christophe, & Ythier, Jean Mercier (eds), *Applications*. Handbook of the Economics of Giving, Altruism and Reciprocity, vol. 2. Elsevier.

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

Becker, Gary S., & Tomes, Nigel. 1979. An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*, **87**(6), 1153–1189.

Behrman, Jere R, Pollak, Robert A, & Taubman, Paul. 1982. Parental Preferences and Provision for Progeny. *Journal of Political Economy*, **90**(1), 52–73.

Blundell, Richard, Costa Dias, Monica, Meghir, Costas, & Shaw, Jonathan. 2016. Female Labor Supply, Human Capital, and Welfare Reform. *Econometrica*, **84**(5), 1705–1753.

Boserup, Simon H., Kopczuk, Wojciech, & Kreiner, Claus T. 2016. The Role of Bequests in Shaping Wealth Inequality: Evidence from Danish Wealth Records. *American Economic Review*, **106**(5), 656–661.

Cox, Donald. 2003. *Private Transfers within the Family: Mothers, Fathers, Sons and Daughters.* Boston College Working Papers in Economics 605. Boston College Department of Economics.

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

De Nardi, Mariacristina, & Yang, Fang. 2014. Bequests and Heterogeneity in Retirement Wealth. *European Economic Review*, **72**(C), 182–196.

De Nardi, Mariacristina, French, Eric, & Jones, John B. 2010. Why Do the Elderly Save? The Role of Medical Expenses. *Journal of Political Economy*, **118**(1), 39–75.

Duffie, Darrell, & Singleton, Kenneth J. 1997. An Econometric Model of the Term Structure of Interest-Rate Swap Yields. *The Journal of Finance*, **52**(4), 1287–1321.

Economist, The. 2017. A hated tax but a fair one. *www.economist.com*, Nov.

Eliason, M., & Ohlsson, H. 2013. Timing of Death and the Repeal of the Swedish Inheritance Tax. *Journal of Behavioral and Experimental Economics*, **45**(C), 113–123.

Elinder, Mikael, Erixson, Oscar, Escobar, Sebastian, & Ohlsson, Henry. 2014 (Nov.). *Estates, Bequests, and Inheritances in Sweden - A Look into the Belinda Databases*. Working Paper Series, Center for Fiscal Studies 2014:14. Uppsala University, Department of Economics.

Elinder, Mikael, Erixson, Oscar, & Waldenström, Daniel. 2016 (Mar.). *Inheritance and Wealth Inequality: Evidence from Population Registers*. IZA Discussion Papers 9839. Institute for the Study of Labor (IZA).

Erixson, Oscar, & Ohlsson, Henry. 2014 (Feb.). *Estate Division: Equal Sharing as Choice, Social Norm, and Legal Requirement*. Working Paper Series, Center for Fiscal Studies 2014:2. Uppsala University, Department of Economics.

Escobar, Sebastian, Ohlsson, Henry, & Selin, Håkan. 2019. *Taxes, frictions and asset shifting: when Swedes disinherited themselves*. Tech. rept.

Fella, Giulio. 2014. A Generalized Endogenous Grid Method for Non-Smooth and Non-Concave Problems. *Review of Economic Dynamics*, **17**(2), 329–344.

French, Eric. 2005. The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour. *Review of Economic Studies*, **72**(2), 395–427.

Glogowsky, Ulrich. 2016. Behavioral Responses to Wealth Transfer Taxation: Bunching Evidence from Germany. *Working Paper*.

Hansen, Nikolaus. 2006. The CMA Evolution Strategy: A Comparing Review. *Pages 75–102 of: Towards a new evolutionary computation*. Springer.

Henrekson, Magnus, Du Rietz, Gunnar, & Waldenström, Daniel. 2014. *Swedish Inheritance and Gift Taxation (1885–2004)*. Working Paper Series 936. Research Institute of Industrial Economics.

Holmlund, Helena, Lindahl, Mikael, & Plug, Erik. 2011. The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. *Journal of Economic Literature*, **49**(3), 615–51.

Iskhakov, Fedor, Jørgensen, Thomas H, Rust, John, & Schjerning, Bertel. 2017. The Endogenous Grid Method for Discrete-Continuous Dynamic Choice Models with (or without) Taste Shocks. *Quantitative Economics*, **8**(2), 317–365.

