NBER WORKING PAPER SERIES

WEALTH TAXATION AND WEALTH ACCUMULATION:
THEORY AND EVIDENCE FROM DENMARK

Katrine Jakobsen
Kristian Jakobsen
Henrik Kleven
Gabriel Zucman

Wealth Taxation and Wealth Accumulation: Theory and Evidence from Denmark
Katrine Jakobsen, Kristian Jakobsen, Henrik Kleven, and Gabriel Zucman
NBER Working Paper No. 24371
March 2018
JEL No. E2,E6,H2,H3

ABSTRACT

Using administrative wealth records from Denmark, we study the effects of wealth taxes on wealth accumulation. Denmark used to impose one of the world's highest marginal tax rates on wealth, but this tax was drastically reduced and ultimately abolished between 1989 and 1997. Due to the specific design of the wealth tax, these changes provide a compelling quasi-experiment for understanding behavioral responses among the wealthiest segments of the population. We find clear reduced-form effects of wealth taxes in the short and medium run, with larger effects on the very wealthy than on the moderately wealthy. We develop a simple lifecycle model with utility of residual wealth (bequests) allowing us to interpret the evidence in terms of structural primitives. We calibrate the model to the quasi-experimental moments and simulate the model forward to estimate the long-run effect of wealth taxes on wealth accumulation. Our simulations show that the long-run elasticity of wealth with respect to the net-of-tax return is sizeable at the top of distribution. Our paper provides the type of evidence needed to assess optimal capital taxation.

Katrine Jakobsen
Department of Economics
University of Copenhagen
DK-1353 Copenhagen
k.m.jakobsen@econ.ku.dk

Kristian Jakobsen
Den Sociale Kapitalfond
Vester Voldgade 108
DK-1552 Copenhagen
ktj@socialkapitalfond.dk

Henrik Kleven
Department of Economics
Princeton University
238 Julis Romo Rabinowitz Building
Princeton, NJ 08544
and NBER
kleven@princeton.edu

Gabriel Zucman
Department of Economics
University of California, Berkeley
530 Evans Hall, #3880
Berkeley, CA 94720
and NBER
zucman@berkeley.edu

1 Introduction

What are the economic effects of taxing household wealth? While an enormous literature estimates the elasticity of labor supply and taxable income, much less is known about how taxes affect the supply of capital. The lack of evidence makes it hard to assess the desirability of taxing household wealth, a proposal that has gained interest following Thomas Piketty's call for a global wealth tax (Piketty 2014) and new evidence of rising wealth inequality in the United States (Saez & Zucman 2016). How would wealth taxes affect the saving and consumption decisions of the rich? How would wealth taxes affect avoidance and evasion decisions? Would they reduce wealth inequality, and by how much?

Answering these questions is difficult due to several empirical challenges. First, while many countries collect data on labor supply and taxable income, very few countries collect individual data on wealth. Second, it has been difficult to find compelling variation in wealth taxation that allows for the estimation of causal effects. What is more, because wealth is always very concentrated — much more than labor income — it is crucial to estimate behavioral responses for the very wealthiest individuals.[1] Sources of exogenous variation at the top of the wealth distribution has so far been elusive. Third, in order to assess the desirability of wealth taxes, and of capital taxes more broadly, it is important to obtain estimates of long-run effects. While tax design always depends on the long run, this is a bigger challenge for capital taxes than for labor taxes due to the dynamic and slow-moving nature of wealth accumulation.

In this paper, we break new ground on these questions. Our laboratory is Denmark, which offers data and quasi-experimental variation that allow us to overcome the challenges described above. Until 1997, Denmark taxed household wealth above an exemption threshold located around the 98th percentile of the household wealth distribution. Through to the 1990s, a dozen of OECD countries levied similar taxes (OECD 1988), but the Danish wealth tax was the largest of its kind. The marginal tax rate on wealth equalled 2.2% up until the late 1980s, which corresponds to a very high rate on the *return* to wealth.[2] The Danish government implemented large changes to the wealth tax starting in 1989 — cutting the marginal rate to 1% and doubling the exemption threshold

---

[1] As we show in the paper, the top 10% owns about half of all wealth in Denmark, while the top 1% owns 20% of all weath. Wealth concentration in the US is even higher (see Saez & Zucman 2016).

[2] For example, assuming a rate of return on wealth equal to 6.6% (which is in the ballpark of empirical estimates at the top), a marginal wealth tax of 2.2% corresponds to a 33.3% tax on the flow of capital income.

for married couples — before eventually abolishing the tax in 1997. These policy changes represent one of the largest natural experiment with wealth taxation ever conducted. Besides this experiment, a key advantage of the Danish setting is that the authorities have been collecting micro-level data on wealth for the entire population since 1980.

Our paper makes three main contributions. The first is to provide quasi-experimental evidence on the effects of the 1989 tax cuts on wealth accumulation. We consider two different empirical strategies and samples. One strategy exploits that, among the very wealthiest households, some face a zero marginal tax rate on wealth due to a *tax ceiling* that limits the total average tax rate from personal taxes (income, social security, and wealth taxes). Therefore, the tax cuts had different impacts on those bound and unbound by the ceiling. This allows us to estimate impacts of wealth taxes on the super wealthy using a difference-in-differences design comparing bound and unbound taxpayers within the top 1%.[3] The other strategy exploits the doubling of the exemption threshold for couples (but not singles), which eliminated wealth taxes among couples located roughly between the 98th and 99th percentiles of the wealth distribution. This allows us to estimate impacts of wealth taxes on the moderately wealthy using a difference-in-differences design comparing couples in the exempted range to couples in other ranges or to singles in the same range.

The quasi-experimental analysis shows that wealth taxes have sizeable effects on taxable wealth, with the effects being much larger at the extreme top of the distribution than further down. We view our evidence as compelling in the sense that, in both of our approaches, the trends in taxable wealth for the treatment and control groups are parallel prior to the reform and then begin to diverge immediately after the reform.[4] The effect on wealth builds up over time and is equal to about 30% after 8 years for the very wealthy (ceiling DD) and about 10% after 8 years for the moderately wealthy (couples DD). These effects include both behavioral and mechanical effects: even if households did not change their behavior in response to wealth taxes, the increase in the after-tax rate of return would mechanically increase wealth over time. We show that the mechanical effects account for about one-quarter of the effect for the ceiling DD and about one-tenth of the

---

[3]The ceiling strategy represents a novel empirical approach in the large literature on behavioral responses to taxes. This approach offers a promising way to identify behavioral responses among the very wealthy that could be implemented in a number of countries with wealth taxes. This is because most countries with wealth taxes (including Norway, Sweden, France, Spain, and Germany) have such ceiling rules.

[4]As we describe in the paper, the pre-trends are parallel after making a non-parametric adjustment for differences in pre-reform portfolio compositions between treatment and control groups. Without adjusting for pre-reform portfolio composition — specifically differences in housing shares and equity shares across treatments and controls — there would be bias from confounding asset price movements.

2

effect for the couples DD.[5] As a result, the behavioral responses to wealth taxes are larger among the very wealthy than among the moderately wealthy, but by less than the raw estimates suggest.

Our second contribution is to develop a theoretical model allowing us to interpret the reduced-form impacts in terms of structural primitives. To keep the model relatively simple, we leave out aspects that are not central to our setting and sample. In particular, because wealthy people tend to be older people — most of those in the top 1% of the wealth distribution are between 50-90 years old — we focus on the savings motives that are central to older, wealthy people. We argue that the lifecycle motive and the bequest motive (or more broadly utility of residual wealth) are important, while the precautionary motive is second order. Within such a model, we demonstrate how the reduced-form impact on wealth is driven by four conceptual effects: a substitution effect on consumption proportional to the Elasticity of Intertemporal Consumption (EIS), a substitution effect on bequests proportional to a bequest elasticity, a wealth effect on the demand for consumption and bequests, and finally the mechanical effect discussed earlier. The importance of the bequest elasticity in determining the reduced-form impacts depends on the weight of the bequest motive in household preferences, and we show that this weight has to be large in order to rationalize the lifecycle profile of wealth among very wealthy people. Therefore, the bequest elasticity is very important for understanding wealth responses at the top.

Our third contribution is to connect the theory and evidence in order to explore the long-run effects of wealth taxes on wealth accumulation. We calibrate the parameters of the model to match the empirical lifecycle profile of wealth at the top of the distribution as well as the quasi-experimental estimates of the short-medium term impacts of wealth taxes. When matching the model to the moderately wealthy in the couples DD — an empirical effect on taxable wealth of about 10% after 8 years — and simulating the model forward, we obtain a 30-year effect of about 20%. When matching the model to the super wealthy in the ceiling DD — an empirical effect on taxable wealth of about 30% after 8 years — the long-run effect is considerably larger, about 70% after 30 years. While this effect may seem large, note that the wealthiest taxpayers have access to particularly effective avoidance and evasion vehicles, making our model of real wealth accumulation less suited for this population (see Alstadsæter *et al.* 2017a).[6] Our estimates for the super

---

[5]The main reason why the mechanical effects are larger in the ceiling DD than in the couples DD is that the former considers households farther above the exemption threshold, so that their *average* after-tax return (which governs the mechanical effect) increases by more.

[6]Using leaked data from HSBC Switzerland and Mossack Fonseca ("Panama Papers"), Alstadsæter *et al.* (2017a) show that essentially all of the wealth in offshore accounts belongs to the top 1% and that most of it belongs to the top 0.1%. If such hidden wealth responds to wealth taxes, this would be picked up by the ceiling DD estimates for the top 1%. It would not be picked up by the couples DD, however, because they are not wealthy enough. Furthermore, for the

wealthy therefore represent upper bounds on real wealth accumulation responses, whereas the estimates for the moderately wealthy arguably capture real responses.

Our paper can be viewed in two ways. One view is that it contributes to a nascent literature studying the effects of wealth taxes (Zoutman 2015; Brülhart *et al.* 2016; Seim 2017). Compared to this literature, we consider a larger natural experiment and we estimate behavioral responses at the very top of the wealth distribution. We provide clear graphical evidence on the short-medium term responses to wealth taxes.[7] Unlike earlier work, our paper provides a tractable dynamic framework to shed light on the theoretical mechanisms driving the estimated reduced-form impacts, and it structurally estimates the model in order to explore the long-term consequences of wealth taxation.

Another view is that our paper provides a first attempt to causally estimate the long-run elasticity of capital supply with respect to capital taxes. From this perspective, it is not crucial that we study wealth taxes per se, but rather that the Danish wealth tax allows us to estimate a key parameter for assessing the efficiency implications of capital taxes more broadly. Saez & Stantcheva (2017) show that the long-run elasticity of capital supply with respect to the after-tax rate of return is a sufficient statistic for optimal capital taxation, but there is virtually no evidence on what a reasonable value of this elasticity might be. Besides the empirical challenges discussed above, a reason for the lack of evidence may be that the seminal theoretical contributions guiding the debate focused on "corner solutions" that did not bring out the key role of the capital supply elasticity. In the Chamley-Judd framework (Chamley 1986; Judd 1985), the optimal capital tax is zero in steady state because long-run capital supply is infinitely elastic. In the Atkinson-Stiglitz framework (Atkinson & Stiglitz 1976), the optimal capital tax is zero because there is no heterogeneity in wealth, conditional on labor income. In other words, in one framework capital taxes are undesirable because they are too costly for efficiency, while in the other framework capital taxes are undesirable because they do not improve equity. But in general capital taxes do pose a trade-off between efficiency and equity, and it is governed by the long-run parameters we estimate here.

In the process of producing the findings described above, we provide a number of bonus contributions. It is worth highlight some of those here. First, our structural approach yields an esti-

moderately wealthy, we present bunching evidence suggesting that avoidance/evasion responses to wealth taxes are very small.

[7]Using a kink point created by the exemption threshold in the Swedish wealth tax, Seim (2017) presents compelling bunching evidence on behavioral responses to wealth taxation. We will also present bunching evidence using the Danish exemption threshold. However, while bunching is useful for estimating avoidance/evasion responses to wealth taxes, it cannot be used to capture *real* responses to wealth taxes (see also Kleven 2016). Because we are primarily interested in real wealth accumulation in this paper, we do not focus on bunching strategies.

mate of the bequest elasticity with respect to the net-of-tax rate on capital. While a large literature discusses the incentive effects of taxes on the size of bequests — typically focusing on estate and inheritance taxes — there is very little empirical evidence on the question. Reviews by Kopczuk (2009, 2013a,b) summarize the few existing estimates and discuss the challenges associated with interpreting them. We estimate the bequest elasticity based on a fundamentally different approach using variation in wealth taxes (rather than wealth *transfer* taxes) on wealthy, older people.[8] Focusing on the sample of moderately wealthy people (to avoid evasion effects as discussed above), we find a bequest elasticity of around 0.5. This elasticity is a key parameter for optimal inheritance tax rates as shown by Piketty & Saez (2013).

Second, to calibrate our model we carefully document the empirical lifecycle profiles of wealth at the top of the distribution. Because we have access to full-population administrative wealth data over a long time horizon, we are able to provide very clean and striking evidence. We show that wealthy people tend to accumulate wealth through most of their lives; only after they reach 80 years of age do their wealth profiles flatten or fall slightly. As a result, people at the top of the wealth distribution tend to die close to their wealth peak. For example, among those who make it to the top 1% of the wealth distribution during their lifetime, the average person is still in the top 2% at age 90 and have almost 20 times the amount of per capita wealth. These findings show how inaccurate the pure lifecycle model is for wealthy individuals, and they would be difficult to rationalize without some form of bequest motive or utility of residual wealth. This part contributes to a literature trying to explain wealth concentration and the lifecycle saving behavior of the rich (e.g. Carroll 2002; Benhabib & Bisin 2017), and especially papers emphasizing bequests (e.g. De Nardi 2004).

The paper is organized as follows. Section 2 describes the data and documents the evolution of wealth concentration in Denmark, section 3 presents quasi-experimental evidence on the effects of wealth taxation, section 4 develops the theoretical model, section 5 combines the model and quasi-experimental evidence in order to structurally estimate long-run effects of wealth taxes, and section 6 concludes.

---

[8]As highlighted by Kopczuk (2009), a key conceptual difficulty when analyzing bequests is precisely that they relate to the stock of wealth, which accumulates over many years and depends on many tax regimes. As a result, it would be very difficult, if not impossible, to come up with a quasi-experimental design that can deliver the full causal effect of taxes on bequests without making any parametric assumptions. This motivates the approach developed here, combining quasi-experimental evidence on short-run effects with a parametric model that can convert those effects into the long run.

2 Danish Household Wealth: Data and Distribution

2.1 Wealth Data

We base our analysis of household wealth on the administrative wealth registry maintained by the Danish Statistical Agency. This registry includes annual wealth data for the entire Danish population since 1980. The Danish authorities initially collected these data to administer the wealth tax, but they continued to do so after the abolition of the wealth tax in 1997. The data is not censored or top-coded, which is a key advantage given our focus on the top of the wealth distribution. We combine the wealth registry with other administrative registries containing data on income and socio-economic characteristics such as occupation and family composition.

The wealth registry includes detailed information on end-of-year financial assets, non-financial assets, and debts. As a rule, these assets are recorded in the registry at their prevailing market prices. Most of these assets and liabilities are reported by third-parties to the Danish government, which makes the data very reliable (see Boserup *et al.* 2014 and Leth-Petersen 2010). For instance, the value of bank deposits is reported by banks; the value of listed stocks and bonds is reported by the financial institutions (banks, mutual funds, and insurance companies) who hold these securities on behalf of their clients; and the value of mortgages is reported by mortgage lenders (banks or specialized mortgage institutions). Non-financial assets are recorded using land and real estate registries. Moreover, before the wealth tax was abolished in 1997, all assets other than those reported by third parties had to be self-reported by households. This included cash, large durables (such as cars, boats, and private planes), non-corporate business assets, unlisted securities (i.e., bearer bonds, unlisted equities, and shares of housing cooperatives), assets held abroad (foreign real estate and foreign bank accounts), and inter-personal debts.

The Danish wealth data are considered of a very high quality, and they have been fruitfully used to study retirement savings (Chetty *et al.* 2014), intergenerational wealth mobility (Boserup *et al.* 2014), and the accuracy of survey responses (Kreiner *et al.* 2015). The data does have two limitations, however. First, they exclude funded pension wealth before 2012, because such assets were not subject to wealth taxation. This is not a major issue for our purposes, because we are primarily interested in the effects of wealth taxation on *taxable* wealth. More broadly, because there are strict limits on the absolute amount that can be invested in tax-preferred pension accounts, pension wealth is always a small fraction of wealth at the top of the distribution, the focus of our analysis. Second, there is a break in the data in 1997 when the wealth tax was abolished. After 1997,

while the Danish administration continued to collect wealth data from third parties, it stopped asking households to self-report assets not reported by third parties.[9] Because of this break in the data, our quasi-experimental analysis of behavioral responses to wealth taxation focuses on the large 1989-reform for which we have consistently measured taxable wealth both before and after the reform.

2.2 Computing Wealth Inequality

To provide context, we start by documenting the evolution of wealth inequality in Denmark over the 1980-2012 period. We compute homogeneous series of wealth shares in which we match 100% of the macroeconomic amount of household wealth at market value recorded in Denmark's household balance sheet. This implies that the wealth levels and wealth shares for Denmark are comparable over time and to existing series for other countries, including those estimated for the United States by Saez & Zucman (2016).[10] In keeping with standard national account concepts, our definition of wealth includes all the non-financial and financial assets that belong to Danish residents, minus debts. In particular, it includes all funded pension wealth, but excludes the present value of future government transfers as well as consumer durables and valuables.[11] Average household wealth per adult was $242,000 in 2012 (using the market exchange rate to convert Danish kroner to US dollars), a level comparable to that of the United States where it is $234,000.

The quality of the Danish data allows us to compute particularly reliable estimates of the wealth distribution. In most countries one has to rely solely on indirect methods to estimate wealth inequality such as the capitalization method or the estate multiplier method (see Zucman 2018 for a survey). In Denmark, by contrast, we directly observe the market value of most wealth components for the entire population in the administrative wealth registry. In order to capture 100% of the macroeconomic amount of household wealth, we supplement the wealth registry as follows. First, we impute funded pension wealth throughout the 1980-2012 period, using the fact

---

[9]This includes non-corporate business assets, but here the data break is only partial. After 1997, non-corporate operating equipment and inventories are no longer recorded in the registry data, but the buildings of non-corporate businesses (which are recorded in the real estate registry) and their financial assets and liabilities (which are reported by third parties) continue to be.

[10]Similar wealth series are being produced for a growing number of countries, as published on the World Wealth and Income Database at `http://WID.world` (Alvaredo *et al.* 2017).

[11]One caveat is that we do not attempt to account for the unreported offshore holdings of Danish households in Switzerland and other tax havens. Zucman (2013) estimates that about 8% of the world's household financial wealth is held offshore globally. We refer the reader to Alstadsæter *et al.* (2017a,b) for an attempt at including offshore assets in the wealth distribution of different countries.

that individual-level pension wealth was added to the administrative data from 2012 onwards.[12]

Second, we impute assets not reported by third parties by capitalizing the respective income flows. Specifically, we compute non-corporate business assets by capitalizing business income (the capitalization rate is equal to the market value of business assets divided by the flow of business income reported on individual income tax returns), while we impute unlisted equities by capitalizing dividend income. Although these imputations introduce some noise at the micro-level, this noise is unlikely to bias our wealth shares in any particular direction. Importantly, we only make these imputations when computing the distribution of wealth in this section. For our main analysis of behavioral responses to wealth taxes, we focus on reported taxable wealth (thus excluding pensions) as this is the most appropriate outcome for this purpose.

2.3   Trends in Wealth Concentration

Figure 1 shows wealth shares in three broad classes: the bottom 50%, the next 40%, and the top 10%. Measured in this way, wealth inequality has been relatively stable in Denmark over the last three decades. Throughout the period, the bottom 50% of the distribution owns a tiny fraction of aggregate wealth: their assets are barely higher than their debts. Therefore, almost all wealth is owned by the richest half of the population, and it is shared about equally between the middle 40% and the top 10%. While the wealth shares in the figure are overall stable, wealth inequality did increase somewhat from the mid 1980s to the early 1990s. During this time, the top 10% wealth share grew while the bottom 50% wealth share shrank. This evolution was driven by the dynamics of asset prices, in particular housing prices, which fell significantly during this period. Because the share of housing in asset portfolios tends to be decreasing in the level of wealth, housing slumps hurt the bottom more than the top, leading to a rise in wealth inequality.

Figure 2 zooms in on the top of the wealth distribution — the sample that is more relevant for our tax reform study below — and contrasts Denmark with the United States. Several insights are worth noting. First, wealth inequality is markedly lower in Denmark than in the US. In 2012, the top 1% accounts for about 20% of total wealth in Denmark, whereas it accounts for almost 40% in the US. Average wealth in the population is similar in the two countries, but the top 1% are twice

---

[12]The imputation is done as follows. In 2012, we observe that about 40% of pension wealth belongs to wage earners while 60% belongs to retirees. We assume that these shares were the same before 2012. We then allocate the pension wealth of workers proportionally to their wage incomes (winsorized at the 99th percentile) and the pension wealth of retirees proportionally to their pension benefits paid out of pension funds. We have checked that the distribution of imputed pension wealth for the year 2012 is very close to the *observed* distribution of pension wealth for that year. Saez & Zucman (2016) use a similar imputation procedure for the United States.

as wealthy in the US as they are in Denmark.[13] Second, the gap between the two countries has widened over time. Top wealth shares were increasing in both countries until the late 1990s, but then they begin to diverge as wealth inequality stabilizes in Denmark while it continues to increase in the US. Third, the similarity between the two countries until the late 1990s and the subsequent divergence look more striking as we move into the extreme tail of the distribution. As shown in the bottom panel of Figure 2, the top 0.1% wealth share in Denmark was only 2-3 percentage points lower than in the US around year 2000, but then starts to diverge very strongly. If we consider top 0.01% wealth shares (not shown), they are essentially the same in the two countries at the turn of the century and then diverge.

To conclude, despite the reduction and ultimate abolition of the wealth tax in Denmark in the 1990s, wealth accumulation at the top of the distribution (relative to the population as a whole) has not picked up speed in Denmark as compared to the US. In other words, the aggregate patterns documented here do not provide a smoking gun for behavioral impacts of wealth taxes. Of course, this does not imply that wealth taxes did not affect wealth accumulation and wealth inequality. It simply means that if the wealth tax cuts caused wealth to grow faster at the top, this unequalizing force must have been offset by confounding equalizing forces. In our analysis of the causal effect of wealth taxation, we do find that lower wealth taxes cause wealth to grow faster.[14]

3 The Effect of Wealth Taxes: Evidence

3.1 Tax Variation and Empirical Strategies

Denmark taxed wealth on an annual basis until 1996. Taxable wealth equalled the total net wealth of households, excluding pension wealth. Taxable wealth components thus included cash, deposits, bonds, equities, housing, large durables and business assets, net of any debts. A number of these components were third-party reported by financial institutions, leaving little scope for tax evasion. But some components were self-reported, namely cash, durables, unlisted equities, non-corporate business assets, and assets held abroad.

Wealth was taxed at a flat rate above an exemption threshold. The exemption threshold varied

---

[13]The average person in the top 1% of the US distribution own net wealth of $9.3 million (roughly 40 times average wealth), as opposed to $4.8 million in Denmark (roughly 20 times average wealth).

[14]One important confounding reason why wealth inequality has stabilized in Denmark (despite wealth tax cuts) is likely to be the sharp rise of pension wealth, from around 50% of national income in the late 1980s to 178% in 2014. Because pension wealth is relatively equally distributed, rising pension wealth tends to reduce inequality. As shown by Chetty *et al.* (2014), automatic employer contributions to retirement accounts increase pension and total wealth substantially for middle-class Danish households.

over time (differentially for singles and couples) as we discuss below, but it was always above the 97th percentile of the household wealth distribution during the period we study. Wealth above the exemption threshold was taxed at 2.2% until two major reforms in the late 1980s and 1990s. Between 1989-91 the tax rate was reduced from 2.2% to 1%, while in 1996-97 the wealth tax was abolished entirely. These tax changes are illustrated in Figure 3.

This setting offers two sources of exogenous tax variation: the kink point at the exemption threshold and the tax reforms. Let us first consider the former. The kink point is very large as a tax rate jump of 2.2% on the stock of wealth translates into a very large tax rate jump on the return to wealth. This would seem to make a bunching strategy potentially compelling.[15] However, while bunching approaches are useful for uncovering evasion and avoidance responses to wealth taxes, they are not useful for uncovering real responses to such taxes. While this may be true even in the context of labor income taxation (see Kleven 2016), it is particularly true in the context of wealth taxation. Taxable wealth depends not only on individual decisions, but also on asset prices that are highly uncertain and move continuously through the tax year. Given such asset price movements, it would be virtually impossible for a taxpayer to bunch at the exemption threshold using real savings responses. Therefore, we will not pursue a bunching strategy as our main approach. We do start out by documenting bunching at the kink, but we view this evidence primarily as informative of avoidance responses around the threshold.