Joulfaian, David. 2005. Choosing Between Gifts and Bequests: How Taxes Affect the Timing of Wealth Transfers. *Journal of Public Economics*, **89**(11-12), 2069–2091.

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

Kendrick, M. Slade. 1939. The Ability-to-Pay Theory of Taxation. *The American Economic Review*, **29**(1), 92–101.

Kopczuk, Wojciech, & Lupton, Joseph P. 2007. To Leave or Not to Leave: The Distribution of Bequest Motives. *The Review of Economic Studies*, **74**(1), 207–235.

Kopczuk, Wojciech, & Slemrod, Joel. 2003. Dying to Save Taxes: Evidence from Estate-Tax Returns on the Death Elasticity. *The Review of Economics and Statistics*, **85**(2), 256–265.

Laitner, John, Silverman, Dan, & Stolyarov, Dmitriy. 2018. The Role of Annuitized Wealth in Post-Retirement Behavior. *American Economic Journal: Macroeconomics*, **10**(3), 71–117.

Light, Audrey, & McGarry, Kathleen. 2004. Why Parents Play Favorites: Explanations for Unequal Bequests. *American Economic Review*, **94**(5), 1669–1681.

Lockwood, Lee. 2012. Bequest Motives and the Annuity Puzzle. *Review of Economic Dynamics*, **15**(2), 226–243.

Lockwood, Lee M. 2018. Incidental Bequests and the Choice to Self-Insure Late-Life Risks. *American Economic Review*.

Ohlsson, Henry. 2007. *Tax Avoidance, a Natural Experiment*. Working Paper Series 2007:13. Uppsala University, Department of Economics.

Ohlsson, Henry. 2011. The Legacy of the Swedish Gift and Inheritance Tax, 1884–2004. *European Review of Economic History*, **15**(03), 539–569.

Pakes, Ariel, & Pollard, David. 1989. Simulation and the Asymptotics of Optimization Estimators. *Econometrica: Journal of the Econometric Society*, 1027–1057.

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

Saez, Emmanuel. 2010. Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, **2**(3), 180–212.

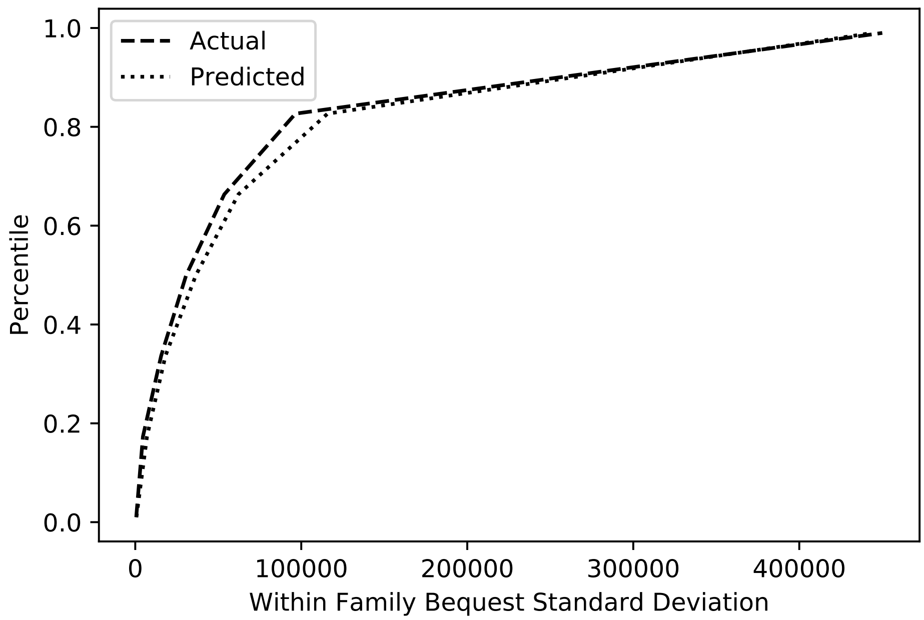
Seim, David. 2017. Behavioral Responses to Wealth Taxes: Evidence from Sweden. *American Economic Journal: Economic Policy*, **9**(4), 395–421.

Slemrod, Joel, & Kopczuk, Wojciech. 2000. *The Impact of the Estate Tax on the Wealth Accumulation and Avoidance Behavior of Donors*. NBER Working Papers 7960. National Bureau of Economic Research, Inc.

Appendices

A Figures

Figure 13: Model Fit



Notes: Figure illustrates a fit of the first stage model. In particular, we plot a distribution of standard deviation moments of within family bequests.

B Dynamic Model - Euler Equations

The Euler equations for the two choice variables in the dynamic problem show the fundamental trade-off that the donor faces when deciding on c and s. First, a change in the expected marginal utility of consumption and gift transfers must correspond to a change in the expected marginal utility from bequeathing at death.

$$u_{ct} - (1 - D_{t+1})\delta E_t [u_{ct+1}] = \delta E_t D_{t+1} [B_{Wt+1}] \tag{13}$$

Secondly, the Euler equation for each element j of the bequest share vector s requires optimality for the current wealth level of each period.

$$B_{s_{jt}} = 0 \tag{14}$$

Incorporating the functional form of the bequest function yields the following Euler equation for c_t and first order condition for the optimal bequest share s to child l:

$$\begin{aligned} &\overbrace{c_t^{-\eta} - (1 - D_{t+1})\delta E_t \left[c_{t+1}^{-\eta}\right]}^{\Delta\text{marginal utility of consumption}} \\ = \delta D_{t+1} E_t &\overbrace{\left[\left(\frac{\lambda_1}{1-\lambda_1}\right)^{\eta}\left(\frac{\lambda_1}{1-\lambda_1}\cdot c_b + \sum_{j=1}^{J}(1-\tau_j)s_j W\right)^{-\eta}\sum_{j=1}^{J} s_j\left(1-\tau_J - W\tau_W\right)\right]}^{\text{marginal utility from bequeathing}} \end{aligned} \tag{15}$$

$$\begin{aligned} &\overbrace{\left(\frac{\lambda_1}{1-\lambda_1}\right)^{\eta}\left[\frac{\lambda_1}{1-\lambda_1}\cdot c_b + \sum_{j=1}^{J}(1-\tau_j)s_j E_t(W_{t+1})\right]^{-\eta}}^{\text{marginal utility from sum of after-tax bequests}} E_t(W_{t+1})\left[1-\tau_l - s_l\tau_{s_l}\right] \\ &= \overbrace{2\lambda_2\left[U^s(s_1 ... s_J) - U^s({s_1}^* ... s_J^*)\right]\frac{\exp(\phi_l)}{s_l}}^{\text{marginal disutility through share deviation}} \end{aligned} \tag{16}$$

with $\sum_{j=1}^{J} s_j = 1$. These two equations guide the optimal behavior of donors and illustrate the identification of λ_1 and λ_2.

C Estimation Details

As described in the main text, the estimation of the model consists of two steps. Firstly, we estimate true allocation preferences. These parameters obtained on the first stage are used to estimate the main parameters from the dynamic model. Overall, the model contains 7 second-stage parameters and 8 first stage parameters. More parameters required to be estimated leads to more iteration of the optimization algorithm. Therefore, reducing the number of parameters estimated on the second stage while solving computationally intensive dynamic model provides large computational gains.

Estimation of the first stage parameters is fairly computationally light and is conducted on the full sample of eligible individuals from bequest dataset. To estimate the second stage model, we draw a 20% random sample. We use 10 cores to solve the model at each iteration of the optimization algorithm. Both estimation procedures match chosen moments that describe features of individuals behavior of interest. Both models have more moments than parameters. We use a two-stage optimal GMM matrix for both models.

Theoretically, our second stage estimator is consistent and asymptotically normally distributed (Pakes & Pollard, 1989; Duffie & Singleton, 1997).

$$\xi \sim N\left(0, (G'_\xi W G_\xi)^{-1} G'_\xi W \left[(1 + \frac{N_d}{N_s})\Omega_\xi + G_\varphi \Omega_\varphi G'_\varphi\right] W G_\xi (G'_\xi W G_\xi)^{-1}\right) \tag{17}$$

where G_φ, G_ξ are the gradient matrices of moments with respect to first stage and second stage parameters, correspondingly. $\Omega_\varphi, \Omega_\xi$ denote moment variance-covariance matrices of the first and second stage, correspondingly, and N_d, N_s are sample and simulation sample size.

To estimate the parameters of the model we use a stochastic global optimizer to explore the parameter space. More precisely, we use the CMA-ES algorithm from different starting values. Then we use the simplex algorithm starting from the best parameters obtained from the global optimizer to refine the solution.