Given these considerations, our main analysis will be based on the tax reform variation. In particular, we focus on the 1989-reform rather than the subsequent elimination of the tax, because of a data limitation discussed earlier: After abolishing the wealth tax in 1997, Statistics Denmark no longer records purely self-reported wealth. This break in the taxable wealth series makes it difficult to study the wealth tax abolishment, and so we focus on the earlier tax cuts that do not have this limitation.[16] To estimate behavioral responses to the 1989-reform, we consider difference-in-differences (DD) approaches in which we compare treatment and control groups in a balanced panel of taxpayers. We develop two DD approaches that we now describe.

The first approach uses that the 1989-reform reduced the tax rate from 2.2% to 1%. To define groups that were differentially affected by this tax cut, we exploit the existence of a ceiling on

---

[15]As discussed earlier, the recent paper by Seim (2017) takes a bunching approach using a wealth tax kink in Sweden.

[16]One could consider using the market-value wealth series constructed in the previous section to provide evidence on the wealth tax abolishment. However, while this series is consistently measured over time, it does not precisely capture *taxable wealth* at the individual level. This series is useful for understanding the evolution of wealth inequality, but it is not sufficiently precise to estimate taxable income responses to tax reform. We therefore prefer to focus on the 1989-reform for which we observe exact taxable wealth for 8 years after the tax cuts.

the total tax liability from all personal taxes (income taxes, social security taxes, and wealth taxes) as a fraction of taxable income. This ceiling — known as *Det Vandrette Skatteloft* ("horizontal tax ceiling") — was in place to limit the total average tax rate on households with large wealth relative to income. The tax ceiling was set at 78% at the time of the 1989-reform. Whenever the total average tax rate exceeded this limit, tax liability would be reduced by the excess amount.[17] For households bound by the ceiling, the marginal tax rate on wealth was equal to zero — before and after the reform — making them a natural control group. Figure A.I in the appendix shows the fraction of taxpayers bound by the ceiling at different quantiles of the wealth distribution. We see that the ceiling starts binding for a substantial fraction of households as we move into the extreme tail of the wealth distribution, allowing us to estimate responses to wealth taxes by the super wealthy. We will compare taxpayers unbound by the ceiling (treatments) to taxpayers bound by the ceiling (controls) within the top percentile of the distribution.[18] We refer to this strategy as the "ceiling DD".

The second approach uses that the 1989-reform increased the exemption threshold for couples relative to singles. Before the reform, singles and couples faced the same nominal exemption threshold for wealth taxation. This is difficult to justify on equity grounds, because a couple is less wealthy in per capita terms than a single individual at the same level of household wealth. To rectify this issue, the exemption threshold for couples relative to singles was doubled between 1989-1992. These threshold changes are illustrated in Panel B of Figure 3. The implication of the reform is that couples in a certain range of the household wealth distribution — roughly those between the 97.6th and the 99.3rd percentiles — became exempt from wealth taxation, allowing us to estimate responses by the moderately wealthy. We will compare couples in the affected range (treatments) to singles in the same range or to couples outside the range (controls). We refer to this strategy as the "couples DD".

As usual, these difference-in-differences specifications rely on the assumption of parallel trends between comparison groups. This assumption is more difficult to satisfy when considering wealth as the outcome than when analyzing more standard outcomes such as labor supply or taxable income. The reason is that the value of wealth depends on asset prices, which move considerably

---

[17]Specifically, wealth tax liability is reduced first (but by no more than 50%) and then income tax liability is reduced next.

[18]A more conventional approach to studying these tax rate changes would be compare taxpayers above and below the exemption threshold. Such difference-in-differences designs have been used in many papers estimating taxable income responses (see Saez *et al.* 2012 for a review). As discussed in this literature, this type of strategy is generally not compelling due to the challenges of dealing with two potential confounders: non-tax changes in inequality (which we documented in the previous section) and mean reversion. For this reason, we do not pursue such a strategy here.

over time and affect groups differently depending on their portfolio composition. For example, if one group owns more equity and less housing than another group, the relative wealth trends in these two groups will be affected by non-tax changes in the prices of stocks and housing. Because our comparison groups do not have the same portfolio composition in the baseline, we have to control for this source of non-parallel trends. We do this in a non-parametric manner as we now describe.

Our empirical analysis starts from DD specifications of the following form

$$\log W_{it} = \sum_{j \neq 1988} \beta_j^C \cdot Year_{j=t} + \sum_{j \neq 1988} \beta_j^T \cdot Year_{j=t} \cdot Treat_i + \gamma_i + X_{it}'\psi + \nu_{it}, \tag{1}$$

where W_{it} denotes the wealth of individual i in year t, $Year_{j=t}$ is a dummy equal to one when the year equals t, $Treat_i$ is a dummy equal to one when individual i belongs to the treatment group, γ_i is an individual fixed effect, and X_{it}' is a set of non-parametric controls that we define below. The assignment of individuals to the treatment group depends on the strategy, either being unbound by the tax ceiling (ceiling DD) or being in a couple in the exempted range (couples DD). In each case, treatment status is based on pre-reform variables and to increase persistence we require that the individual has the same status in three consecutive pre-reform years (1986-88).[19] For each strategy presented below, we plot the evolution of wealth in the control and treatment groups after absorbing the non-parametric controls, i.e. we show the estimates $\hat{\beta}_t^C$ and $\hat{\beta}_t^C + \hat{\beta}_t^T$. The difference between these two series ($\hat{\beta}_t^T$) is our reduced-form estimate, or intention-to-treat (ITT) estimate, of the effect of wealth tax reform in year t.

To absorb non-tax trends driven by baseline differences in portfolio composition across groups, we calculate pre-reform (1988) shares of housing and equities in total wealth and include the following set of controls: (i) dummies for each pre-reform housing share decile interacted with dummies for each year, (ii) dummies for each pre-reform equity share decile interacted with dummies for each year. In other words, we allow each decile of the housing share and equity share distributions to have their own non-parametric time trends unrelated to wealth tax reform. Furthermore, we include (iii) dummies for each pre-reform income decile interacted with dummies for each year. This last set of controls absorbs any confounding effects of heterogeneity in savings rates across income levels.

A closely related alternative to these non-parametric controls for pre-reform portfolio compo-

---

[19]Therefore, individuals who switch status (say, between bound and unbound or between being couples and singles) during the three pre-reform years are dropped from the estimation sample.

sition and income would be DFL reweighting (DiNardo *et al.* 1996) on those same variables. In this case, we would reweight observations so that the two groups have the same distribution of portfolio shares and income in the pre-reform year, and implement a raw difference-in-differences estimation on the reweighted groups. We present results based on DFL reweighting and show that this gives similar results to those based on equation (1). Our DFL specifications will be less granular than the regression controls described above in order to avoid imprecision from cells with few or zero observations, but we will compare the two approaches at the same level of granularity.

Finally, in order to link the empirical evidence to parameters of models (as we do later), it is critical to translate the intention-to-treat (ITT) estimates into treatment-on-the-treated (TOT) estimates. Even though we assign treatment status using three pre-reform years (rather than just one), the treatment and control groups are not perfectly persistent over the post-reform period. We carefully document the persistence of comparison groups over time, and turn our ITT effects into TOT effects using a Wald estimator.

3.2   Bunching Evidence

We start by documenting bunching at the kink. As discussed above, bunching is useful for estimating avoidance or evasion responses to wealth taxes, but not for capturing real wealth accumulation responses. Figure 4 presents bunching evidence for the full population and pooling all years in which the wealth tax was in place. Panel A shows the observed distribution around the kink in bins of $20,000$ Danish Kroner (about $20,000/6 = 3,333$ US Dollars), while Panel B adds an estimate of the counterfactual distribution absent the kink. The counterfactual distribution is obtained by fitting a polynomial to the observed distribution, excluding data in a range around the kink, and extrapolating the fitted distribution to the kink.[20] This is the standard approach to estimating counterfactual distributions in the bunching literature as developed by Chetty *et al.* (2011) and Kleven & Waseem (2013). In the figure, we also report an estimate of excess bunching b, defined as the total excess mass around the kink (the area between the observed and counterfactual distributions) scaled by the height of the counterfactual distributions at the kink. This statistic can be interpreted as the number of bins by which bunchers are moving on average, and it is proportional to the compensated tax elasticity as first demonstrated by Saez (2010). We refer to the review by Kleven (2016) for details and discussions of the approach.

---

[20]The counterfactual distribution shown in the figure is based on a 5th-order polynomial and an excluded range of 6 bins around the kink (with more bins below than above due to the asymmetry in observed bunching).

The figure shows that there is some bunching at the exemption threshold, but that the amount is very modest. The bunching estimate $b = 0.386$ implies that bunchers reduce taxable wealth by 0.386 bins or $7,720$ Danish kroner.[21] This is about 0.3% of wealth at the threshold, which is a very small response considering the size of the tax incentive. The wealth tax rate jumps by 2.2% in the early years and by 1% in the later years, which translates into very large drops in the marginal after-tax return. The implied elasticity is therefore tiny. As a comparison, the Swedish study by Seim (2017) found somewhat larger bunching at a smaller kink, but even there the estimated elasticity of taxable wealth was small.

These findings suggest that avoidance and evasion responses to changes in wealth tax rates are modest. Two points are important to highlight regarding this interpretation. First, the finding of small avoidance responses applies to the sample of *moderately* wealthy people around the kink, not necessarily to the super wealthy located higher up. Second, the finding that avoidance *responses* to wealth tax changes are small does not necessarily imply that avoidance *levels* are small. A general insight from the compliance literature is the large evasion or avoidance in levels do not necessarily translate into large elasticities with respect to the tax rate (see e.g., Slemrod & Yitzhaki 2002; Kleven *et al.* 2011b).

In the appendix, we provide two additional figures on bunching. Figure A.II investigates heterogeneity in bunching across samples with different opportunities to avoid or evade. It compares employees and self-employed individuals (Panels A-B) as well as "ordinary" and "non-ordinary" taxpayers (Panels C-D). Non-ordinary taxpayers refer to individuals with extended tax returns due to complicated tax matters (the self-employed, those with wealth abroad, etc.). The figure shows that bunching is larger in samples with larger avoidance opportunity, but that bunching remains modest in all samples. Figure A.III compares bunching before the 1989-reform (when the tax rate jumped by 2.2% at the kink) and after the 1989-reform (when the tax rate jumped by 1% at the kink). As one would expect, bunching is much smaller after the reform than it was before. This provides prima facie evidence that taxpayers understood the incentives created by the reform and responded to them.

---

[21]When constructing the bunching diagram, we adjust the wealth levels in each year to 2014 values (using the national inflation index), and so the wealth response of $7,720$ DKK is measured in 2014 terms.

3.3   Ceiling DD: Responses by the Very Wealthy

We now turn to our preferred empirical strategies, starting with the ceiling DD laid out in section 3.1. As described there, we investigate responses to the 1989-reform by comparing taxpayers who are unbound by the tax ceiling (treatments) to taxpayers who are bound by the tax ceiling (controls). The treatment group experiences a reduction in the marginal wealth tax from 2.2% to 1%, while the control group experiences no change. Because the tax ceiling becomes effective only at the very top of the wealth distribution (see Figure A.I), we compare bound and unbound taxpayers located in the top 1% during the three years leading up to the reform (1986-88). We assign taxpayers to treatment and control groups based on their ceiling status in the same three pre-reform years, thus dropping taxpayers who frequently switch ceiling status. To further increase the persistence of the groups, we also drop observations who are only *marginally* bound by the tax ceiling. In the baseline specification, the bound group includes those whose wealth tax liability would have to fall by at least 20% for them to become unbound.

Figure 5 documents the evolution of taxable wealth in these comparison groups, absorbing trends related to pre-reform differences in portfolio composition or income using specification (1). Panel A shows the time series of log taxable wealth in the unbound group (red dots) and the bound group (black squares) between 1985-1996, with both series normalized to zero in the year before the reform. Panel B shows the differences between the unbound and bound groups in each year (relative to the pre-reform year), i.e. our difference-in-differences estimates of the effect of the reform on taxable wealth. Two key insights emerge from the figure. First, the two series are on parallel trends in the years leading up to the reform and then begin to diverge immediately after the reform. This clearly suggests that the lowering of the wealth tax rate had a positive effect on taxable wealth. Second, the effect on taxable wealth is growing over time, exactly as we would expect if the reform changed the real savings rate. Conversely, if the entire effect were a one-time avoidance adjustment (for example, a repatriation of hidden wealth abroad), we would see a different pattern.

The results in Figure 5 represent reduced-form or intention-to-treat (ITT) effects. The treatment assignment is based on pre-reform ceiling status, which is not perfectly persistent over time. We may think of pre-reform ceiling status as an instrument for current ceiling status (the latter being endogenous), in which case the figure shows estimates based on regressing the outcome directly on the instrument. Figure 6 converts the intention-to-treat (ITT) effects into treatment-on-the-treated (TOT) effects. Panel A documents the persistence in treatment status by showing, for the bound

and unbound groups, the fraction who retains their ceiling status over time. By construction, the unbound group is 100% unbound in the three pre-reform years, while the bound group is 0% unbound in those years. Over time, taxpayers may change their status either by switching between bound and unbound or by falling below the exemption threshold. The figure shows that the unbound group is very persistent (as less than 10% have changed status in 1996), while the bound group is much less persistent (as more than 40% have changed status in 1996). Therefore, eight years after the reform, there is a 50% difference in the fraction unbound between the two groups.

Panel B converts the ITT series into a TOT series by dividing the former with the differences in the fraction from Panel A (Wald estimator). For example, given the 50% difference in treatment intensity in 1996, the ITT effect is scaled by about two in order to obtain the TOT effect for that year. When implementing this across time, the dynamically growing effect on taxable wealth is enhanced, an implication of the gradual reduction in persistence. While the ITT series appear to stabilize towards the end of the time window, this is not true for the TOT series. The treatment effect on log wealth is equal to 0.322 in 1996, an increase of about 30% over 8 years.

When considering the effects on wealth, it is important to keep in mind that these include both behavioral and mechanical effects. The tax reform raises the after-tax rate of return on wealth, which would increase wealth accumulation even if behavior were fixed. How much of the effects can be explained by such mechanical effects? This is not an entirely trivial question to answer due to two complications. The first complication is that the mechanical tax savings cannot be based on observed wealth as this includes any behavioral responses, but must be based on a measure of counterfactual wealth. Consistent with the difference-in-differences design, we impute counterfactual wealth after the reform as observed wealth before the reform (in 1988) plus the growth rate in wealth experienced by the control group. The second complication is that the tax savings earned in a given year will grow over time according to a rate of return that is not directly observed in the data. In the simulation presented below, we assume an annual rate of return equal to 7%. This falls within the range of existing estimates of wealth returns at the top of the distribution (see e.g., Fagereng *et al.* 2016) and it corresponds to what we assume in the calibration exercise presented later. Based on these assumptions, we calculate the cumulative mechanical tax savings in each year due to the reduction in the wealth tax rate from 2.2% to 1%.[22] The details of this calculation

---

[22]Besides the tax savings from the reduced tax rate, we also account for the fact that a minor fraction of the treatment group benefit from the increased exemption threshold for couples. Because we are considering taxpayers within the top 1% and the threshold for couples is raised only just above the 99th percentile cutoff, this part of the reform has a

are provided in Appendix A.

The results are presented in Figure 7. The figure shows the series of total and behavioral effects on wealth, with the differences between the two being the mechanical effects. We see that the mechanical effects are sizeable, but that most of the effects come from changed behavior. The mechanical effect on log wealth in 1996 equals 0.073, which is less than a quarter of the total effect. The reason why the mechanical effects are relatively modest has to do with the progressive nature of wealth tax: taxes are saved only above the exemption threshold located around the 98th percentile of the wealth distribution. That is, while the behavioral responses are governed by the change in the *marginal* after-tax return (which is very large), the mechanical effect is governed by the change in the *average* after-tax return (which is more modest). This is a nice feature of the quasi-experiment we are analyzing. If we had considered similar rate changes in a proportional wealth tax, the mechanical effects would have been much larger.

In the appendix, we provide three robustness checks. The first check in Figure A.IV considers an alternative wealth measure in which we exclude self-reported, easy-to-evade components. Specifically, we subtract cash, durables, and foreign wealth from taxable wealth. We refer to this measure as "third-party reported wealth", although some self-reported components remain.[23] The estimated effects on third-party reported wealth are essentially the same as on taxable wealth. While this suggests that real asset investments are changing in response to the tax cuts, it does *not* rule out evasion and avoidance responses. It is possible that the increase in third-party reported assets is driven in part by shifting from previously unreported items. The best evidence we have of small avoidance and evasion responses is the bunching evidence presented above, but this applies to the moderately wealthy taxpayers around the threshold, not to the very wealthy analyzed here.

The second check in Figure A.V investigates a control group that consists only of those who are *strongly* bound by the tax ceiling, thereby increasing the persistence of the control group. Specifically, we include only households whose wealth tax liability would have to fall by at least 50% (as opposed to 20% in the baseline) for them to become unbound. Reassuringly, the results are essentially unchanged when considering this alternative specification, and this holds for both taxable wealth and for third-party reported wealth. The intention-to-treat effects are somewhat larger, but because the fractions treated are correspondingly larger, the treatment-on-the-treated effects are the same.

---

relatively small impact in the ceiling design.

[23]Unlisted equities and non-corporate business assets are both self-reported. We are not able to exclude these components from taxable wealth, because we do not observe them at the individual level throughout the period.

Finally, Figure A.VI compares the results from our baseline regression specification (1) to the results obtained from DFL reweighting. That is, instead of controlling for differences in pre-reform portfolio composition and income in a regression framework, we reweight the groups to have the same pre-reform distribution of those variables. The figure compares two different DFL specifications (left panels) to the corresponding regression controls (right panels). We consider a parsimonious specification (adjusting for housing share deciles only) and a richer specification (adjusting for housing share, equity share, and income quintiles). In the richer specification, we consider quintiles rather than deciles to reduce imprecision from cells with few or zero observations.[24] The main insight from the figure is that the DFL approach produces results that are similar to the regression approach in terms of pre-trends and the dynamic build-up of the effect after the reform. While the magnitude of the effects is almost the same in the parsimonious specification, the DFL approach yields somewhat smaller effects in the rich specification.

3.4   Couples DD: Responses by the Moderately Wealthy

In this section we investigate responses by the moderately wealthy. We exploit that the 1989-reform doubled the exemption threshold for couples, thus eliminating wealth taxation for couples between the 97.6th and the 99.3rd percentiles of the household wealth distribution. We compare these households to two alternative control groups: (i) couples located outside the exempted range (either above or below within the top 5%), (ii) singles located inside the exempted range. As above, the groups are based on variables (marital status and wealth level) in the three pre-reform years 1986-88. The advantages of the first specification is that we are comparing couples to couples (which is good for comparability), and that the identifying tax variation is large as the treatments have their wealth tax eliminated while the controls experience almost no change.[25] The advantages of the second specification is that we are comparing households at the same wealth level, but on the other hand the identifying tax variation is smaller as the control group benefit from the tax rate cut from 2.2% to 1%. The fact that there is a significant tax change in the comparison group may make it harder to pick up any effects, and it implies that we have to assume homogeneous responsiveness across groups. Therefore, on balance, we prefer the first specification and present this one first.

---

[24]Reweighting simultaneously on housing share, equity share, and income deciles would imply $10^3 = 1000$ cells (as opposed to $5^3 = 125$ cells when doing it in quintiles).

[25]To be precise, the part of the control group located above the exempted range (about 20% of the control group) gets a wealth tax cut from 2.2% to 1%. So in aggregate there is a small change in the control group.

Figures 8-10 present our findings from comparing couples inside and outside the exempted range. The three figures are constructed in the same way as the three preceding figures for the ceiling DD. In Figure 8, we see that the treatment and control groups are on almost perfectly parallel trends before the reform, but that they begin to diverge after the reform. There is a delay of 1-2 years in the effect, which can be explained by the fact that the threshold change was implemented gradually (between 1989-92) while we are comparing groups based on the full implementation. As a result, the treatment group is only marginally treated in the first years after the reform. Over time there is a gradually increasing effect on the stock of wealth, but the magnitude of the effect is clearly smaller than those obtained for the ceiling DD. One reason may be that the moderately wealthy are less responsive than the very wealthy — and we will end up concluding that this is indeed the case — but there are two other reasons why the effects in Figure 8 are relatively modest: one is related to persistence and the other to mechanical effects.

Figure 9 investigates the persistence of the treatment and control groups, and converts the ITT effects into TOT effects. The treatment intensity in the treatment group starts from 100% and falls gradually over time (due to changes in marital status and wealth levels), while the treatment intensity in the control group starts from 0% and increases gradually over time. In the last year of the event window, the difference in treatment intensities is slightly less than 40%. The couples DD is therefore less persistent than the ceilling DD, explaining some of the difference in the results from the two strategies. When we convert the intention-to-treat effects into treatment-on-the-treated effects in Panel B, the effects become considerably larger and the gradual dynamic build-up is strongly enhanced. The effect on the stock of wealth after eight years is about 10%.

Figure 10 splits the total effect on wealth into behavioral and mechanical effects. The basic methodology is the same as for the ceiling DD (with details provided in Appendix A): we calculate annual tax savings for the treatment group using a measure of counterfactual wealth, and we simulate cumulative tax savings assuming an annual rate of return equal to 7%. The figure shows that the mechanical effect is much smaller for the couples DD than for the ceiling DD, providing an additional reason why the total effect is smaller. The main explanation is that the couples DD identifies responses for moderately wealthy households, who were not very far above the exemption threshold. As a result, the reform did not change their *average* return very much — while changing their marginal return by a lot — and so most of the effect is behavioral and very little of it is mechanical.[26] We estimate that the behavioral effect on wealth in 1996 is almost 10%,

_____

[26]There is an additional, quantitatively less important reason why the mechanical effect is smaller in the couples DD,

or about 40% of the behavioral effect we found for the very wealthy in the ceiling DD.

The appendix provides three robustness checks. Figure A.VII presents estimates from the same specification, but considering third-party reported wealth as the outcome. The figure shows ITT and TOT estimates, and it makes the split between behavioral and mechanical effects. The empirical patterns are similar to those for taxable wealth, but the effects are slightly smaller. Figure A.VIII presents estimates from a specification where we use singles in the same wealth range as couples (97.6th-99.3rd percentiles) to form a control group. In this case, the effects are considerably smaller (about half the size) and the empirical patterns are not as persuasive. As discussed above, the likely reason is that the control group in this specification also gets a substantial wealth tax cut (from 2.2% to 1%), implying that we are identifying off the differential tax cuts in the two groups and rely on an assumption of homogeneous responsiveness. Finally, Figure A.IX compares the regression-based approach to DFL reweighting. We see that the two different strategies produce almost identical dynamic patterns and magnitude of effects, which is reassuring.

Table 1 summarizes the main results from our quasi-experimental analysis. It shows estimates from both the ceiling DD and the couples DD, and it shows the different specifications of the control group and the non-parametric controls for pre-trends. All of the estimates are statistically significant at the 1% level. The effects from the ceiling DD are consistently larger than the effects from the couples DD. Some of this can be explained by differences in mechanical effects, but not all of it. Therefore, we find that the very wealthy people are are more responsive than moderately wealthy people. The dynamic patterns we have documented are consistent with real savings effects, and it is conceivable that very wealthy people are more elastic in their savings behavior. At the same time, it is likely that the very wealthy respond to wealth taxes partly through avoidance or evasion. This channel of response may be less available for the moderately wealthy, and in fact our bunching evidence suggests very small avoidance responses for this population. In the next section we develop a model of responses to wealth taxes in which we focus on the real savings channel.

namely that a minor fraction of the control group (those above the exempted range) benefits from the tax rate cuts and therefore also experiences mechanical tax savings. We net out the tax savings in the control group in Figure 10, which further reduces the mechanical effect.

4 The Effect of Wealth Taxes: Theory

4.1 Lifecycle Model With Utility of Residual Wealth

In this section we develop a model for studying the effects of wealth taxation on the wealthy. Our goal is to construct a model that is sufficiently simple to derive analytical results, but at the same time rich enough to facilitate interpretation of the empirical results and to allow for informative calibration exercises. To understand what the key features of such a model should be, we highlight two empirical facts that will be documented in detail in the next section. First, wealthy people tend to be older people. For example, almost 80% of those in the top 1% of the wealth distribution are above age 50, as opposed to only 31% in the general population. Second, wealthy people continue to accumulate wealth into very old ages and therefore die with lots of wealth.

To match these empirical facts, our model incorporates utility of residual wealth. This may be interpreted as capturing a bequest motive — and we will refer to it as such — but it may also capture other utility-of-wealth motivations (see Saez & Stantcheva 2017 for a discussion of different mechanisms). The specific mechanism is not important for our purposes.[27] While our model accounts for the bequest motive as well as the standard lifecycle motive for saving, it abstracts from precautionary savings and uncertainty. The precautionary savings motive matters for the lower tail of the distribution, but it is second order for understanding savings behavior at the top of the wealth distribution (see e.g., Carroll 2002; De Nardi 2004).[28]

Households live for T periods and their preferences are specified as follows

$$\frac{\sigma}{\sigma - 1} \sum_{t=0}^{T} \delta^t \left(c_t\right)^{\frac{\sigma-1}{\sigma}} + \delta^T V\left(W_{T+1}\right), \tag{2}$$

where c_t is consumption in period t, W_{T+1} is wealth at the end of life (bequests), σ is the elasticity

---

[27]We will use the model (together with the quasi-experimental moments) to estimate the long-run responsiveness of wealth accumulation to wealth taxation. This depends on the curvature of utility from wealth (which we estimate structurally), but it does not depend directly on the specific reason for utility from wealth. On the other hand, the specific reason may matter for *normative* policy analysis. For example, utility of wealth due to warm-glow of bequests will in general have different optimal tax implications than utility of wealth due to social status, because the former is associated with positive externalities (calling for Pigouvian subsidies) while the latter is associated with negative externalities (calling for Pigouvian taxes). However, in either case, the long-run elasticity of capital supply that we estimate is a key parameter, because it determines the fiscal externality against which we would trade-off the potential benefits from redistribution and externalities.

[28]We also abstract from labor supply responses to wealth taxes. Existing evidence from Denmark suggests that labor supply is relatively inelastic to labor taxes (Kleven & Schultz 2014; Kleven 2014), suggesting that labor supply is also inelastic to capital taxes and perhaps especially for the population of older, wealthy people studied here. Furthermore, our explorations of the data shows no evidence of labor supply responses to wealth taxation when using the same empirical strategies (ceiling DD and couples DD) as those used for estimating savings responses.

of intertemporal substitution (EIS), and δ is the discount factor. To capture utility of bequests, we adopt the following parameterization

$$V\left(W_{T+1}\right) = A\frac{\alpha}{\alpha-1}\left(\frac{W_{T+1}}{A}\right)^{\frac{\alpha-1}{\alpha}}, \tag{3}$$

where A determines the strenght of the bequest motive (under $A = 0$ the model corresponds to the pure lifecycle model) and α is a bequest elasticity. This is a warm-glow bequest motive as introduced by Andreoni (1989, 1990) and used for studying estate taxation by for example Farhi & Werning (2010), Piketty & Saez (2013), and Kopczuk (2013a). For simplicity of exposition, we abstract from estate taxes (as our focus is on wealth taxes rather than on wealth transfer taxes) and model warm glow as a function of gross wealth.

In each period, there is a tax rate τ on household wealth above an exemption threshold $\bar{W}$. For someone with wealth above the exemption threshold in period t, the budget constraint is given by

$$\begin{aligned} c_t &= y_t + RW_t - \tau R\left(W_t - \bar{W}\right) - W_{t+1} \\ &= y_t + (1-\tau)RW_t + \tau R\bar{W} - W_{t+1}, \end{aligned} \tag{4}$$

where y_t is (exogenous) labor income net of income tax, W_t is wealth at the beginning of the period, and R is the gross rate of return. We assume that R is time-invariant, but this is straightforward to generalize and has no important implications for our results. The second line of the budget constraint (4) is a "virtual income" representation: it writes the budget as if the net-of-tax return equals $(1-\tau)R$ on all units of wealth, but provides a lump-sum income of $\tau R\bar{W}$ to compensate for the fact that the tax is not paid below the threshold. Combining all the per-period budget constraints, we can express the lifetime budget constraint as

$$\sum_{t=0}^{T}\frac{c_t}{\left((1-\tau)R\right)^t} + \frac{W_{T+1}}{\left((1-\tau)R\right)^T} = \sum_{t=0}^{T}\frac{y_t}{\left((1-\tau)R\right)^t} + \sum_{t=1}^{T}\frac{\tau R\bar{W}}{\left((1-\tau)R\right)^t} + W_0^n, \tag{5}$$

where $W_0^n \equiv (1-\tau)RW_0 + \tau R\bar{W}$ is initial (exogenous) wealth after tax.

Households maximize lifetime utility (2)-(3) subject to the lifetime budget constraint (5) with respect to consumption and bequests. The first-order conditions for c_t and c_{t+1} yield the standard Euler equation,

$$c_{t+1} = \left(\delta(1-\tau)R\right)^{\sigma}c_t, \tag{6}$$

while the first-order conditions for W_{T+1} and c_T give

$$W_{T+1} = A c_T^{\alpha/\sigma}. \tag{7}$$

The solution to the model is described by the lifetime budget (5), the Euler equations (6) for all t, and the bequest condition (7). These conditions determine $c_0, ..., c_T$ and W_{T+1}. Wealth W_t in any given period can then be backed out using the per-period budget constraints.

Using the Euler equations in each period, we can write consumption in period t and bequests in terms of consumption in period 0, i.e.

$$c_t = (\delta (1 - \tau) R)^{t\sigma} c_0, \tag{8}$$

$$W_{T+1} = A (\delta (1 - \tau) R)^{T\alpha} c_0^{\alpha/\sigma}. \tag{9}$$

Inserting these conditions into (5), we can express the choice of c_0 as

$$\sum_{t=0}^{T} q_t \cdot c_0 + q_b \cdot c_0^{\alpha/\sigma} = \sum_{t=0}^{T} \frac{y_t}{((1-\tau) R)^t} + \sum_{t=1}^{T} \frac{\tau R \bar{W}}{((1-\tau) R)^t} + W_0^n, \tag{10}$$

where $q_t \equiv \frac{(\delta(1-\tau)R)^{t\sigma}}{((1-\tau)R)^t}$ denotes present-value expenditures on consumption in period t relative to period 0, and $q_b \equiv \frac{A(\delta(1-\tau)R)^{T\alpha}}{((1-\tau)R)^T}$ denotes present-value expenditures on bequests relative to consumption in period 0. This expression is useful for characterizing the effects of wealth taxes.

4.2   The Effect of Wealth Taxes

Consider a permanent change in the wealth tax rate, $d\tau$, holding the exemption threshold $\bar{W}$ constant. The tax change is announced in period 0, and may affect wealth from the end of this period, W_1. Initial after-tax wealth W_0^n is pre-determined. We investigate the effect on households who are above the threshold $\bar{W}$ (and stay above over time), as opposed to the effect on households who are sometimes below and sometimes above the threshold over their lifetime. The former scenario is simpler to analyze, and it fits our quasi-experimental setting in which we estimate responses by households above the exemption threshold. The potential response by those who are below the exemption threshold, but expect to rise above it in the future, is not captured by our empirical design and would be very hard to estimate in general.[29]

---

[29]Such responses — if they are empirically relevant — cannot be persuasively identified using within-country variation, because everyone is potentially treated and the treatment intensity is unobservable to the econometrician. One

We characterize analytically how the reduced-form effect of changing the wealth tax rate —
what we have estimated empirically — relates to the structural parameters of the model. We start
by deriving the effect of taxes on first-period wealth W_1, and then show how the effect accumulates
over time. The effect of taxes includes both substitution and wealth effects. To characterize the
wealth effect, it is useful to define the amount of initial resources a household would have to
receive to be able to afford an unchanged bundle of consumption and bequests when the net-of-
tax return changes. This can be obtained by differentiating the lifetime budget constraint (5) with
respect to $1 - \tau$, holding behavior $\{c_t\}_0^T$, W_{T+1} constant but allowing initial wealth to adjust. We
denote this compensating change in initial wealth by dW_0^C.

We may state the following proposition:

Proposition 1 (First-Period Reduced-Form Effect). *Consider a permanent change in the wealth tax rate*
τ from period 0 onwards. The reduced-form elasticity of first-period wealth W_1 with respect to the net-of-tax
rate $1 - \tau$ can be expressed as

$$\frac{dW_1}{d\left(1 - \tau\right)} \frac{1 - \tau}{W_0} = \sigma \cdot \left\{ \frac{\sum_{t=0}^{T} t q_t}{\sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\alpha/\sigma - 1}} - \frac{c_0}{W_0} \right\} + \alpha \cdot \left\{ \frac{T q_b}{\sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\alpha/\sigma - 1}} - \frac{c_0^{\alpha/\sigma}}{W_0} \right\}$$
$$+ \frac{dW_0^C}{d\left(1 - \tau\right)} \frac{1 - \tau}{W_0} \cdot \left\{ \frac{1}{\sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\alpha/\sigma - 1}} \right\}, \tag{11}$$

where $q_t \equiv \frac{(\delta(1-\tau)R)^{t\sigma}}{((1-\tau)R)^t}$, $q_b \equiv \frac{A(\delta(1-\tau)R)^{T\alpha}}{((1-\tau)R)^T}$, and $dW_0^C \leq 0$ is a compensating wealth change allowing
the household to afford an unchanged bundle of consumption and bequests when $1 - \tau$ changes. The first
term is a substitution effect on consumption (positive), the second term is a substitution effect on bequests
(positive), and the third term is the wealth effect (negative).

Proof: See Appendix B. □

This result shows that the reduced-form elasticity of wealth (one period after the tax change) is
an involved function of all the parameters of the model. There are three qualitative effects on
wealth accumulation. First, there is a substitution effect on consumption. A larger net-of-tax return
induces households to shift consumption to later in life, thereby increasing wealth accumulation.
This effect is proportional to the EIS σ.[30] Second, there is a substitution effect on bequests. A larger

---

would have to make comparisons across economies with different wealth taxes, raising a number of empirical chal-
lenges. The evidence on Denmark vs the US presented in section 2.3 is an example of such a cross-country comparison,
and our discussion of this evidence illustrates the confounders that make interpretation difficult.

[30]Our wording is somewhat loose here. The effect is only proportional in σ when taking the complicated term in
braces (which itself depends on σ) as given.

net-of-tax return reduces the price on bequests, further increasing wealth accumulation. This effect is proportional to the bequest elasticity α, and it depends on the weight of the bequest motive A. The bequest effect vanishes when A (and therefore q_b) goes to zero, but can be important when A is large. We show later that A has to be very large to rationalize the lifecycle profiles of wealthy people, which puts the bequest elasticity α at center stage. Finally, there is the wealth effect. A larger net-of-tax return increases lifetime resources, which leads to larger consumption and lower savings. The presence of the wealth effect implies that the total reduced-form effect is ambiguous in sign.

A simplifying special case is where bequests and consumption goods are equally elastic, i.e. $\alpha = \sigma$. This is a natural benchmark assumption if bequests are viewed as future consumption (for the next generation). With $\alpha = \sigma$, we obtain the following result:

Proposition 2 (First-Period Reduced-Form Effect Simplified). *Assuming $\alpha = \sigma$ (bequest elasticity equals the EIS), the reduced-form elasticity of first-period wealth W_1 with respect to the net-of-tax rate $1 - \tau$ simplifies to*

$$\frac{dW_1}{d(1-\tau)} \frac{1-\tau}{W_0} = \sigma \cdot \left\{ \frac{\sum_{t=0}^{T} t q_t + T A q_T}{\sum_{t=0}^{T} q_t + A q_T} \frac{c_0}{W_0} \right\} + \frac{dW_0^C}{d(1-\tau)} \frac{1-\tau}{W_0} \cdot \left\{ \frac{1}{\sum_{t=0}^{T} q_t + A q_T} \right\}, \quad (12)$$

where $q_t \equiv \frac{(\delta(1-\tau)R)^{t\sigma}}{((1-\tau)R)^t}$ and $dW_0^C \leq 0$ is the compensating wealth change allowing the household to afford an unchanged bundle of consumption and bequests when $1 - \tau$ changes. The first term is the substitution effect on both consumption and bequests (positive) and the second term is the wealth effect (negative).

Proof: Follows from setting $\alpha = \sigma$ (and thus $q_b = A q_T$) in equation (11). □

This result provides a simpler characterization of the reduced-form elasticity. There is only one substitution term (capturing both consumption and bequest responses) and it is proportional to the structural elasticity $\sigma = \alpha$. We do not rely on the assumption of $\sigma = \alpha$ in the structural estimation presented below, but identify these two parameters separately. Given the lack of evidence on bequest elasticities (see Kopczuk 2009, 2013b), this is a contribution in itself.

The one-period effect derived above is helpful for establishing economic intuition, but our empirical analysis provides estimates over more than one year. We have estimates of the reduced-form elasticity $\frac{dW_t}{d(1-\tau)} \frac{1-\tau}{W_t}$ in each year t over 8 years. How does the theoretical effect of wealth taxes accumulate over time? We now consider the effect in any period t as a function of the effect in period 1 provided above.

From the per-period budget constraint (4), we have

$$W_t = y_{t-1} + (1 - \tau) R W_{t-1} + \tau R \bar{W} - c_{t-1}.$$

We may substitute out W_{t-1} using the one-period lagged equation, and then substitute out W_{t-2} using the two-period lagged equation, and so on. This process allows us to write

$$W_t = [(1 - \tau) R]^{t-1} W_0^n + \sum_{j=0}^{t-1} (y_j - c_j) [(1 - \tau) R]^{t-1-j} + \sum_{j=1}^{t-1} \tau R \bar{W} [(1 - \tau) R]^{t-1-j}. \quad (13)$$

Thus, W_t equals the sum of initial net wealth W_0^n with returns, t periods of savings $y_j - c_j$ with returns, and the virtual income adjustment with returns. Using this expression, we may state the following proposition:

Proposition 3 (Period-t Reduced-Form Effect). *Consider a permanent change in the wealth tax rate τ from period 0 onwards. The reduced-form elasticity of wealth in period t, W_t, with respect to the net-of-tax rate $1 - \tau$ can be expressed as*

$$\frac{dW_t / W_t}{d(1 - \tau) / (1 - \tau)} = dM + dB, \quad (14)$$

where dM is a mechanical effect given by

$$
\begin{aligned}
dM &= (t - 1) [(1 - \tau) R]^{t-1} \frac{W_0^n}{W_t} + \sum_{j=0}^{t-1} (t - 1 - j) [(1 - \tau) R]^{t-1-j} \frac{y_j - c_j}{W_t} \\
&\quad - \sum_{j=1}^{t-1} \left(1 - \frac{\tau (t - 1 - j)}{1 - \tau} \right) [(1 - \tau) R]^{t-j} \frac{\bar{W}}{W_t}, \quad (15)
\end{aligned}
$$

and dB is a behavioral effect given by

$$dB = \sum_{j=0}^{t-1} \frac{q_j}{[(1 - \tau) R]^{1-t}} \left\{ \frac{dW_1 / W_0}{d(1 - \tau) / (1 - \tau)} \frac{W_0}{W_t} - j\sigma \frac{c_0}{W_t} \right\}. \quad (16)$$

In equation (16), the first term in braces, $\frac{dW_1 / W_0}{d(1-\tau)/(1-\tau)}$, is characterized in Proposition 1 for the general case and in Proposition 2 for the special case of $\alpha = \sigma$.

Proof: See Appendix B. $\square$

While the first-period effect consists only of behavioral terms, the t-period effect consists of both behavioral and mechanical terms. The mechanical term reflects that, as $1 - \tau$ increases, the individual earns larger net-of-tax returns on initial wealth and savings in each period. This increases

26

wealth over time, holding consumption/savings behavior fixed. As a result, estimating significant reduced-form effects on wealth does not necessarily imply that people are elastic in their consumption or bequest behavior. However, as we saw in the previous section, the mechanical effect is a minor fraction of the total effect in our empirical setting due to the progressive design of the wealth tax. In addition to the mechanical effect, there will be behavioral effects as captured by the expression in (16). These consist of the one-period effect derived earlier (substitution and wealth effect) accumulating over time, along with additional substitution effects on consumption from period 1 onwards.

What have we learned from this theoretical exercise? First, the elasticity of wealth with respect to taxes is very far from a structural parameter. It is endogenous to all the other parameters of the dynamic setting, and its size depends mechanically on the time horizon of the estimate. Second, the magnitude of any reduced-form elasticity is most naturally assessed in terms of the structural elasticities of intertemporal substitution (σ) and bequests (α) implied by it. This requires a theoretical model and taking a stand on other parameters. It is in general possible to observe "large" effects under modest structural elasticities due to the mechanical effects. Third, the theoretical characterization allows us to calibrate the model to quasi-experimental moments obtained over the short-medium run, and then use the calibrated model to assess long-run effects. Such long-run simulations would rely on parametric assumptions, but in a way that respects shorter-run non-parametric moments. Given the gradual, dynamic nature of wealth accumulation, it would be difficult (or impossible) to capture the long-run effects without a parametric model.

5 Connecting Theory and Evidence: Long-Run Effects

5.1 Empirical Lifecycle Profiles of Wealth

To calibrate our model, we start by studying the empirical lifecycle profiles of wealth among households at the top of the distribution. Besides informing our calibration exercise, these lifecycle profiles will provide qualitative insights on the savings behavior of the rich and contribute to an area where we have relatively little evidence. Because we have household-level information on wealth for the full population over a long time period, we are able to present particularly clean and striking evidence.

Figure 11 shows lifecycle profiles of wealth between the ages of 20 to 90 for the full population

(in Panel A) and for the top percentiles of the population (in Panel B).[31] We highlight the following key points regarding the construction of the figure. First, the graphs show profiles of *taxable* wealth and therefore do not include pension wealth. Pension wealth is not very important at the top of the distribution (which is our main interest), but they are significant when considering the full population. Second, the graphs show profiles of *normalized* log wealth. Denoting log wealth for individual i at age a in year t by $\log(W_{iat})$, we define normalized log wealth as $\omega_{iat} \equiv \log(W_{iat}) - \mathsf{E}\left[\log(W_{iat}) \mid t\right]$. That is, we normalize log wealth for each individual in each year by the average wealth in the population in that year, so that the lifecycle profiles are not confounded by asset price inflation.[32] Third, we consider *unbalanced* panels of individuals, because this allows us to show very wide age ranges. We consider balanced panels in narrower age ranges below. Finally, when showing individuals in the top percentiles of the distribution, we select individuals who are in the top $p\%$ *at some point during their lives*, but not necessarily in every year. That is, we do not condition the sample on being in the top $p\%$ at each age, but allow them to build up wealth gradually and draw down wealth at the end of life. This way of selecting the sample is more informative for understanding the lifecycle savings behavior of the wealthy.[33]

The following key findings emerge from Figure 11. The average person in the population (Panel A) accumulates wealth until just after age 60 and draws down wealth thereafter. In other words, the average person reaches her wealth peak around the age of retirement, consistent with the predictions of a pure lifecycle model without any bequest motive or utility of wealth.[34] At the same time, wealth is still higher than average wealth in the population (the horizontal line at zero) at age 90, suggesting that the pure lifecycle savings motive is not the only factor at play.

When turning to the wealthiest segments of the population (Panel B), the picture changes drastically. Wealthy individuals tend to accumulate wealth through most of their lifetime. Not until after the age of 80 do we observe any draw-down of wealth, and even then the draw-down is only marginal. For example, those who reach the top 1% of the wealth distribution during their lifetime, surpasses the exemption threshold for the wealth tax (demarcated by the horizontal line) around age 50 and stay well above it until age 90. Figure A.X in appendix shows the same type of graph,

---

[31]Here we consider individuals (rather than households) as the units of analysis, but split household wealth in married couples equally between the spouses.

[32]This normalization implies that the mean of ω_{iat} in each year t equals zero. The mean of ω_{iat} in each age bin is calculated as $\mathsf{E}\left\{\mathsf{E}\left[\omega_{iat} \mid a, t\right] \mid a\right\}$, i.e. we first calculate means in bins of age a and year t, and then we calculate means in age bins over all years. This gives different calendar years the same weight in the calculated means.

[33]However, we do condition the sample on being in the top $p\%$ for at least three years (not necessarily three consecutive years) in order to avoid noise from transitory wealth shocks.

[34]This pattern would be reinforced by including pension wealth, because in general people pay into such schemes during their working lives and draw benefits during retirement.

but with wealth percentiles (instead of amounts) on the y-axis and including the extreme top of the distribution. There we see that those who reach the top 1% during their lifetime are, on average, located above the 98th percentile cutoff at the age of 90. Those who reach the top 0.5% are located at the 99th percentile cutoff at age 90, while those who reach the top 0.1% are located at the 99.8th percentile cutoff at age 90. There must be some form of utility of residual wealth (due to a bequest motive or another mechanism) in order to rationalize these empirical patterns.

Have these lifecycle patterns changed over time? Specifically, have they changed from before to after the 1989-reform that we analyze in this paper? Figure 12 investigates this question by comparing the lifecycle profiles of wealth before the reform (1980-88) and after the reform (1989-96).[35] In this figure, we focus on individuals in the top 1% of the wealth distribution, selected in the same way as the top 1% in the previous graph. We focus on the age range 60-90 during which the top 1% is always above the wealth tax threshold on average. The series are otherwise constructed in the same way as in the previous figure. We see that there are no striking changes in the lifecycle profile over time: both profiles stay quite flat into very old ages and provide no indication that the wealth tax cuts increased wealth accumulation (or reduced draw-down) for people close to death. This time series comparison cannot be taken as causal evidence of the effect of the reform, but we do use one aspect of it in the calibration below. Specifically, in the calibration of the bequest motive, we use the fact that wealth accumulation in the very last years of life (after the age of 80) is unchanged from before to after the reform.

A potential concern with these graphs, especially in terms of interpreting what happens at very old ages, is that they are based on unbalanced samples. People drop out of the samples as they die, which may affect the wealth profiles due to selection on mortality. As we move out into the tail of the age distribution, we are increasingly considering people who live long and such people may tend to be wealthier. If so, this would understate the within-person wealth draw-down at the end of life. To investigate this issue, Figure A.XI shows lifecycle profiles of wealth for the top 1% in a balanced sample. Panel A of the figure compares wealth profiles in balanced and unbalanced samples between 70-90 years of age, while Panel B compares wealth profiles before the reform (1980-88) and after the reform (1989-96) in the balanced sample. Panel A shows that the balanced sample do feature stronger wealth draw-down at the end of life, consistent with some selection on mortality. However, the extent of this draw-down is small. Moreover, it is driven

---

[35]We define the after-reform period as 1989-96 in order to avoid any confounding effects from the wealth tax abolishment in 1997 and from the change in measured wealth after the abolishment of the wealth tax.

largely by changes after the wealth tax abolishment in 1997, which can be seen from Panel B. This graph shows that both before and after the 1989-reform — but within the wealth tax period 1980-96 — the wealth profile among the very old is almost completely flat even in the balanced sample. Overall, considering balanced samples do not change the fundamental insights from the unbalanced samples.

5.2 Calibration

To study the long-run effect of wealth taxes on the wealthy, we calibrate the model from section 4 to fit the empirical lifecycle profile of wealth at the top along with the quasi-experimental estimates of the impact of wealth tax reform. Because our theoretical model is a representative agent framework, we calibrate it to fit the average wealth profile in specific wealth ranges that vary by experiment. In particular, the quasi-experimental analysis considered two different empirical strategies and samples: the couples DD gave treatments effects roughly between the 98th and 99th percentile cutoffs, while the ceiling DD gave treatment effects above the 99th percentile cutoff. The lifecycle profiles in Figure A.X show that the couples DD is captured well by the "top 1% sample" (i.e., the sample of those who reach the top 1% at some point in their lives) during the age range from 60 to 90. The figure also shows that the ceiling DD is captured well by the "top 0.3% sample" during the same age range. Therefore, we show calibrations for each of these experiments and samples over a 30-year lifespan, where age 61 corresponds to period $t = 0$ and age 90 corresponds to period $t = T$. Wealth at death, W_{T+1}, is thus defined as wealth at the end of the 90th year.

To fit the empirical lifecycle profile of wealth in the baseline with $\tau = 0.022$, we set initial wealth W_0 equal to observed wealth at age 60 and calibrate the bequest parameter A so that end-of-life wealth W_{T+1} equals observed wealth at age 90. The calibration of A uses the optimal bequest condition (9) under $\tau = 0.022$. Thus, our calibrations ensure that the baseline wealth path matches the observed start and end points over the considered age range. To capture the shape of the wealth profile between periods 0 and T, we calibrate the discount factor δ given a reasonable value of the gross rate of return R. The calibrated values are $R = 1.07$ and $\delta \approx 0.96$.[36] As we shall see, our parsimonious model is able to fit the empirical lifecycle profile very well.

We calibrate the EIS σ and the bequest elasticity α to match the quasi-experimental moments from the analysis of the 1989-reform. The structural parameters must be matched to the treatment-

---

[36] As we do not explicitly account for income taxes, R should be interpreted as the gross rate of return after the taxation of capital income. Even so, $R = 1.07$ falls well within the range of estimates returns at the top of the wealth distribution (see Fagereng *et al.* 2016).

on-the-treated (TOT) effects, and we exploit the full dynamic pattern of those effects. We have estimates for each year between 1989-96, giving us eight moments to identify two parameters. The reason why the dynamic pattern of the estimates is informative is that the two elasticities σ, α have different implications for the time path of behavioral responses: the EIS channel is relatively strong in the short run, while the bequest elasticity channel is relatively strong in the long run. Therefore, the two elasticities determine, not just the overall magnitude of the effects, but also the concavity or convexity of the effects over time. When σ is larger (smaller) relative to α, the time path of effects is more concave (convex).

Specifically, we calibrate σ and α by minimizing a standard quadratic loss function, i.e.

$$\mathcal{L}(\sigma, \alpha) = \sum_{t=1}^{8} \left[\text{TOT}_t - \Delta \text{log} W_t(\sigma, \alpha) \right]^2, \tag{17}$$

where TOT_t is the estimated treatment effect in year t and $\Delta \text{log} W_t(\sigma, \alpha)$ is the model-predicted effect in year t under the structural primitives σ, α. Recall that we have TOT estimates of both the *total* effect and the *behavioral* effect, where the latter excludes the mechanical effect. We will match on the total effect, letting the mechanical and behavioral effects be a "free variables". In general, our model does not exactly match the mechanical effect in the data, in part because our representative agent approach does not explicitly model the different tax parameters for singles and couples. We consider a unique threshold $\bar{W}$ and a unique tax rate τ, translating the tax reform (part of which changed the exemption threshold for couples relative to singles) into an implied average change in the tax rate $\Delta \tau$. To put it differently, while our empirical estimates are based on experiments that change tax rates partly through household-specific threshold changes, our calibration exercises translate this into a simpler experiment that changes the tax rate on a representative individual.

Although we could estimate σ and α based solely on equation (17), we bring in an additional empirical moment to disciplin the calibration. As we saw in the previous section, the age profile of wealth (among the wealthy) tend to be flat towards the end of life. Specifically, we showed that the wealth profile is roughly flat after the age of 80, and that this is true both before and after the wealth tax reform. Calibrations that imply strongly increasing or decreasing wealth during the last years of life do not seem reasonable in light of these facts. As a result, we require that the average wealth growth during the last 10 years of life remain the same after the reform as before the reform. The estimation of σ and α is therefore based on minimizing equation (17) subject to this requirement on the wealth path during the last years of life.

To summarize, our calibration procedure consists of two interconnected steps. In the first step we calibrate one set of parameters (such as the weight on bequests A and the discount factor δ) to match the wealth profile in the baseline with $\tau = 0.022$, taking the structural elasticities σ and α as given. In the second step we calibrate σ and α to match the quasi-experimental estimates of the effects of tax reform, taking the parameters from the first step as given. We loop back and forth between these two steps until we converge to a fixed point where the elasticities found in the second step correspond to those used as input in the first step.

5.3 Simulating the Long-Run Effects of Wealth Taxation

We first consider the effects of wealth taxation on the moderately wealthy (couples DD), i.e. on taxpayers between the 98th and 99th percentile cutoffs of the wealth distribution. As discussed earlier, this population features only small avoidance responses to wealth taxation, which makes our model of real wealth accumulation particularly appropriate. Recall that the absence of strong avoidance responses in this sample is suggested by two empirical facts: (i) the finding in this paper of very small bunching at the exemption threshold, and (ii) the finding by Alstadsæter *et al.* (2017a) that hidden wealth is substantial only within the top 1% (and mostly within the top 0.1%).

Our findings are shown in Figure 13. Panel A shows three wealth paths: the empirical wealth path (dotted black line), the simulated wealth path before the reform (solid black line), and the simulated wealth path after the reform (solid red line). The simulated wealth path before the reform is calibrated to fit the empirical wealth path. The simulated wealth path after the reform is based on a reduction of the wealth tax rate by 1.84 percentage points, corresponding to the differential tax cut between the treatment and control groups in the couples DD. Our calibration ensures that the differences between the before-reform and after-reform wealth paths respect the quasi-experimental estimates during the first 8 years. This can be seen in Panel B, which compares the simulated effects on log wealth (solid orange line) to the estimated effects (dotted blue line). This panel also splits the total effect into the underlying mechanical and behavioral effects (dashed grey lines).

The following insights are worth highlighting. First, the effect of the wealth tax reduction on the stock of wealth grows for about 20 years and then starts to stabilize. At the end of life, wealth is about 20% higher than it would have been absent the reform. Second, while this effect is sizeable, the underlying tax incentive driving it is very large. Writing the net-of-tax rate of return as $(1 - \tau) R - 1$, the reform-induced change in the return is equal to $-\Delta \tau \cdot R / ((1 - \tau) R - 1) =$

32

42%. Therefore, the long-run elasticity of wealth with respect to the net-of-tax rate of return is about 0.5. Third, another way of gauging the size of the effect is by considering the "structural primitives" implied by the calibration. These are summarized in Table 2, which shows that the EIS is 0.75 and the bequest elasticity is 0.51. The bequest elasticity weighs heavily in the long-run effect, because the weight on the bequest motive is large ($A = 13.53$).[37]

Finally, the figure illustrates that the mechanical effect becomes increasingly important over time. In fact, in this calibration, the positive effect on end-of-life wealth is driven entirely by the mechanical effect, while the behavioral effect is negative. To understand this, Figure A.XII in the appendix shows the consumption paths before and after the reform. Because of the substitution effects of tax cuts (governed by σ and α), households shift consumption from earlier periods to later periods. At the same time, from the wealth effect, households want to increase consumption in all periods. The two effects together imply that consumption is reduced (savings are increased) only for a short period of time. Therefore, when wealth taxes are reduced, the households save less during most of their lives.[38] The compounded impact of these savings reductions eventually turn the behavioral effect negative. Even with the negative behavioral effects, the strong mechanical effects prevent wealth from falling and allow households to die with considerably larger wealth than they would have in the presence of the wealth tax.

We now turn to the very wealthiest taxpayers (ceiling DD), namely those within the top 1% of the wealth distribution. The very wealthy featured stronger responses than the moderately wealthy in the quasi-experimental analysis, and it is likely that some or all of this additional responsiveness represents avoidance or evasion. While our theoretical framework does not explicitly characterize the avoidance decision, the model may still be useful for extrapolating the estimated short-run responses forward in time. But it does mean that the structural parameters — for example the EIS — will pick up evasion responses and therefore cannot be interpreted in the usual way.

The results are presented in Figure 14, which is constructed in the same way as the preceding

---

[37]Figure A.XIII in the appendix illustrates how the calibration of α and σ works. The two parameters are set to ensure that the model fits both the quasi-experimental moments during years 1,...,8 and the observed flatness of the wealth profile at the end of life (before and after the reform). While different combinations of the two parameters can provide a reasonable fit of the short-term effects (a lower α can be compensated for by a higher σ), the additional requirement on the end-of-life profile nails both parameters. As discussed in the previous section, the concavity or convexity of the estimated effects over time may in itself be informative of the relative magnitudes of α and σ. While this is true in general, our calibration is not very robust when using only those moments, because the time path of the estimates happens to be almost linear. If we had observed stronger concavity or convexity, the composition of the effect on α and σ would have been better identified.

[38]Or rather, because we are looking at a sample of older people who dissave, the effect of the tax cut is make them *dissave more* during most of their remaining life.

figure. Here the reform experiment is a reduction in the wealth tax rate by 1.56 percentage points, corresponding to the average tax cut in the treatment group of unbound taxpayers. As before, the model does a good job of fitting the baseline wealth profile to the empirical profile (Panel A) and the wealth responses to the quasi-experimental moments (Panel B). The effect of cutting wealth taxes on long-run wealth is equal to about 70%. Given that the percentage change in the net-of-tax rate of return is equal to $-\Delta\tau \cdot R/\left((1-\tau)R - 1\right) = 36\%$, the implied long-run elasticity of wealth is about 2. Slightly more than half of this effect is mechanical, while the rest is behavioral. The mechanical effects are stronger in this experiment than in the previous one (because the population is wealthier), but because the total effects on which we are matching are also stronger, the behavioral effects remain positive throughout. The calibrated "structural elasticities" σ, α are very large in this case (as shown in Table 2), but as we have discussed they are partly picking up avoidance responses.

6   Conclusion

In this paper we address one of the most important unanswered questions in public finance: What is the effect of capital taxation on capital supply? The answer to this question is critical for assessing the desirability of taxing capital income or wealth. There is an existing empirical literature studying different aspects of capital taxation — including a handful of papers on wealth taxation and many more on wealth *transfer* taxation — but there is no consensus on what might be a reasonable range for the elasticity of capital supply, particularly in the long run. Saez & Stantcheva (2017) show that this elasticity parameter provides a sufficient statistic for optimal capital taxation, but they do not cite any empirical evidence on its value. As discussed in the beginning, the lack of evidence in this area can be explained by a number of methodological difficulties. There are major empirical challenges related to both measurement and identification, as well as conceptual challenges related to the modeling of savings motives and the dynamic nature of wealth accumulation. Through a combination of quasi-experimental analysis based on administrative wealth records, theoretical modeling and calibration, we have tried to make progress on this question.

For the bigger picture, it is worth discussing what our estimates may be missing. We highlight three potential limitations. First, our estimates capture the effect of wealth taxes on those who are already wealthy, as opposed to the forward-looking effect on those who *aspire* to become wealthy. The aspiration effect of wealth taxes, *even if it is quantitatively important*, is extremely difficult to

identify empirically. One would have to compare the savings behavior of the *potentially* wealthy across economies that vary permanently in their levels of capital taxation, but such cross-country analyses are typically not persuasive.

Second, our estimates capture the effect of wealth taxes conditional on staying in Denmark. In other words, we do not consider the potential migration response to wealth taxes at the top. While there some evidence on migration responses to *labor* income taxes at the top (see Kleven *et al.* 2011a, 2014; Akcigit *et al.* 2016), there is virtually no evidence on migration responses to capital or wealth taxes.[39] Such responses are difficult to study due to lack of statistical power: we are studying the extreme tail of the wealth distribution, and given the low frequencies of international migration, very few individuals are moving country from year to year. Furthermore, moving to another country is arguably not the most natural response to wealth taxes. Because capital tends to be more mobile than people, it would be more natural to move wealth across borders (which would be picked up by our taxable wealth estimates) than to move the household across borders.

Third, our quasi-experimental approach captures the effect of wealth taxes in partial equilibrium. The changes in wealth accumulation that we find may have implications for asset prices and wage rates, making the general equilibrium effect different from the partial equilibrium effect. While this is important to keep in mind, it is a limitation of any quasi-experimental study (i.e., of any well-identified study). Our partial equilibrium estimates provide a set of moments that economists can target when calibrating general equilibrium models.

Finally, while we have motivated our contribution by its importance for assessing the desirability of capital taxation, we do not undertake an optimal tax analysis in this paper. Given the machinery required, this is beyond the scope of this paper. We hope that future work on optimal capital taxation will build on the empirical work presented here.

---

[39]An exception is the paper by Martínez (2017), who provides evidence on migration responses to reductions in income and wealth taxes in the Swiss canton of Obwalden. These estimates capture migration across tax jurisdictions within a country, as opposed to migration across international borders. On the latter, Zucman (2008) considered migration responses in response to a French wealth tax, but found little evidence of any effect.

References

AKCIGIT, UFUK, BASLANDZE, SALOMÉ, & STANTCHEVA, STEFANIE. 2016. Taxation and the International Mobility of Inventors. *American Economic Review*, **106**(10), 2930–2981. 35

ALSTADSÆTER, ANNETTE, JOHANNESEN, NIELS, & ZUCMAN, GABRIEL. 2017a. Tax Evasion and Inequality. NBER Working Paper No. 23772. 3, 7, 32

ALSTADSÆTER, ANNETTE, JOHANNESEN, NIELS, & ZUCMAN, GABRIEL. 2017b. Who Owns the Wealth in Tax Havens? Macro Evidence and Implications for Global Inequality. NBER Working Paper No. 23805. 7

ALVAREDO, FACUNDO, CHANCEL, LUCAS, PIKETTY, THOMAS, & SAEZ, EMMANUEL. 2017. *The World Wealth and Income Database (WID.world)*. Website. 7

ANDREONI, JAMES. 1989. Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy*, **97**(6), 1447–1458. 22

ANDREONI, JAMES. 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *Economic Journal*, **100**(401), 464–477. 22

ATKINSON, ANTHONY B., & STIGLITZ, JOSEPH E. 1976. The Design of Tax Structure: Direct Versus Indirect Taxation. *Journal of Public Economics*, **6**(1-2), 55–75. 4

BENHABIB, JESS, & BISIN, ALBERTO. 2017. Skewed Wealth Distributions: Theory and Empirics. *Journal of Economic Literature*, forthcoming. 5

BOSERUP, SIMON H., KOPCZUK, WOJCIECH, & KREINER, CLAUS T. 2014. Stability and Persistence of Intergenerational Wealth Formation: Evidence from Danish Wealth Records of Three Generations. Working Paper. 6

BRÜLHART, MARIUS, GRUBER, JONATHAN, KRAPF, MATTHIAS, & SCHMIDHEINY, KURT. 2016. Taxing Wealth: Evidence from Switzerland. NBER Working Paper No. 22376. 4

CARROLL, CHRISTOPHER D. 2002. Why Do the Rich Save So Much? *In:* SLEMROD, J. B. (ed), *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*. Harvard University Press: Cambridge, MA. 5, 21

CHAMLEY, CHRISTOPHE. 1986. Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. *Econometrica*, **54**(3), 607–622. 4

CHETTY, RAJ, FRIEDMAN, JOHN N., OLSEN, TORE, & PISTAFERRI, LUIGI. 2011. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *Quarterly Journal of Economics*, **126**(2), 749–804. 13

CHETTY, RAJ, FRIEDMAN, JOHN N., LETH-PETERSEN, SØREN, NIELSEN, TORBEN H., & OLSEN, TORE. 2014. Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark. *Quarterly Journal of Economics*, **129**(3), 1141–1219. 6, 9

DE NARDI, MARIACRISTINA. 2004. Wealth Inequality and Intergenerational Links. *Review of Economic Studies*, **71**(3), 743–768. 5, 21

DINARDO, JOHN, FORTIN, NICOLE M., & LEMIEUX, THOMAS. 1996. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, **64**(5), 1001–1044. 13

FAGERENG, ANDREAS, GUISO, LUIGI, MALACRINO, DAVIDE, & PISTAFERRI, LUIGI. 2016. Heterogeneity and Persistence in Returns to Wealth. NBER Working Paper No. 22822. 16, 30

FARHI, EMMANUEL, & WERNING, IVÁN. 2010. Progressive Estate Taxation. *Quarterly Journal of Economics*, **125**(2), 635–673. 22

JUDD, KENNETH L. 1985. Redistributive Taxation in a Simple Perfect Foresight Model. *Journal of Public Economics*, **28**(1), 59–83. 4

KLEVEN, HENRIK. 2014. How Can Scandinavians Tax So Much? *Journal of Economic Perspectives*, **28**(4), 77–98. 21

KLEVEN, HENRIK J. 2016. Bunching. *Annual Review of Economics*, **8**(1), 435–464. 4, 10, 13

KLEVEN, HENRIK J., & SCHULTZ, ESBEN A. 2014. Estimating Taxable Income Responses Using Danish Tax Reforms. *American Economic Journal: Economic Policy*, **6**(4), 271–301. 21

KLEVEN, HENRIK J., & WASEEM, MAZHAR. 2013. Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *Quarterly Journal of Economics*, **128**(2), 669–723. 13

KLEVEN, HENRIK J., LANDAIS, CAMILLE, & SAEZ, EMMANUEL. 2011a. Taxation and International Migration of Superstars: Evidence from the European Football Market. *American Economic Review*, **103**(5), 1892–1924. 35

KLEVEN, HENRIK J., KNUDSEN, MARTIN, KREINER, CLAUS T., PEDERSEN, SØREN, & SAEZ, EMMANUEL. 2011b. Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, **79**(3), 651–692. 14

KLEVEN, HENRIK J., LANDAIS, CAMILLE, SAEZ, EMMANUEL, & SCHULTZ, ESBEN A. 2014. Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark. *Quarterly Journal of Economics*, **129**(1), 333–378. 35

KOPCZUK, WOJCIECH. 2009. Economics of Estate Taxation: A Brief Review of Theory and Evidence. *Tax Law Review*, **63**(1), 139–157. 5, 25

KOPCZUK, WOJCIECH. 2013a. Incentive Effects of Inheritances and Optimal Estate Taxation. *American Economic Review, Papers and Proceedings*, **103**(3), 472–477. 5, 22

KOPCZUK, WOJCIECH. 2013b. Taxation of Intergenerational Transfers and Wealth. *In:* AUERBACH, A. J., CHETTY, R., FELDSTEIN, M. S., & SAEZ, E. (eds), *Handbook of Public Economics*, vol. 5. Elsevier: Amsterdam, NL. 5, 25

KREINER, CLAUS T., LASSEN, DAVID D., & LETH-PETERSEN, SØREN. 2015. Measuring the Accuracy of Survey Responses Using Administrative Register Data: Evidence from Denmark. *In:* CARROLL, C. D., CROSSLEY, T. F., & SABELHAUS, J. (eds), *Improving the Measurement of Household Consumption Expenditures*. University of Chicago Press: Chicago, IL. 6

LETH-PETERSEN, SØREN. 2010. Intertemporal Consumption and Credit Constraints: Does Total Expenditure Respond to an Exogenous Shock to Credit? *American Economic Review*, **100**(3), 1080–1103. 6

MARTÍNEZ, ISABEL. 2017. Beggar-Thy-Neighbour Tax Cuts: Mobility After a Local Income and Wealth Tax Reform in Switzerland. LISER Working Paper Series 2017-08. 35

OECD. 1988. *Taxation of Net Wealth, Capital Transfers and Capital Gains of Individuals*. Organization for Economic Cooperation and Development (OECD): Paris, France. 1

PIKETTY, THOMAS. 2014. *Capital in the Twenty First Century*. Harvard University Press: Cambridge, MA. 1

PIKETTY, THOMAS, & SAEZ, EMMANUEL. 2013. A Theory of Optimal Inheritance Taxation. *Econometrica*, **81**(5), 1851–1886. 5, 22

SAEZ, EMMANUEL. 2010. Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, **2**(3), 180–212. 13

SAEZ, EMMANUEL, & STANTCHEVA, STEFANIE. 2017. A Simpler Theory of Optimal Capital Taxation. *Journal of Public Economics*, forthcoming. 4, 21, 34

SAEZ, EMMANUEL, & ZUCMAN, GABRIEL. 2016. Wealth Inequality in the United States Since 1913: Evidence from Capitalized Income Tax Data. *The Quarterly Journal Of Economics*, **131**(2), 519–578. 1, 7, 8

SAEZ, EMMANUEL, SLEMROD, JOEL B., & GIERTZ, SETH H. 2012. The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature*, **50**(1), 3–50. 11

SEIM, DAVID. 2017. Behavioral Responses to Wealth Taxes: Evidence from Sweden. *American Economic Journal: Economic Policy*, **9**(4), 395–421. 4, 10, 14

SLEMROD, JOEL B., & YITZHAKI, SHLOMO. 2002. Tax Avoidance, Evasion, and Administration. *In:* AUERBACH, A. J., & FELDSTEIN, M. S. (eds), *Handbook of Public Economics*, vol. 3. Elsevier: Amsterdam, NL. 14

ZOUTMAN, FLORIS T. 2015. The Effect of Capital Taxation on Household Savings. Working Paper. 4

ZUCMAN, GABRIEL. 2008. *Les hauts patrimoines fuient-ils l'ISF? Une estimation sur la période 1995-2006*. Master's Thesis, Paris School of Economics. 35

ZUCMAN, GABRIEL. 2013. The Missing Wealth of Nations: Are Europe and the U.S. Net Debtors or Net Creditors? *Quarterly Journal of Economics*, **128**(3), 1321–1364. 7

ZUCMAN, GABRIEL. 2018. Measuring Wealth in a Globalized World. *Annual Review of Economics*, in preparation. 7

FIGURE 1: DISTRIBUTION OF WEALTH IN DENMARK, 1980-2012
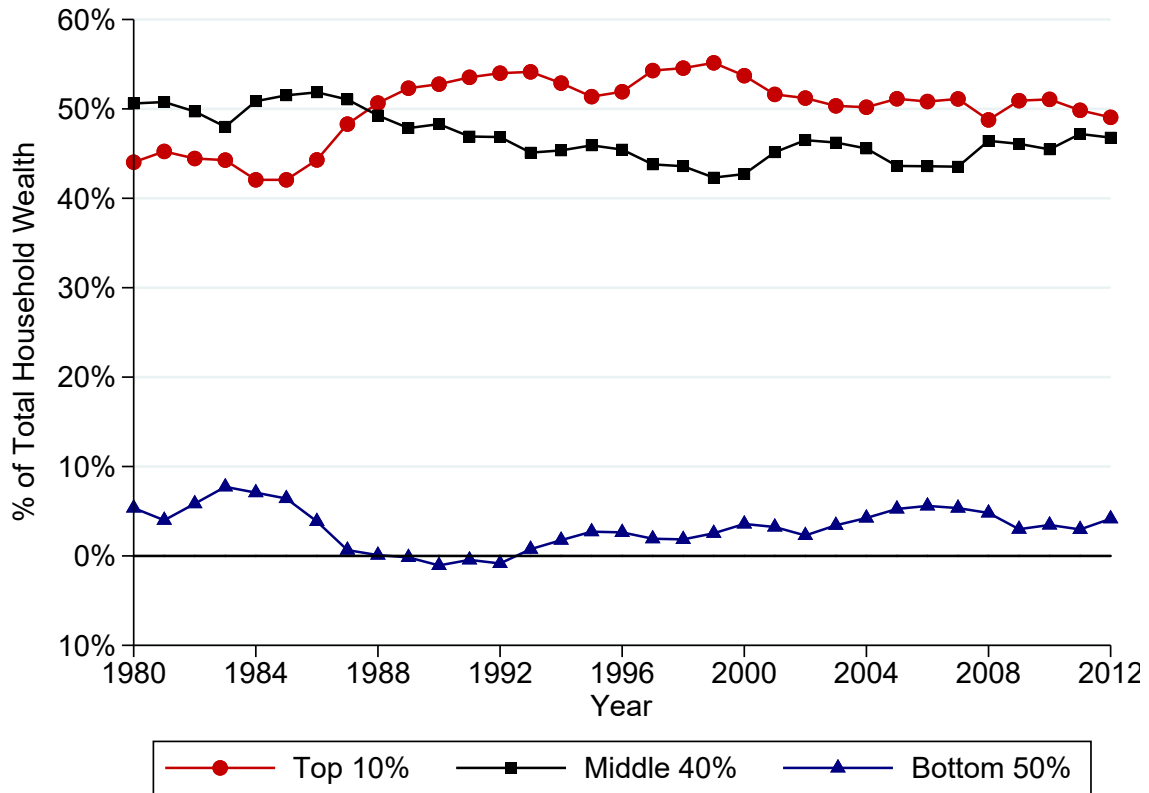
Notes: This figure shows the share of total household wealth in Denmark owned by the bottom 50% of the distribution, the middle 40% (adults between the median and the 90th percentile), and the top 10%. The unit of observation is the adult individual (aged 20 or above), splitting household wealth in married couples equally among the spouses. Wealth includes all financial and non-financial assets, net of any debts. It matches the total amount of household wealth recorded in Denmark's household balance sheet.

FIGURE 2: TOP 1% AND TOP 0.1% WEALTH SHARES IN DENMARK VS THE UNITED STATES

A: Top 1% Wealth Share

B: Top 0.1% Wealth Share
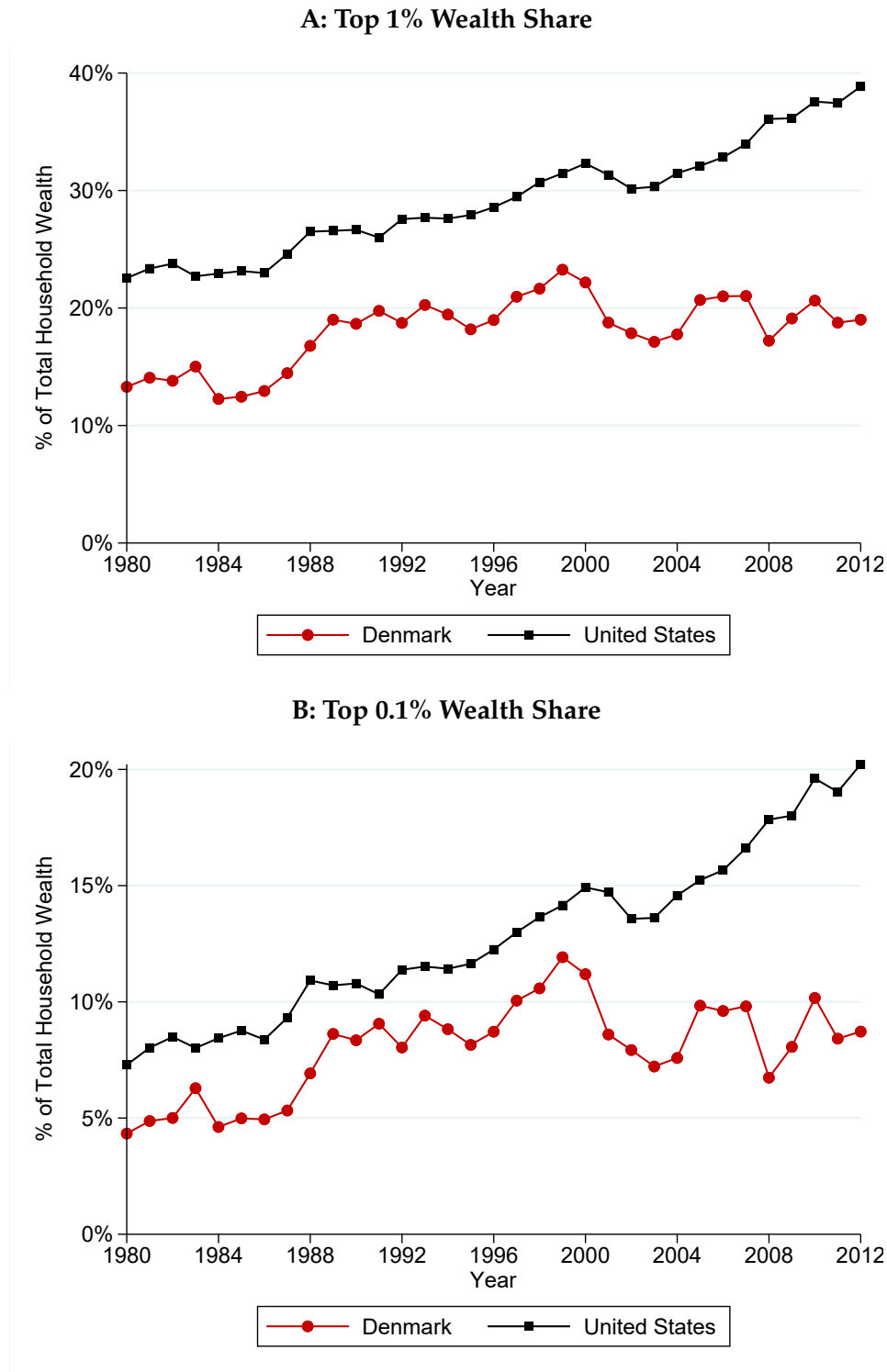
Notes: This figure shows the share of total household wealth owned by the top 1% (Panel A) and the top 0.1% (Panel B) in Denmark vs United States. In both countries, the unit of observation is the adult individual (aged 20 or above), splitting household wealth in married couples equally among the spouses. Wealth includes all financial and non-financial assets, net of any debts, and it adds up to the total amount of household wealth recorded in Denmark's and the United States' household balance sheets.

A: Marginal Tax Rate

B: Exemption Threshold
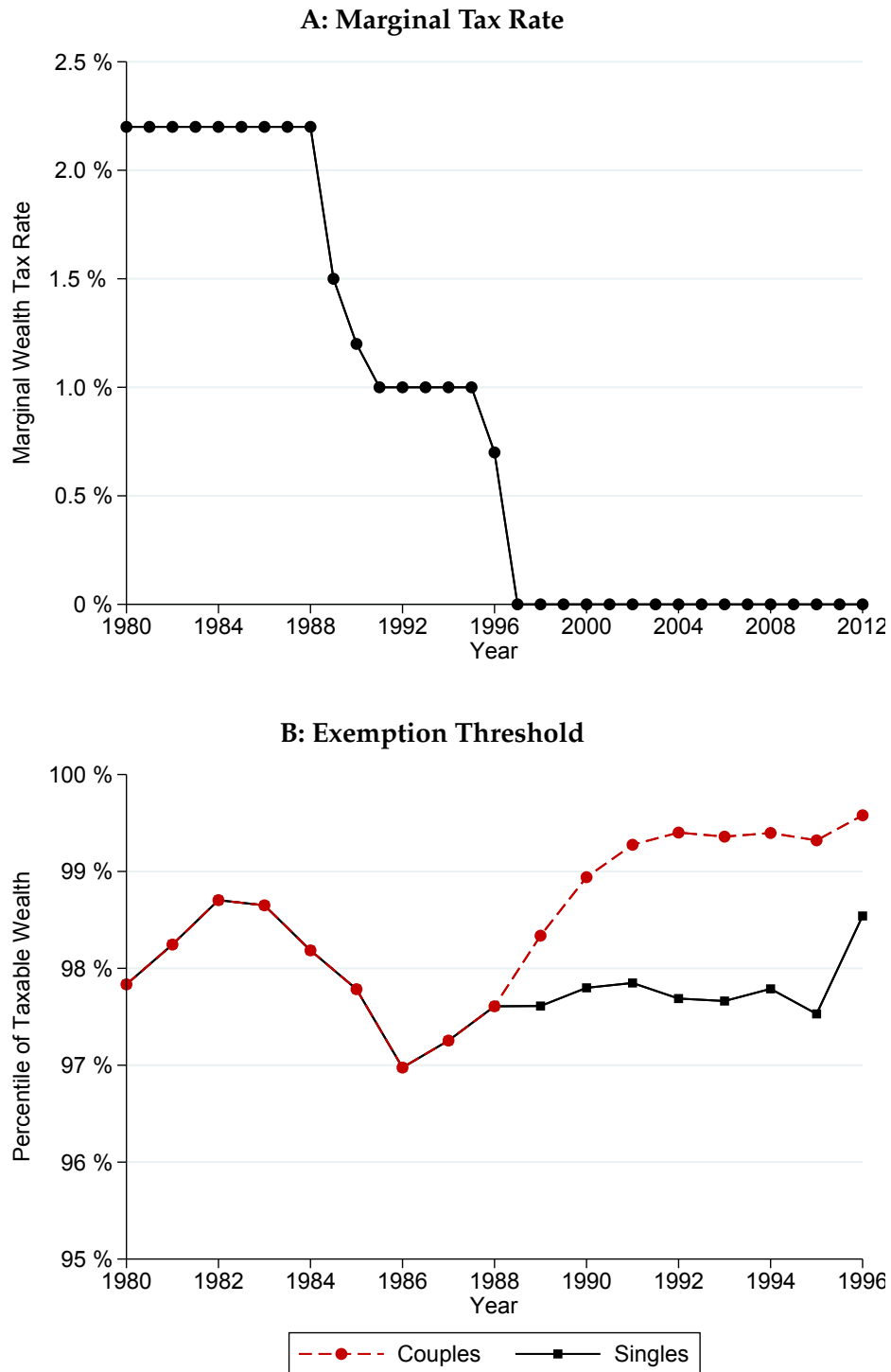
Notes: This figure shows the evolution of the marginal tax rate (Panel A) and the exemption threshold (Panel B) in the Danish wealth tax.

FIGURE 4: BUNCHING AT THE KINK

A: Observed Distribution
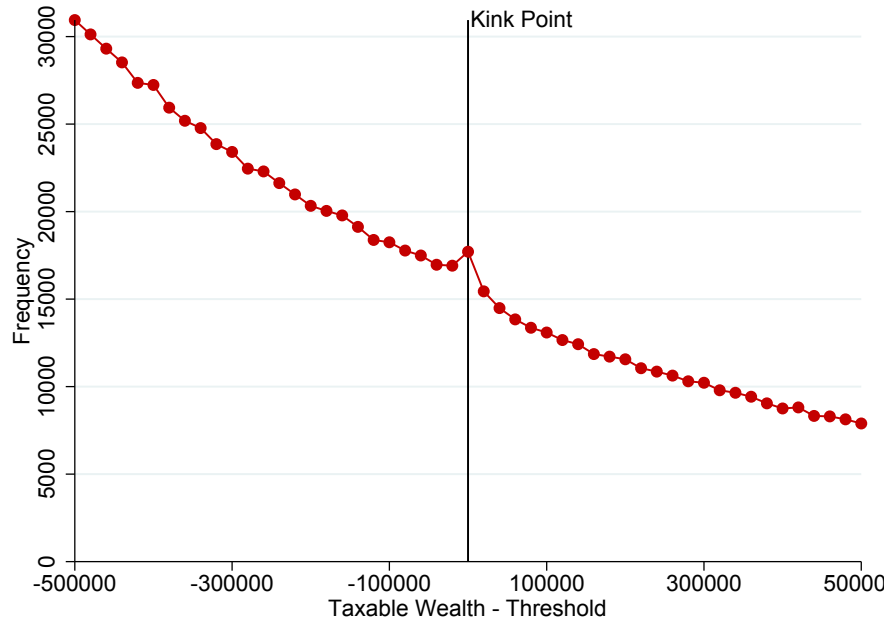
B: Observed vs Counterfactual Distribution
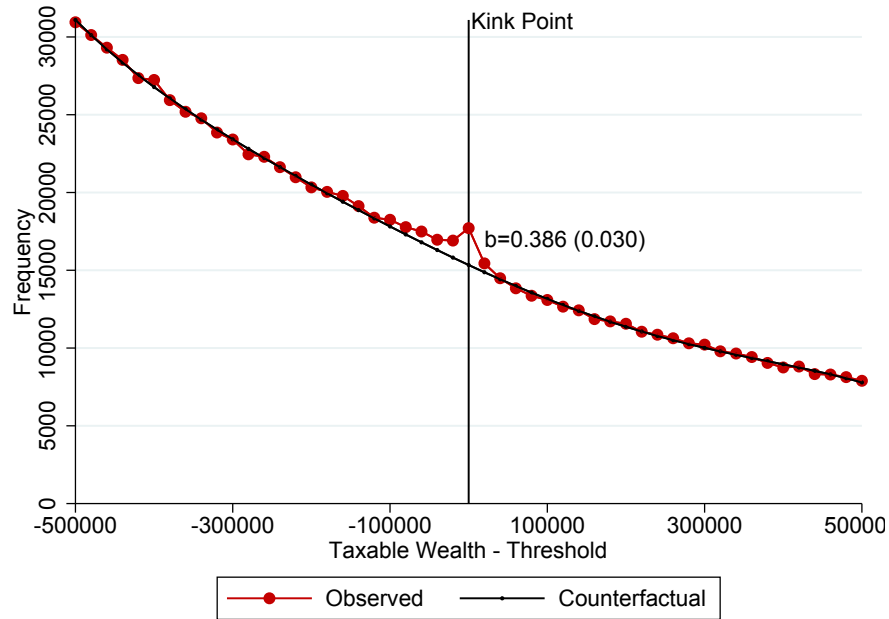
Notes: This figure shows the distribution of taxable wealth in bins of 20,000 DKK (about 20,000/6 USD Dollars) around the exemption threshold (kink point) of the Danish wealth tax, pooling the years 1980-1996. Taxable wealth is measured in terms of its distance to the threshold in each year (putting the kink point at zero), and we adjust for inflation to ensure comparability across years. Panel A shows the observed distribution, while Panel B adds an estimate of the counterfactual distribution absent the kink. The counterfactual is obtained by fitting a 5th-order polynomial to the observed distribution, excluding 6 bins around the kink. The bunching estimate b equals excess mass around the kink, scaled by the height of the counterfactual distribution at the kink.

A: Time Series of Bound and Unbound Households

B: Difference Between Bound and Unbound Households
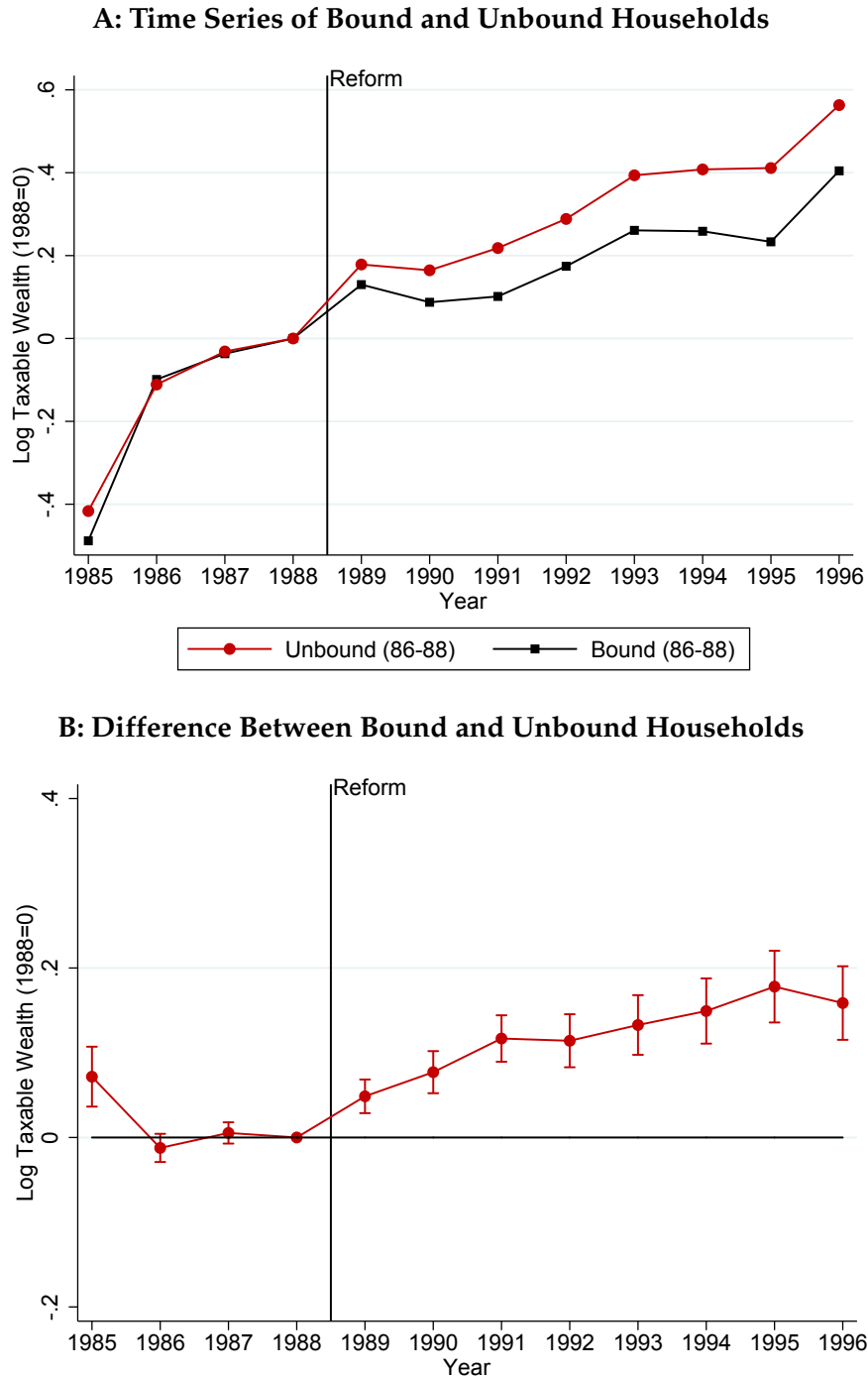
Notes: The figure shows the (intention-to-treat) effects of wealth tax reform on taxable wealth based on the "ceiling DD" approach. The estimation sample is a balanced panel of households observed in all years 1985-1996 and located in the top 1% of the wealth distribution before the reform. Panel A shows the evolution of taxable wealth in the unbound group (tax rate cuts) and the bound group (no tax rate cuts), normalized to zero in the pre-reform year 1988, after absorbing the impact of pre-reform differences in portfolio composition and income using specification (1). Panel B shows the differences between these two series, i.e. our difference-in-differences estimates. The 95% confidence intervals are based on robust standard errors clustered at the household level.

FIGURE 6: CEILING DD: TREATMENT EFFECT ON THE TREATED
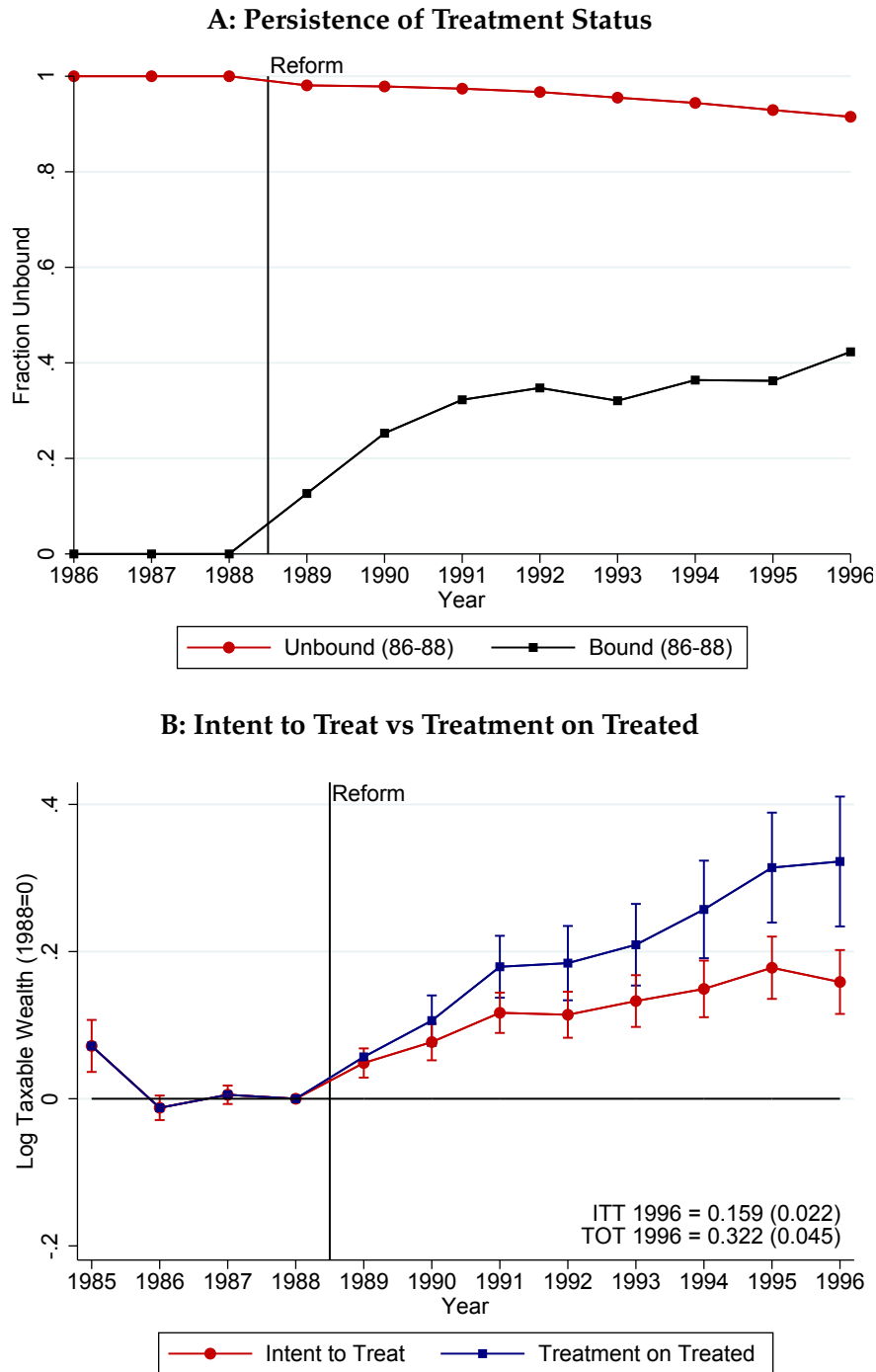
A: Persistence of Treatment Status

B: Intent to Treat vs Treatment on Treated

ITT 1996 = 0.159 (0.022)
TOT 1996 = 0.322 (0.045)

Notes: The figure converts the intention-to-treat (ITT) estimates into treatment-on-the-treated (TOT) estimates. Panel A shows the persistence in treatment status over time. The unbound group switches status either if they become bound or if their wealth falls below the (pre-reform) exemption threshold. The bound group switches status if they become unbound or if their wealth falls below the (pre-reform) exemption threshold. Panel B converts the ITT series into a TOT series by dividing the ITT estimates with the differences in fraction treated from Panel A. The 95% confidence intervals of the TOT estimates are based on the standard errors of the ITT estimates divided by the differences in fraction treated from Panel A.

## FIGURE 7: CEILING DD: BEHAVIORAL VS MECHANICAL EFFECTS
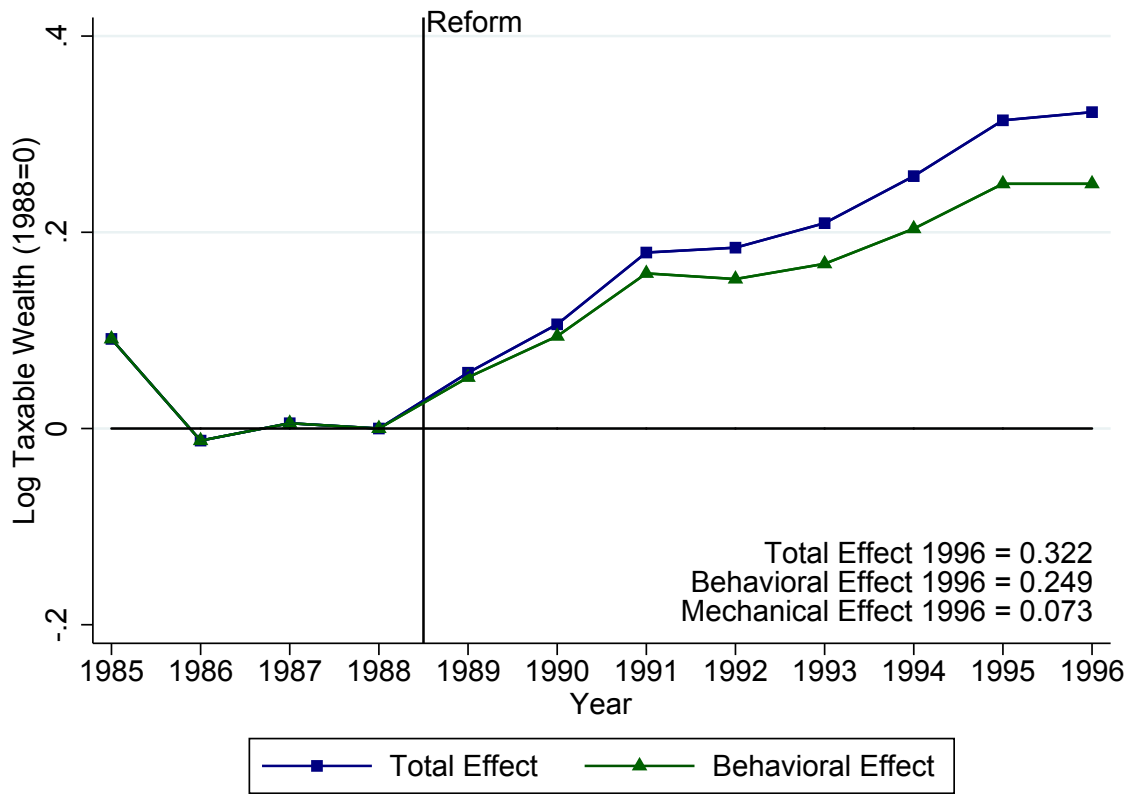
Notes: The figure shows the composition of the total effect of wealth tax reform into mechanical and behavioral effects. The calculation of the mechanical effects is described in Appendix A. The behavioral effect equals the total effect minus the mechanical effect.

A: Time Series of Couples Inside and Outside Range

B: Difference Between Couples Inside and Outside Range

Notes: The figure shows the (intention-to-treat) effects of wealth tax reform on taxable wealth based on the "couples DD" approach. The estimation sample is a balanced panel of married couples observed in all years 1985-1996 and located in the top 5% of the wealth distribution before the reform. Panel A shows the evolution of taxable wealth for couples inside the exempted range and couples outside the exempted range, normalized to zero in the pre-reform year 1988, after absorbing the impact of pre-reform differences in portfolio composition and income using specification (1). Panel B shows the differences between these two series, i.e. our difference-in-differences estimates. The 95% confidence intervals are based on robust standard errors clustered at the household level.

FIGURE 9: COUPLES DD: TREATMENT EFFECT ON THE TREATED
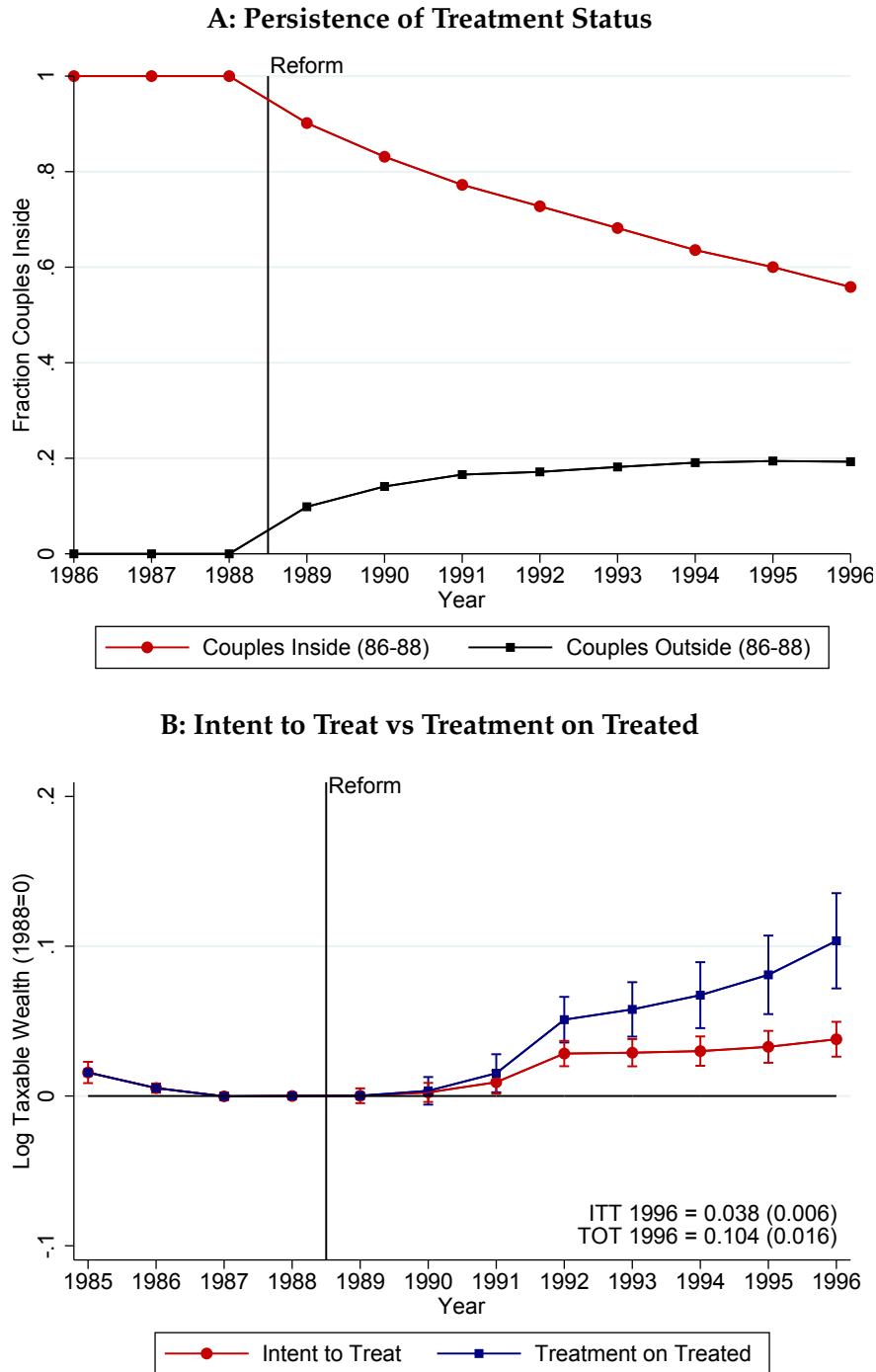
A: Persistence of Treatment Status

B: Intent to Treat vs Treatment on Treated

Notes: The figure converts the intention-to-treat (ITT) estimates into treatment-on-the-treated (TOT) estimates. Panel A shows the persistence in treatment status over time. Couples inside the exempted range switch treatment status if their wealth falls below or rises above the exempted range, and conversely for couples outside the exempted range. Panel B converts the ITT series into a TOT series by dividing the ITT estimates with the differences in fraction treated from Panel A. The 95% confidence intervals of the TOT estimates are based on the standard errors of the ITT estimates divided by the differences in fraction treated from Panel A.

FIGURE 10: COUPLES DD: BEHAVIORAL VS MECHANICAL EFFECTS
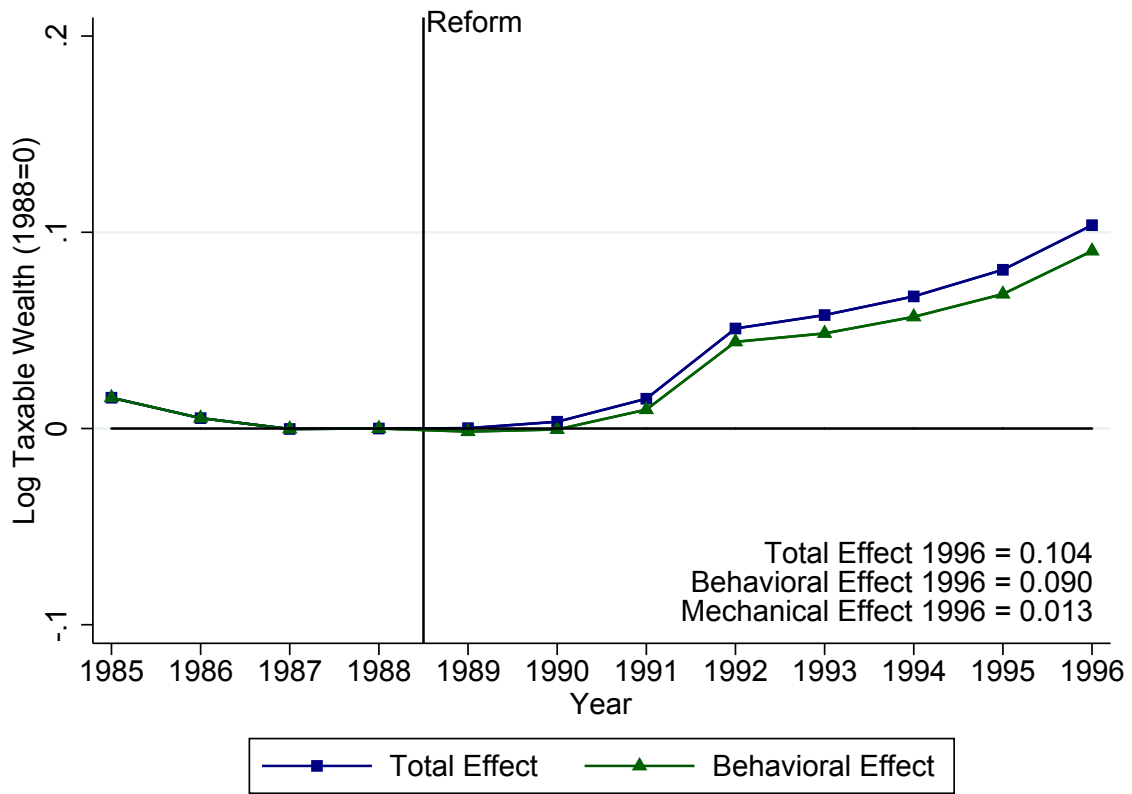
Notes: The figure shows the composition of the total effect of wealth tax reform into mechanical and behavioral effects. The calculation of the mechanical effects is described in Appendix A. The behavioral effect equals the total effect minus the mechanical effect.

FIGURE 11: EMPIRICAL LIFECYCLE PROFILES OF WEALTH
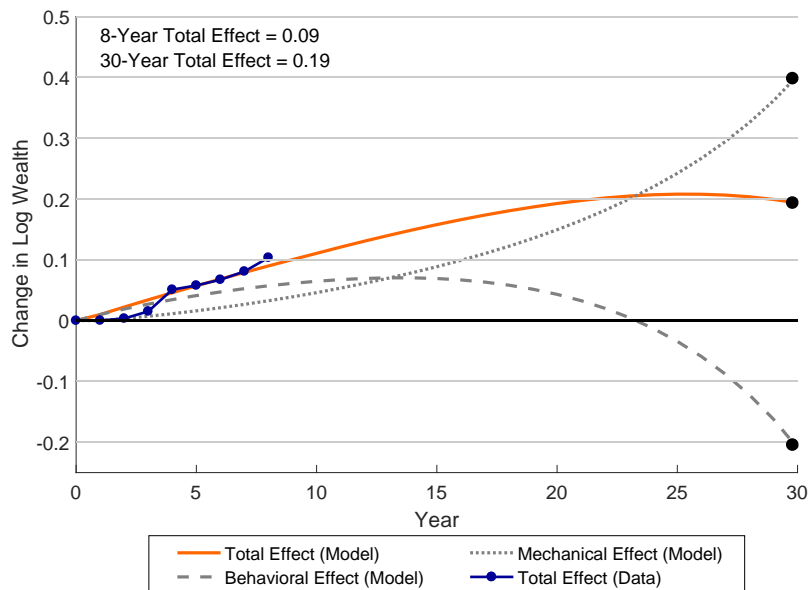
A: Full Population
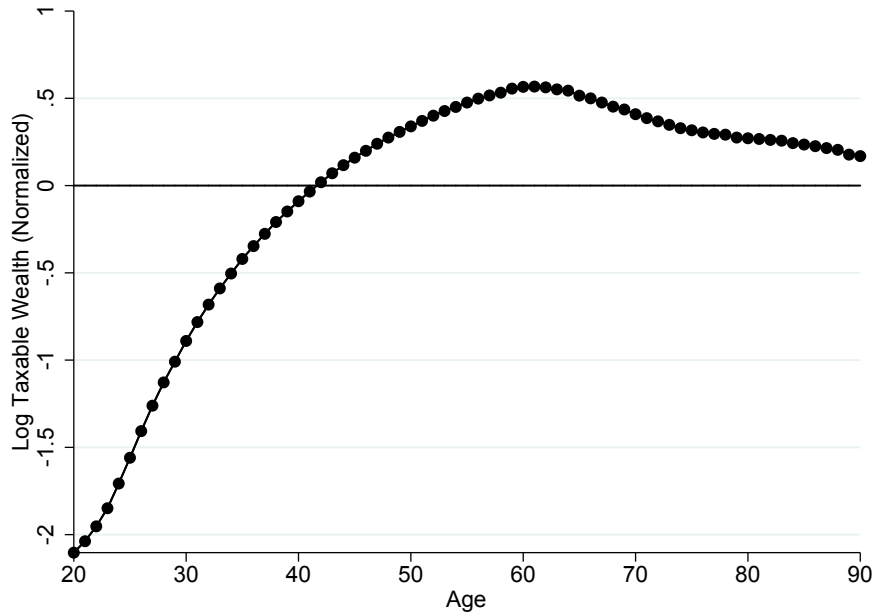
B: Top Percentiles
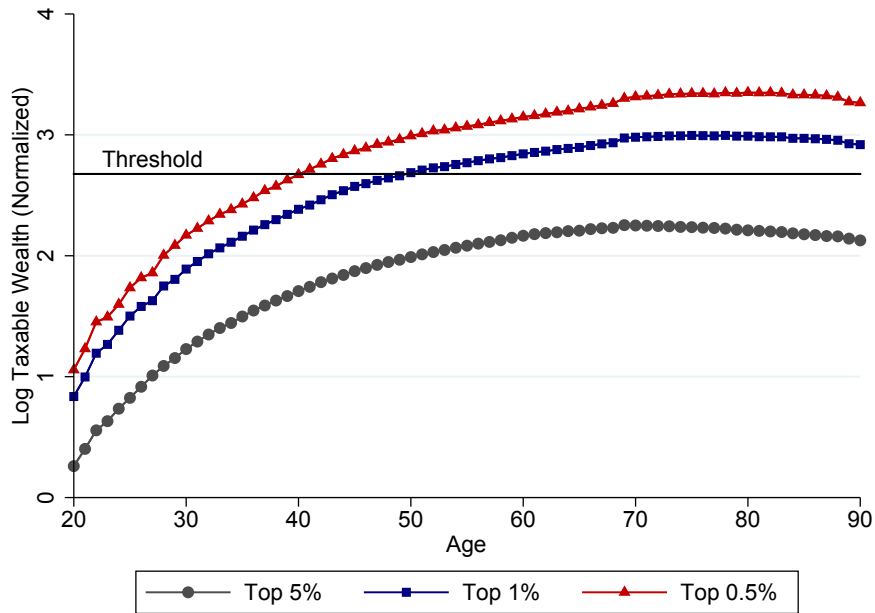
Notes: The figure shows lifecycle profiles of taxable wealth between the ages of 20-90 in an unbalanced panel of individuals over the period 1980-2012. To avoid any confounding effects of inflation as people grow older, we normalize log wealth for each individual in each year by the average wealth in the population in that year. The graphs show averages of this normalized wealth measure in different age bins. Panel A shows the lifecycle profile in the full population, while Panel B shows lifecycle profiles in the top percentiles of the population. The top-percentile samples include individuals who are in the top $p\%$ for at least three years of their observed lifespan, keeping them in the data for their entire observed lifespan.

FIGURE 12: EMPIRICAL LIFECYCLE PROFILES OF WEALTH FOR THE TOP 1% BEFORE AND AFTER THE REFORM
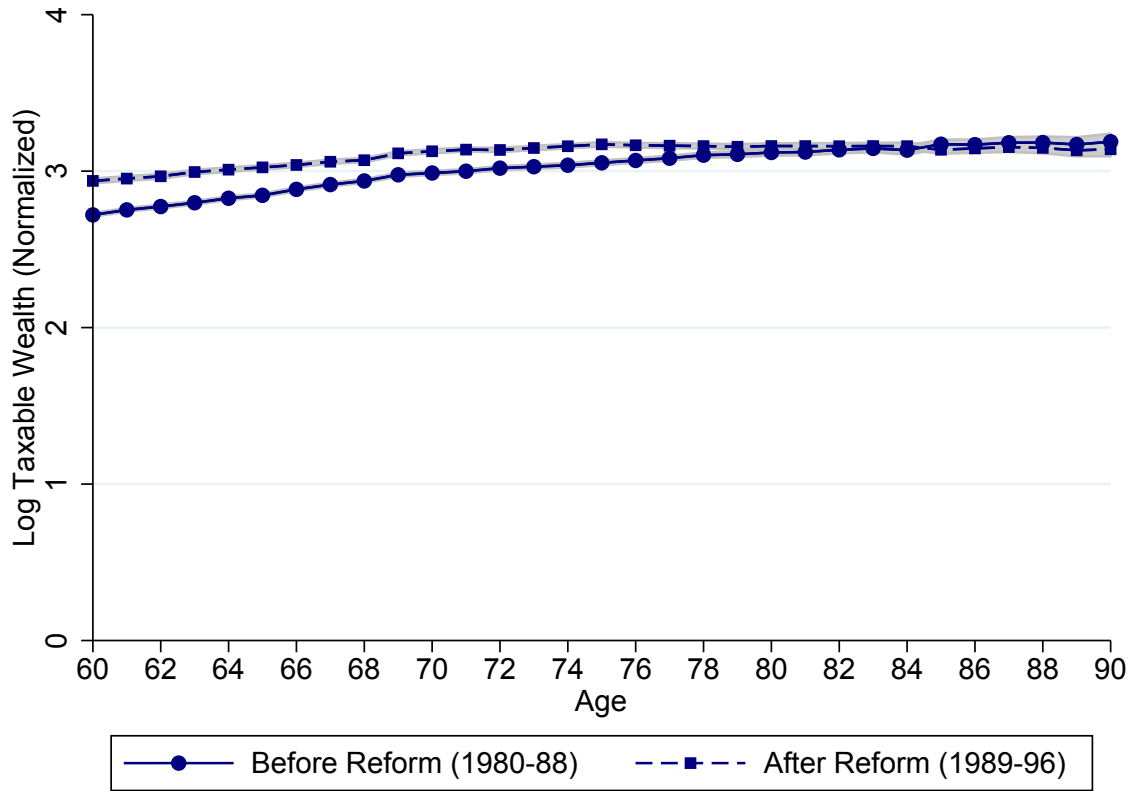
Notes: The figure shows lifecycle profiles of taxable wealth between the ages of 60-90 in an unbalanced panel of individuals before the reform (1980-1988) and after the reform (1989-96), respectively. To avoid any confounding effects of inflation as people grow older, we normalize log wealth for each individual in each year by the average wealth in the population in that year. The graphs show averages of this normalized wealth measure in different age bins. The sample consists of individuals who make it to the top 1% for at least three years of their observed lifespan, keeping them in the data for their entire observed lifespan. The grey-shaded areas represent 95% confidence intervals.

A: Observed and Simulated Wealth Paths

B: Long-Run Effects of Wealth Tax Cut

Notes: The figure shows the long-run effects of wealth tax cuts when calibrating our model to the sample and estimates from the couples DD. These are effects for the moderately wealthy (between the 98th and 99th percentile cutoffs). The reform experiment cuts the wealth tax rate by 1.84 percentage points, corresponding to the differential tax cut between the treatment and control groups. Panel A shows the empirical lifecycle profile of wealth, the simulated lifecycle profile before the reform (calibrated to fit the empirical profile), and the simulated lifecycle profile after the reform. Panel B illustrates the total effects, the mechanical effects, and the behavioral effects on taxable wealth over 30 years, demonstrating that the model matches the quasi-experimental estimates over the initial 8 years.

FIGURE 14: LONG-RUN EFFECTS OF CUTTING WEALTH TAXES: THE VERY WEALTHY (CEILING DD)
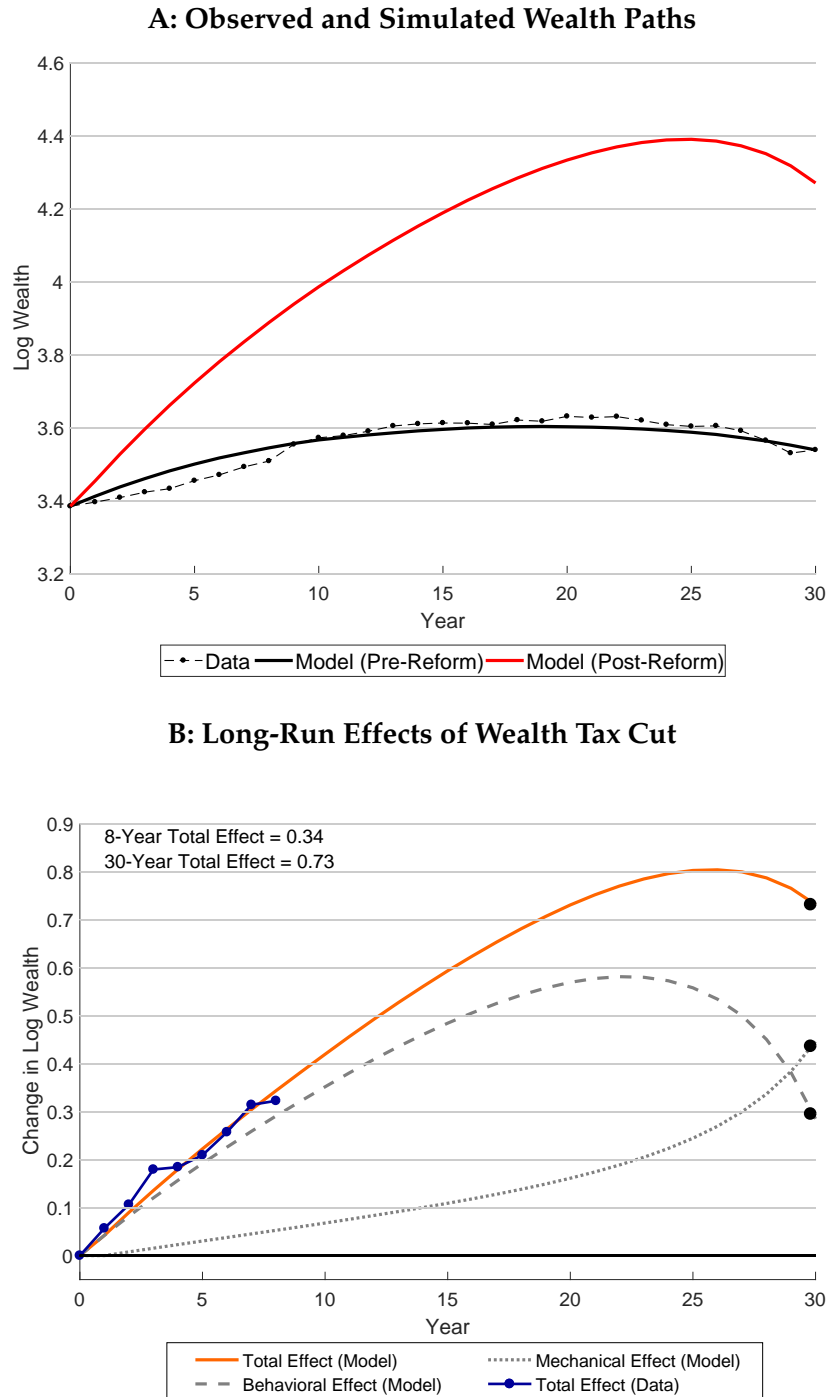
A: Observed and Simulated Wealth Paths

B: Long-Run Effects of Wealth Tax Cut

Notes: The figure shows the long-run effects of wealth tax cuts when calibrating our model to the sample and estimates from the ceiling DD. These are effects for the very wealthy (within the top 1%). The reform experiment cuts the wealth tax rate by 1.56 percentage points, corresponding to the tax cut for the average person in the treatment group. Panel A shows the empirical lifecycle profile of wealth, the simulated lifecycle profile before the reform (calibrated to fit the empirical profile), and the simulated lifecycle profile after the reform. Panel B illustrates the total effects, the mechanical effects, and the behavioral effects on taxable wealth over 30 years, demonstrating that the model matches the quasi-experimental estimates over the initial 8 years.

TABLE 1: DIFFERENCE-IN-DIFFERENCES ESTIMATES OF WEALTH RESPONSES TO WEALTH TAX CUTS

	Ceiling DD (Very Wealthy)				Couples DD (Moderately Wealthy)			
	Unbound vs Bound		Unbound vs Very Bound		Couples vs Couples		Couples vs Singles	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Taxable Wealth								
ITT Effect in 1996	0.093	0.159	0.131	0.191	0.037	0.038	0.109	0.025
	(0.021)	(0.022)	(0.031)	(0.032)	(0.006)	(0.006)	(0.008)	(0.008)
TOT Effect in 1996	0.189	0.322	0.226	0.330	0.100	0.104	0.198	0.045
	(0.043)	(0.045)	(0.053)	(0.055)	(0.016)	(0.016)	(0.014)	(0.014)
Observations	138,384	138,384	126,336	126,336	520,452	520,452	340,428	340,428
Panel B: Third-Party Wealth								
ITT Effect in 1996	0.098	0.165	0.129	0.189	0.030	0.029	0.118	0.031
	(0.022)	(0.023)	(0.032)	(0.034)	(0.006)	(0.006)	(0.008)	(0.008)
TOT Effect in 1996	0.200	0.335	0.224	0.328	0.083	0.079	0.214	0.057
	(0.021)	(0.047)	(0.056)	(0.058)	(0.017)	(0.017)	(0.014)	(0.015)
Observations	138,165	138,165	126,152	126,152	519,939	519,939	340,198	340,198
Household FE	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X
Equity Shares × Year FE		X		X		X		X
Housing Shares × Year FE		X		X		X		X
Income Deciles × Year FE		X		X		X		X

Notes: The table summarizes our estimates of the intent-to-treat (ITT) and treatment-on-the-treated (TOT) effects of the 1989 wealth tax cuts. We show the effects on taxable wealth (Panel A) and on third-party reported wealth (Panel B) obtained from the ceiling DD (very wealthy) and the couples DD (moderately wealthy). We show estimates across different specifications of the control group and of the non-parametric controls. Columns (1), (3), (5), and (7) show plain DD estimates based on specifications with only household and year fixed effects. Columns (2), (4), (6), and (8) show our preferred DD estimates based on specification (1) in which we control for the impact of pre-reform portfolio composition and income. The estimation sample is a balanced panel of households observed in all years 1985-1996. Robust standard errors are clustered at the household level.

TABLE 2: SIMULATION ANALYSIS OF LONG-RUN EFFECTS

	Couples DD (Moderately Wealthy)		Ceiling DD (Very Wealthy)	
	8-Year Effect	30-Year Effect	8-Year Effect	30-Year Effect
	(1)	(2)	(3)	(4)
Panel A: Effects on Wealth				
Total Effect	0.09	0.19	0.34	0.73
Mechanical Effect	0.03	0.40	0.05	0.44
Behavioral Effect	0.06	-0.21	0.29	0.29
Panel B: Calibration				
R	1.07		1.07	
δ	0.956		0.958	
A	13.53		18.17	
σ	0.75		5.07	
α	0.51		3.23	

Notes: The table summarizes the simulation results shown in Figures 13-14 and the underlying calibrations of parameters. We show the 8-year effects on wealth (corresponding to the last year in which we can directly estimate those effects) and the 30-year effects on wealth.

A Mechanical Effects of Wealth Tax Reform

To compute the mechanical effects of the 1989-reform, we distinguish between single and married individuals as they were treated differently by the reform. We index marital status by $m = 0, 1$. As described above, the mechanical tax savings of the treated must be based on a measure of counterfactual wealth. Focusing on representative individuals (single and married, respectively) in the treatment group, we define counterfactual wealth in year t as average actual wealth in 1988 (pre-reform year) adjusted for the wealth growth in the control group between 1988 and year t. That is, the counterfactual wealth for marital status m in year t is given by

$$\widetilde{W}_t^m = \begin{cases} \mathsf{E}\left[W_{it}^m \mid m, t\right] & \text{if } t = 1988 \\ \widetilde{W}_{t-1}^m \times G_t^m & \text{if } t > 1988 \end{cases} \tag{18}$$

where G_t^m denotes the growth rate for marital status m in the control group between year $t - 1$ and t.

When calculating the tax savings, we take into account that the tax rate was gradually reduced from 2.2 % to 1% between 1989-91, while the exemption threshold was gradually increased for couples relative to singles between 1989-1992.[40] We denote by $\Delta \tau_t$ the tax rate reduction between year t and 1988. Furthermore, we denote by $\overline{W}_t^m$ the exemption threshold for marital status m in year t. With this notation, the mechanical tax savings in year t due to the 1989-reform can be defined as

$$
\begin{aligned}
\Delta \mathrm{tax}_t^m = \ & (1-m) \ \times \ \mathbf{I}\left[\widetilde{W}_t^0 \geq \overline{W}_t^0\right] \times \left(\Delta\tau_t \cdot \left(\widetilde{W}_t^0 - \overline{W}_t^0\right)\right) \\
& +m \quad \times \quad \Big\{ \mathbf{I}\left[\overline{W}_t^0 \leq \widetilde{W}_t^1 < \overline{W}_t^1\right] \times \left(0.022 \cdot \left(\widetilde{W}_t^1 - \overline{W}_t^0\right)\right) \\
& \qquad\qquad + \mathbf{I}\left[\widetilde{W}_t^1 \geq \overline{W}_t^1\right] \times \left(0.022 \cdot \left(\overline{W}_t^1 - \overline{W}_t^0\right) + \Delta\tau_t \cdot \left(\widetilde{W}_t^1 - \overline{W}_t^1\right)\right) \Big\}.
\end{aligned}
\tag{19}
$$

The first line captures tax savings for singles and shows that they save $\Delta\tau_t$ on counterfactual wealth above the singles threshold. The second and third lines capture tax savings for couples. Couples who have counterfactual wealth between the singles threshold $\overline{W}_t^0$ and the couples threshold $\overline{W}_t^1$ save 2.2% on wealth above $\overline{W}_t^0$. Couples who have counterfactual wealth above the couples threshold $\overline{W}_t^1$ save 2.2% on the difference between the thresholds, $\overline{W}_t^1 - \overline{W}_t^0$, and another $\Delta\tau_t$ on wealth above the couples threshold.

---

[40]Notice that there were additional tax cuts (a rate reduction and a threshold increase) in 1996, the last year of the wealth tax before its abolishment. We do not include the mechanical effects of these tax changes, because we are interested in studying the effects of the 1989-reform on its own.

Equation (19) gives the mechanical tax savings within a given year t, but those tax savings will grow over time according to a rate of return that is not directly observed in the data. We assume a gross rate of return equal to $R = 1.07$, broadly in line with empirical estimates of rates of return and consistent with the calibration exercise in section 5. The cumulative tax savings for marital status m in year t can be defined as

$$\Delta\text{cumtax}_t^m = \begin{cases} \Delta\text{tax}_t^m & \text{if } t = 1989 \\ \Delta\text{tax}_t^m + \Delta\text{cumtax}_{t-1}^m \times R & \text{if } t > 1989 \end{cases} \tag{20}$$

The last step of the procedure is to translate the cumulative tax savings into a log effect on wealth. This is done as follows

$$\Delta\log W_t = \mathsf{E}\left[\log\left(\widetilde{W}_t^m + \Delta\text{cumtax}_t^m\right) - \log\left(\widetilde{W}_t^m\right) \mid t\right], \tag{21}$$

where $\mathsf{E}\left[\cdot\right]$ is the average across singles and married individuals, using the marriage share in the pre-reform year 1988.

The procedure described above provides a measure of the mechanical effect on individuals in the treatment group. We note that, while the control group does not save any taxes in the ceiling DD (as those bound by the ceiling face a fixed average tax rate), the control group in the couples DD does save some taxes. When we compare couples located in the exempted range (between $\overline{W}_t^0$ and $\overline{W}_t^1$) to couples located either below or above that range (below $\overline{W}_t^0$ or above $\overline{W}_t^1$), there will be tax savings for those located above. The tax savings are relatively small for the control group as a whole, because those located above $\overline{W}_t^1$ represent only about 20% of the control group. We calculate the mechanical tax savings for the control group using the same approach as for the treatment group.[41] The mechanical effect on wealth (that we subtract from the total estimated difference-in-differences effect) equals the effect for the treatment group minus the effect for the control group.

---

[41] The calculation of tax savings for the control group using equations (19)-(20) is based on their *actual* wealth rather than on their *counterfactual* wealth. This is not exactly right because their actual wealth includes the mechanical effect as well as a potential behavioral response. But the error thus introduced is likely to be minor. Having calculated Δcumtax_t^m for the control group, we convert it into a log effect on wealth using the equation $\Delta\log W_t = \mathsf{E}\left[\log\left(\widetilde{W}_t^m\right) - \log\left(\widetilde{W}_t^m - \Delta\text{cumtax}_t^m\right) \mid t\right]$, where $\widetilde{W}_t^m \equiv \mathsf{E}\left[W_{it}^m \mid m, t\right]$ denotes the average actual wealth.

B Proofs

To prove Proposition 1, we totally differentiate equation (10) with respect to $1 - \tau$ and c_0. After rearranging terms, this gives

$$\frac{dW_1}{d(1-\tau)} \left\{ \sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\frac{\alpha}{\sigma}-1} \right\} = \sum_{t=0}^{T} \frac{\partial q_t}{\partial(1-\tau)} \cdot c_0 + \frac{\partial q_b}{\partial(1-\tau)} \cdot c_0^{\frac{\alpha}{\sigma}}$$
$$+ \sum_{t=0}^{T} \frac{tRy_t}{((1-\tau)R)^{t+1}} + \sum_{t=1}^{T} \left\{ \frac{[(1-\tau)R + t\tau R]R\bar{W}}{((1-\tau)R)^{t+1}} \right\}, \quad (22)$$

where we have used $dc_0 = -dW_1$. From the definitions $q_t = \frac{(\delta(1-\tau)R)^{t\sigma}}{((1-\tau)R)^t}$ and $q_b = \frac{A(\delta(1-\tau)R)^{T\alpha}}{((1-\tau)R)^T}$, we obtain

$$\frac{\partial q_t}{\partial(1-\tau)} = \frac{\sigma t q_t}{1-\tau} - \frac{tR(\delta(1-\tau)R)^{t\sigma}}{((1-\tau)R)^{t+1}}, \quad (23)$$

$$\frac{\partial q_b}{\partial(1-\tau)} = \frac{\alpha T q_b}{1-\tau} - \frac{TRA(\delta(1-\tau)R)^{T\alpha}}{((1-\tau)R)^{T+1}}, \quad (24)$$

which allows us to rewrite equation (22) as follows

$$\frac{dW_1}{d(1-\tau)} \left\{ \sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\frac{\alpha}{\sigma}-1} \right\} = \sum_{t=0}^{T} \frac{\sigma t q_t}{1-\tau} \cdot c_0 + \frac{\alpha T q_b}{1-\tau} \cdot c_0^{\frac{\alpha}{\sigma}}$$
$$- \sum_{t=0}^{T} \frac{tR(\delta(1-\tau)R)^{t\sigma}}{((1-\tau)R)^{t+1}} \cdot c_0 - \frac{TRA(\delta(1-\tau)R)^{T\alpha}}{((1-\tau)R)^{T+1}} \cdot c_0^{\frac{\alpha}{\sigma}} \quad (25)$$
$$+ \sum_{t=0}^{T} \frac{tRy_t}{((1-\tau)R)^{t+1}} + \sum_{t=1}^{T} \left\{ \frac{[(1-\tau)R + t\tau R]R\bar{W}}{((1-\tau)R)^{t+1}} \right\}.$$

The first line on the RHS represents substitution effects from the EIS σ and the bequest elasticity α. The second and third lines represent wealth effects. To see this, we take the lifetime budget constraint (5) and totally differentiate it with respect to $(1-\tau)$ and W_0^n, holding behavior constant. This gives the compensating change in initial wealth W_0^n that would allow the household to buy an unchanged bundle of consumpion and bequests when taxes change. Denoting this compensating change by dW_0^C, we may write

$$\frac{dW_0^C}{d(1-\tau)} = - \sum_{t=0}^{T} \frac{tRc_t}{((1-\tau)R)^{t+1}} - \frac{TRW_{T+1}}{((1-\tau)R)^{T+1}}$$
$$+ \sum_{t=0}^{T} \frac{tRy_t}{((1-\tau)R)^{t+1}} + \sum_{t=1}^{T} \left\{ \frac{[(1-\tau)R + t\tau R]R\bar{W}}{((1-\tau)R)^{t+1}} \right\}. \quad (26)$$

The expression on the RHS of (26) is identical to the second and third lines of (25) once we account for the fact that, from household optimization, $c_t = (\delta (1 - \tau) R)^{t\sigma} c_0$ and $W_{T+1} = A (\delta (1 - \tau) R)^{T\alpha} c_0^{\frac{\alpha}{\sigma}}$. This allows us to rewrite equation (25) to

$$
\frac{dW_1}{d (1 - \tau)} \left\{ \sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\frac{\alpha}{\sigma} - 1} \right\} = \sum_{t=0}^{T} \frac{\sigma t q_t}{1 - \tau} \cdot c_0 + \frac{\alpha T q_b}{1 - \tau} \cdot c_0^{\frac{\alpha}{\sigma}} + \frac{dW_0^C}{d (1 - \tau)}, \tag{27}
$$

which can be immediately rearranged to give the expression in equation (11) of Proposition 1. The expression has three terms: a substitution effect on consumption governed by σ, a substitution effect on bequests governed by α, and the wealth effect captured by the compensating wealth change $\frac{dW_0^C}{d(1-\tau)}$. Equation (12) of Proposition 2 follows from setting $\alpha = \sigma$ (which implies $q_b = A q_T$) in equation (11). In this case the two substitution effects can be consolidated into one effect governed by a single structural elasticity $\alpha = \sigma$.

To prove Proposition 3, we totally differentiate equation (13) with respect to $1 - \tau$, W_t, and c_j. This gives

$$
\begin{aligned}
\frac{dW_t / W_t}{d (1 - \tau) / (1 - \tau)} &= (t - 1) [(1 - \tau) R]^{t-1} \frac{W_0^n}{W_t} + \sum_{j=0}^{t-1} (t - 1 - j) [(1 - \tau) R]^{t-1-j} \frac{y_j - c_j}{W_t} \\
&\quad - \sum_{j=1}^{t-1} \left(1 - \frac{\tau (t - 1 - j)}{1 - \tau} \right) [(1 - \tau) R]^{t-j} \frac{\bar{W}}{W_t} \\
&\quad - \sum_{j=0}^{t-1} [(1 - \tau) R]^{t-1-j} \frac{dc_j / W_t}{d (1 - \tau) / (1 - \tau)}.
\end{aligned} \tag{28}
$$

In this expression, the first three terms correspond to the mechanical effect dM shown in equation (15) of Proposition 3. The last term corresponds to the behavioral effect dB in equation (16). To see this, we start from the Euler equation shown in (8): $c_j = (\delta (1 - \tau) R)^{j\sigma} c_0$. We totally differentiate this condition and rearrange terms, which gives

$$
\frac{dc_j / W_t}{d (1 - \tau) / (1 - \tau)} = j\sigma (\delta (1 - \tau) R)^{j\sigma} \frac{c_0}{W_t} + (\delta (1 - \tau) R)^{j\sigma} \frac{dc_0 / W_t}{d (1 - \tau) / (1 - \tau)}. \tag{29}
$$

Using $dc_0 = -dW_1$, we can rewrite this to

$$
\begin{aligned}
\frac{dc_j / W_t}{d (1 - \tau) / (1 - \tau)} &= j\sigma (\delta (1 - \tau) R)^{j\sigma} \frac{c_0}{W_t} - (\delta (1 - \tau) R)^{j\sigma} \frac{dW_1 / W_0}{d (1 - \tau) / (1 - \tau)} \frac{W_0}{W_t} \\
&= (\delta (1 - \tau) R)^{j\sigma} \left\{ j\sigma \frac{c_0}{W_t} - \frac{dW_1 / W_0}{d (1 - \tau) / (1 - \tau)} \frac{W_0}{W_t} \right\}.
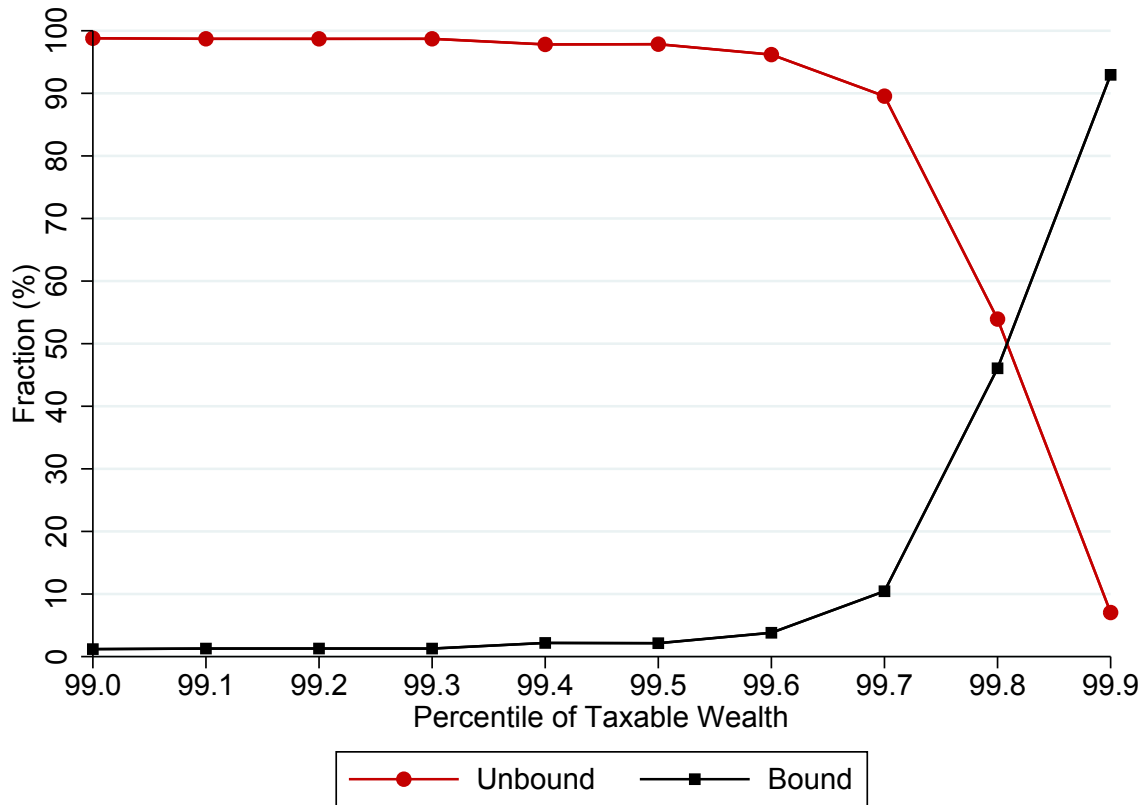\end{aligned} \tag{30}
$$

Finally, inserting (30) into (28) and using $q_j \equiv \frac{(\delta(1-\tau)R)^{j\sigma}}{((1-\tau)R)^j}$, we obtain

$$\frac{dW_t/W_t}{d(1-\tau)/(1-\tau)} = (t-1)\left[(1-\tau)R\right]^{t-1}\frac{W_0^n}{W_t} + \sum_{j=0}^{t-1}(t-1-j)\left[(1-\tau)R\right]^{t-1-j}\frac{y_j-c_j}{W_t}$$
$$-\sum_{j=1}^{t-1}\left(1-\frac{\tau(t-1-j)}{1-\tau}\right)\left[(1-\tau)R\right]^{t-j}\frac{\bar{W}}{W_t} \qquad (31)$$
$$+\sum_{j=0}^{t-1}\frac{q_j}{\left[(1-\tau)R\right]^{1-t}}\left\{\frac{dW_1/W_0}{d(1-\tau)/(1-\tau)}\frac{W_0}{W_t} - j\sigma\frac{c_0}{W_t}\right\},$$

which is the result in Proposition 3.

C Appendix Figures

FIGURE A.I: FRACTIONS BOUND AND UNBOUND BY TAX CEILING

Notes: The figure shows the fraction of households bound and unbound by the tax ceiling at different quantiles of the wealth distribution (within the top 1%) in the pre-reform year, 1988.

FIGURE A.II: HETEROGENEITY IN BUNCHING

A: Employees

B: Self-Employed

C: "Ordinary" Taxpayers

D: "Non-Ordinary" Taxpayers
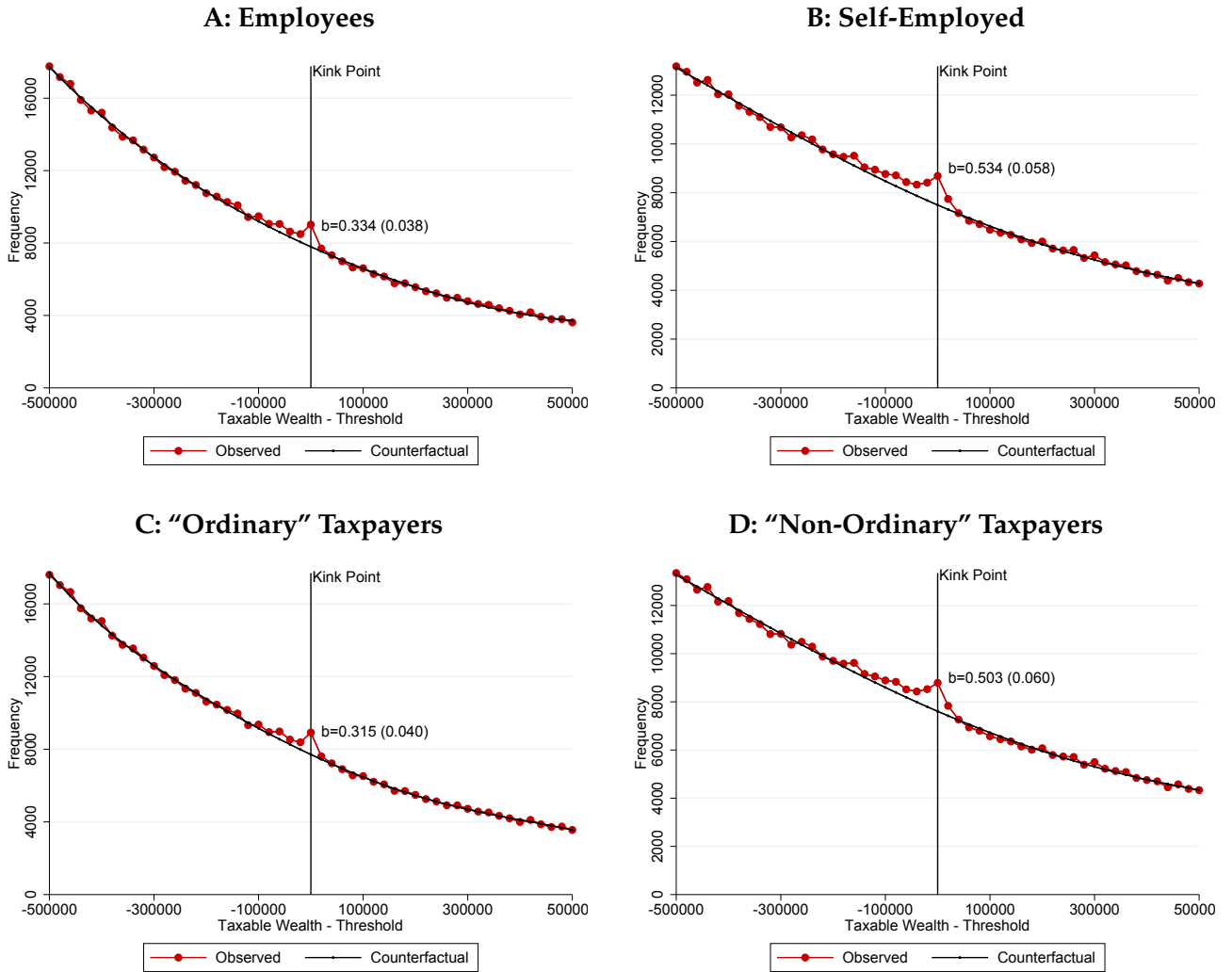
Notes: The figure shows heterogeneity in bunching across samples with different opportunities to avoid or evade taxes. Panels A and B compare employees and self-employed individuals, while Panels C and D compare "ordinary" and "non-ordinary" taxpayers. Non-ordinary taxpayers are defined as those who have self-employment income or who have foreign wealth or income. The diagrams are otherwise constructed in the same way as Figure 4, Panel B. The bunching estimate b varies with opportunity to evade in the expected direction, but the differences are not very large.

FIGURE A.III: BUNCHING BEFORE AND AFTER THE 1989-REFORM

A: Before 1989-Reform
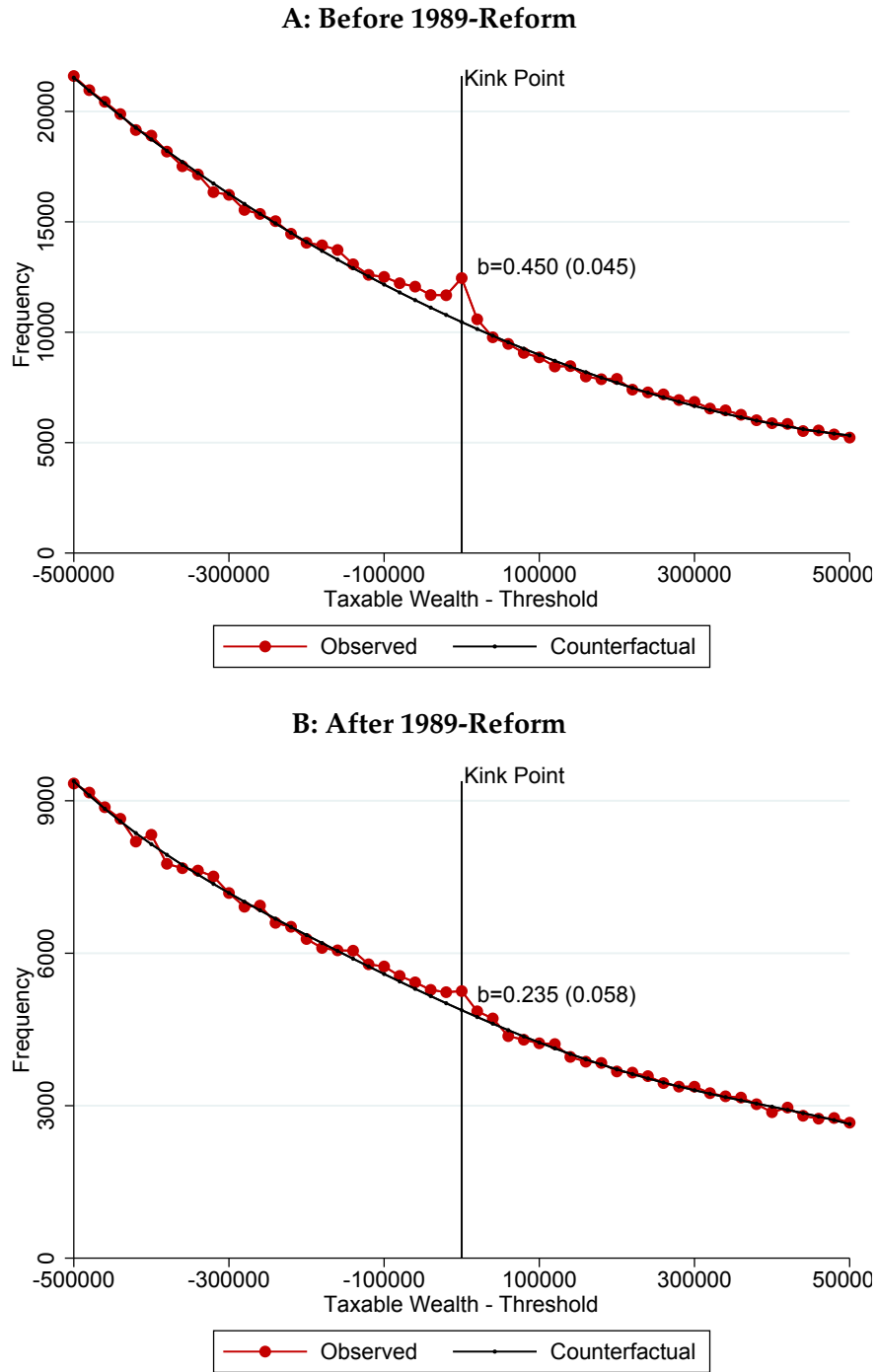
B: After 1989-Reform

Notes: The figure shows bunching at the kink before the 1989-reform (pooling 1980-88) and after the 1989-reform (pooling 1989-96). The diagrams are otherwise constructed in the same way as Figure 4, Panel B. As expected, the amount of bunching b is larger before the reform when the size of the wealth tax kink was larger.

FIGURE A.IV: CEILING DD USING THIRD-PARTY REPORTED WEALTH
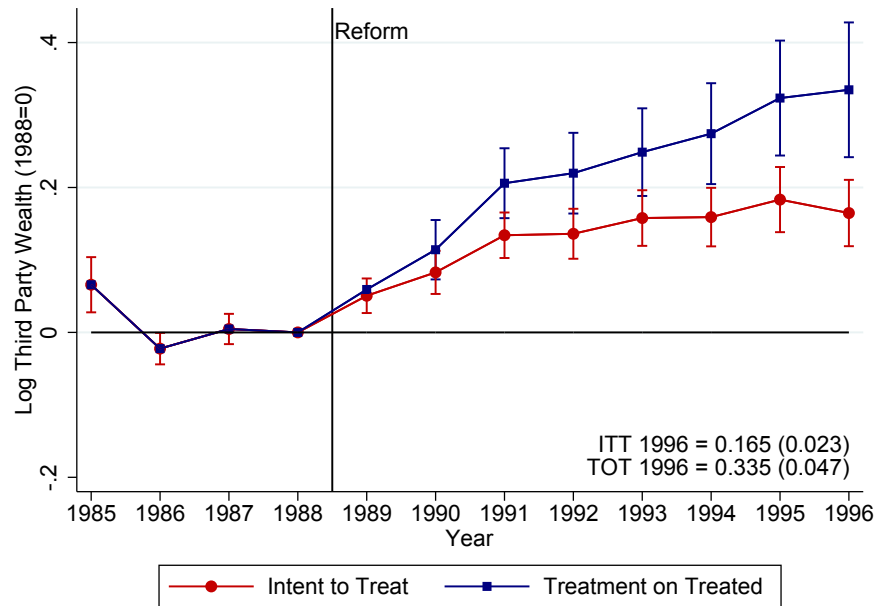
A: Intent to Treat vs Treatment on Treated

B: Behavioral vs Mechanical Effects

Notes: The figure shows the effects of wealth tax reform on "third-party reported wealth" based on the ceiling DD approach. Third-party reported wealth is defined as (taxable wealth – cash – durables – foreign wealth). Panel A shows intent-to-treat (ITT) estimates and treatment-on-the-treated (TOT) estimates of the total effect of tax reform. Panel B decomposes the total effect into mechanical and behavioral effects. Besides the difference in wealth measure, the two panels are constructed in the same way as the analogous diagrams shown earlier.
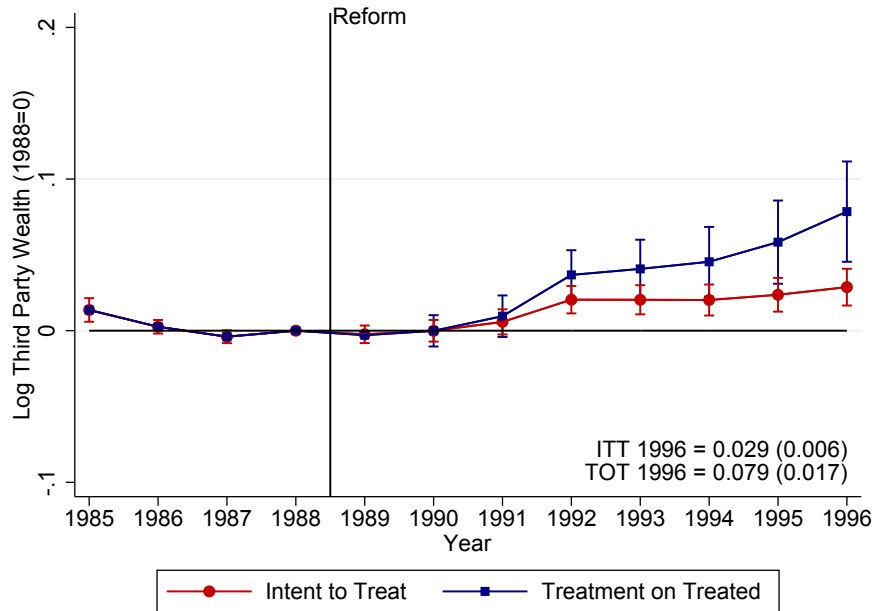
A: Taxable Wealth

B: Third-Party Reported Wealth

Notes: The figure shows the effects of wealth tax reform based on a ceiling DD where we use only strongly bound households as controls. Specifically, the control group consists of households whose wealth tax liability would have to fall by at least 50% (as opposed to 20% in the baseline) for them to become unbound by the tax ceiling. We show intent-to-treat (ITT) and treatment-on-the-treated (TOT) estimates for both taxable wealth (Panel A) and third-party reported wealth (Panel B). Besides the difference in control group, the two panels are constructed in the same way as the analogous diagrams shown earlier.

FIGURE A.VI: CEILING DD UNDER DFL REWEIGHTING VS REGRESSION CONTROLS

A: DFL Reweighting
Housing Share Deciles

ITT 1996 = 0.097 (0.026)
TOT 1996 = 0.196 (0.052)

B: Non-Parametric Controls
Housing Share Deciles

ITT 1996 = 0.086 (0.022)
TOT 1996 = 0.175 (0.044)

C: DFL Reweighting
Housing Share, Equity Share, and Income Quintiles

ITT 1996 = 0.118 (0.029)
TOT 1996 = 0.239 (0.060)

D: Non-Parametric Controls
Housing Share, Equity Share, and Income Quintiles

ITT 1996 = 0.146 (0.022)
TOT 1996 = 0.297 (0.044)

Notes: The figure compares DD estimates obtained from our baseline specification (1) to DD estimates obtained from DFL reweighting. While the baseline specification is based on controlling non-parametrically for differences in pre-reform variables that may impact wealth trends, the DFL specification reweights treatment and control groups to have the same pre-reform distribution of those variables. The figure compares two different DFL specifications (left panels) to the corresponding regression controls (right panels). We consider parsimonious specifications in Panels A-B (adjusting for housing share deciles only) and richer specifications in Panels C-D (adjusting for housing share, equity share, and income quintiles).

FIGURE A.VII: COUPLES DD USING THIRD-PARTY REPORTED WEALTH
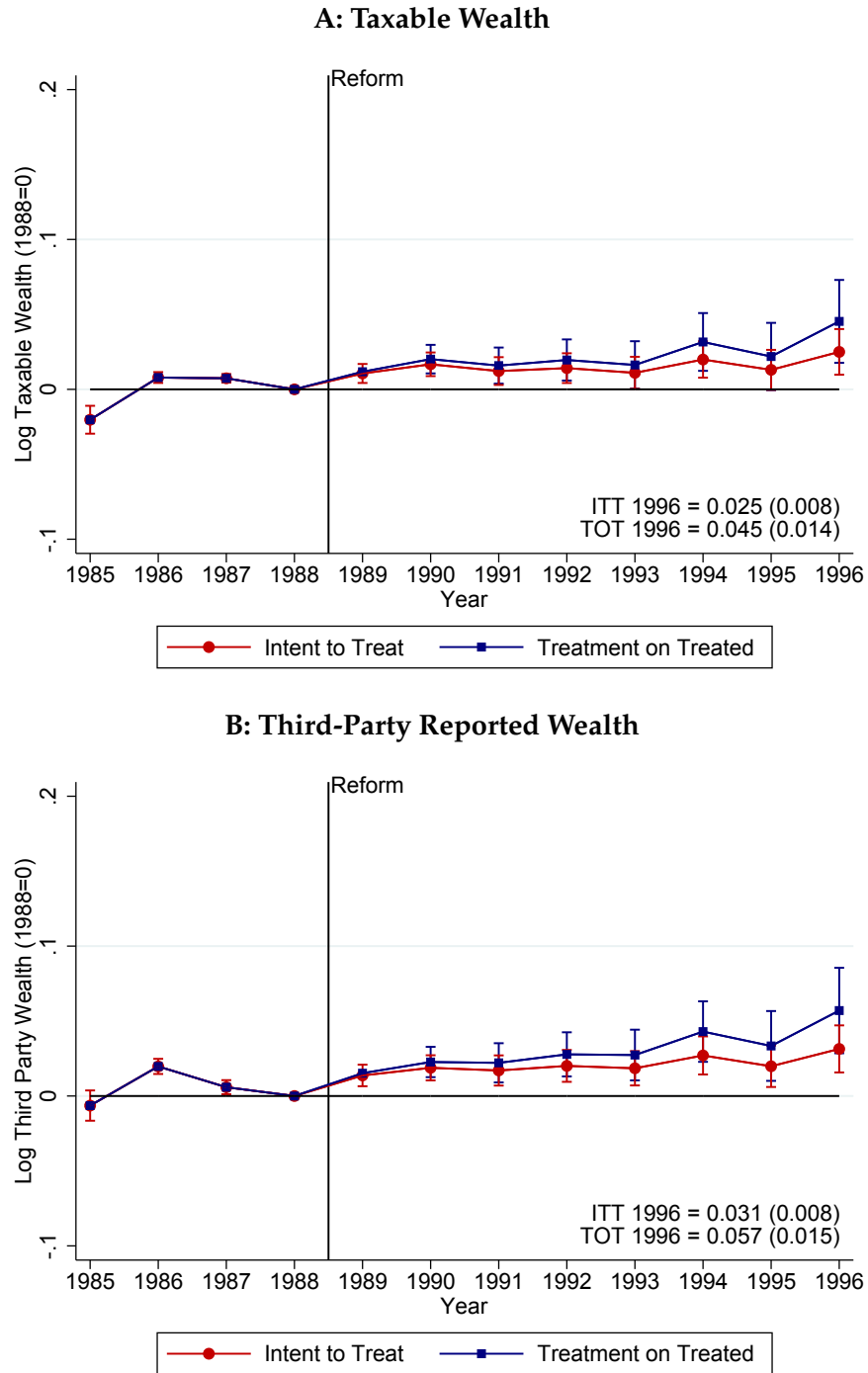
A: Intent to Treat vs Treatment on Treated

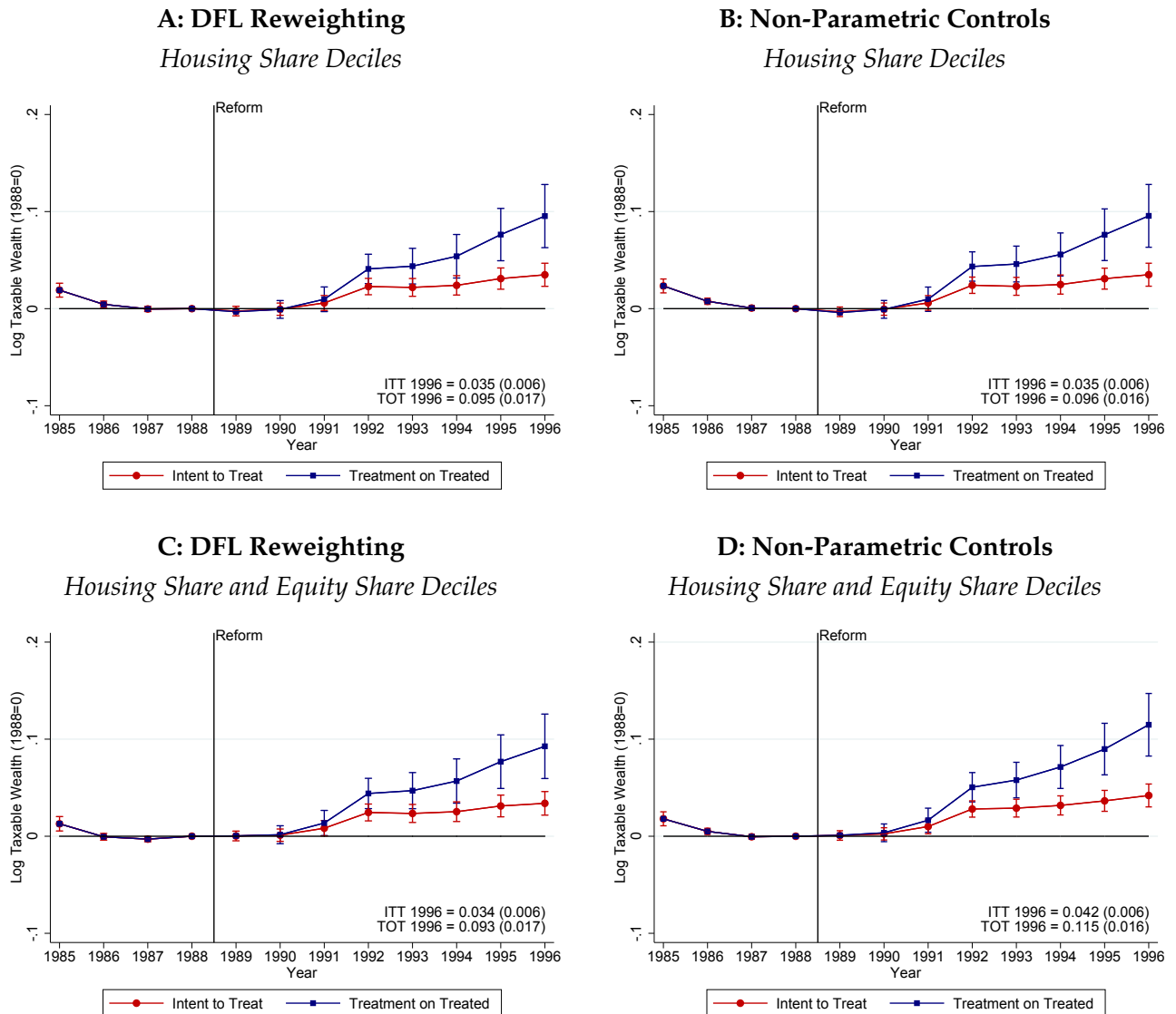### B: Behavioral vs Mechanical Effects

Notes: The figure shows the effects of wealth tax reform on "third-party reported wealth" based on the couples DD approach. Third-party reported wealth is defined as (taxable wealth – cash – durables – foreign wealth). Panel A shows intent-to-treat (ITT) estimates and treatment-on-the-treated (TOT) estimates of the total effect of tax reform. Panel B decomposes the total effect into mechanical and behavioral effects. Besides the difference in wealth measure, the two panels are constructed in the same way as the analogous diagrams shown earlier.

A: Taxable Wealth

B: Third-Party Reported Wealth

Notes: The figure shows the effects of wealth tax reform based on a couples DD where we use singles inside the exempted wealth range as controls. That is, while the baseline specification compares couples inside the exempted range to couples outside the exempted range, we are now comparing couples to singles in the same wealth range. Both of these groups get wealth tax cuts, but the singles get a smaller cut than couples. We show intent-to-treat (ITT) and treatment-on-the-treated (TOT) estimates for both taxable wealth (Panel A) and third-party reported wealth (Panel B). Besides the difference in control group, the two panels are constructed in the same way as the analogous diagrams shown earlier.

A: DFL Reweighting

Housing Share Deciles

B: Non-Parametric Controls

Housing Share Deciles

C: DFL Reweighting

Housing Share and Equity Share Deciles

D: Non-Parametric Controls

Housing Share and Equity Share Deciles

Notes: The figure compares DD estimates obtained from our baseline specification (1) to DD estimates obtained from DFL reweighting. While the baseline specification is based on controlling non-parametrically for differences in pre-reform variables that may impact wealth trends, the DFL specification reweights treatment and control groups to have the same pre-reform distribution of those variables. The figure compares two different DFL specifications (left panels) to the corresponding regression controls (right panels). We consider parsimonious specifications in Panels A-B (adjusting for housing share deciles only) and richer specifications in Panels C-D (adjusting for both housing share and equity share deciles).

FIGURE A.X: EMPIRICAL LIFECYCLE PROFILES OF WEALTH AT THE EXTREME TOP (IN TERMS OF PERCENTILES)
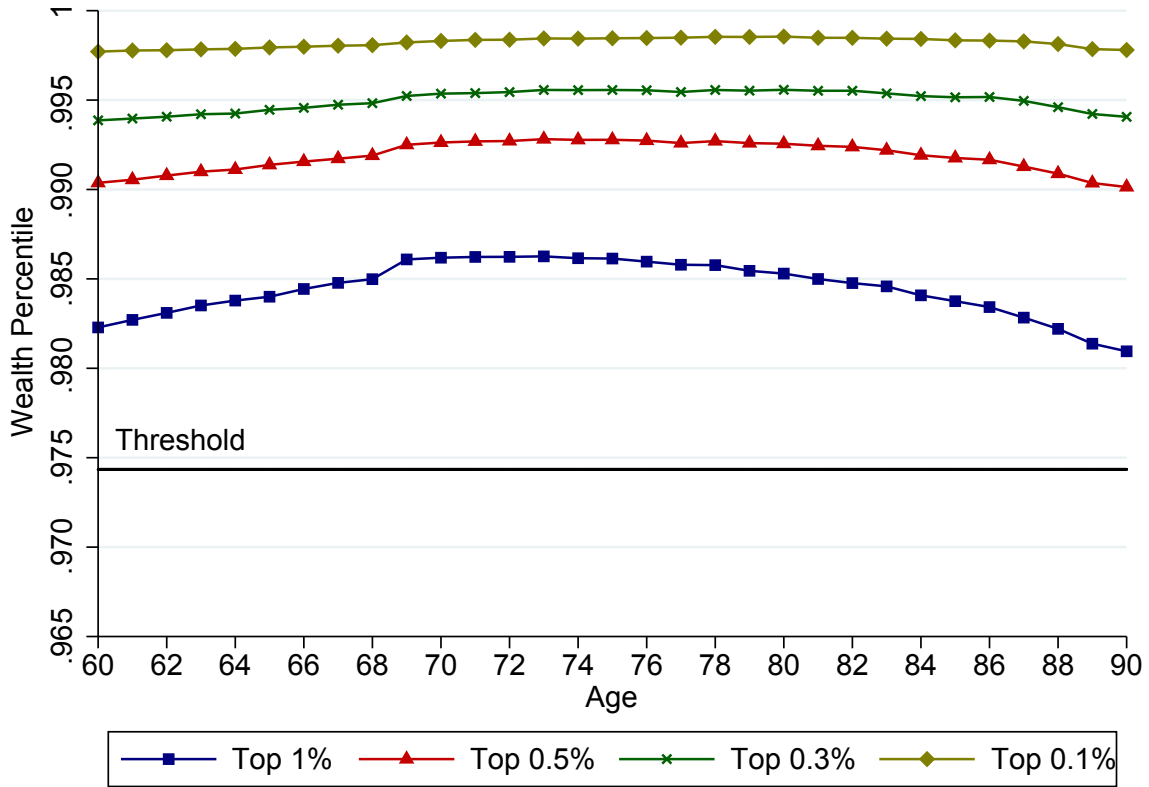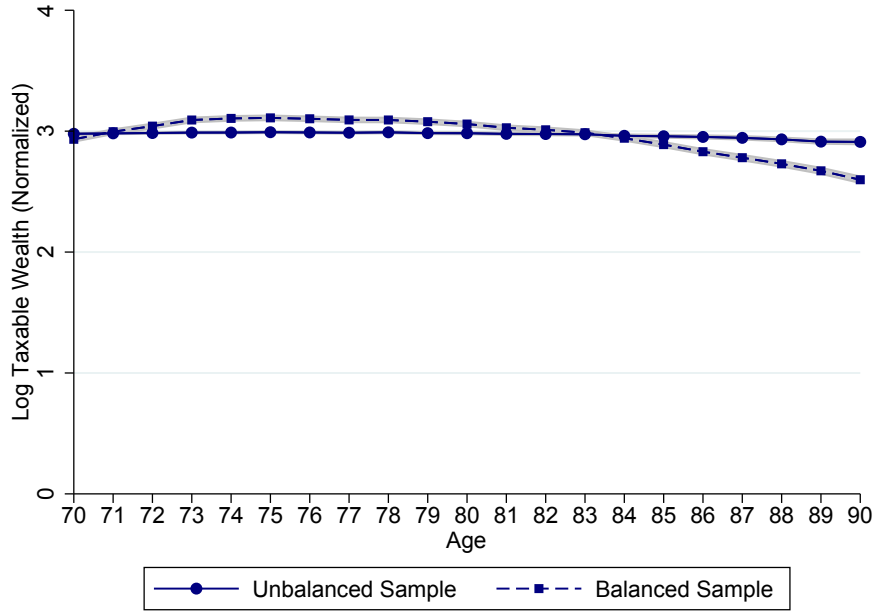

Notes: The figure shows lifecycle profiles of taxable wealth between the ages of 60-90 in an unbalanced panel of individuals over the period 1980-2012. The figure is constructed in the same way as Figure 11, except that we include the extreme top of the distribution (the top 0.3% and 0.1%) and that the y-axis is expressed in terms of wealth percentiles instead of wealth amounts. The lifecycle profile for the top $p\%$ includes individuals who reach the top $p\%$ during their observed lifespan (for at least three years), but who are not necessarily there in every year. This explains why the "top $p\%$" is always below the pth percentile cutoff on average. The graph shows that the top-1% sample (located between the 98th and 99th percentile cutoffs throughout) corresponds well to the treatment group in the couples DD, while the top-0.3% sample (located in the middle of the top 1% throughout) corresponds well to the treatment group in the ceiling DD.
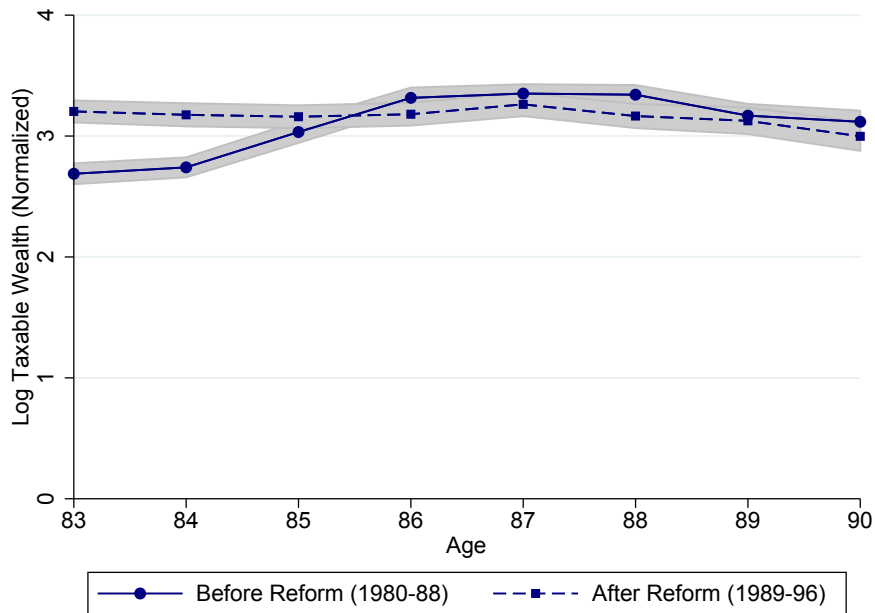
A: Balanced vs Unbalanced

## B: Before Reform vs After Reform (Both Balanced)

Notes: The figure shows lifecycle profiles of taxable wealth for the top 1% in a balanced panel. Panel A compares the lifecycle profiles in the balanced and unbalanced panels between the ages of 70 to 90. Panel B compares the lifecycle profiles before and after the tax reform in the balanced panel between the ages of 83 to 90. The figure is otherwise constructed as the previous lifecycle graphs, including the fact that we consider averages of normalized log wealth (to avoid the confounding effects of inflation) and the way in which we select the top 1% sample (those who make it to the top 1% during the lifetime, but who are not necessarily there every year). The grey-shaded areas represent 95% confidence intervals.

FIGURE A.XII: CONSUMPTION PROFILES BEFORE AND AFTER WEALTH TAX REFORM

**A: Couples DD (Moderately Wealthy)**
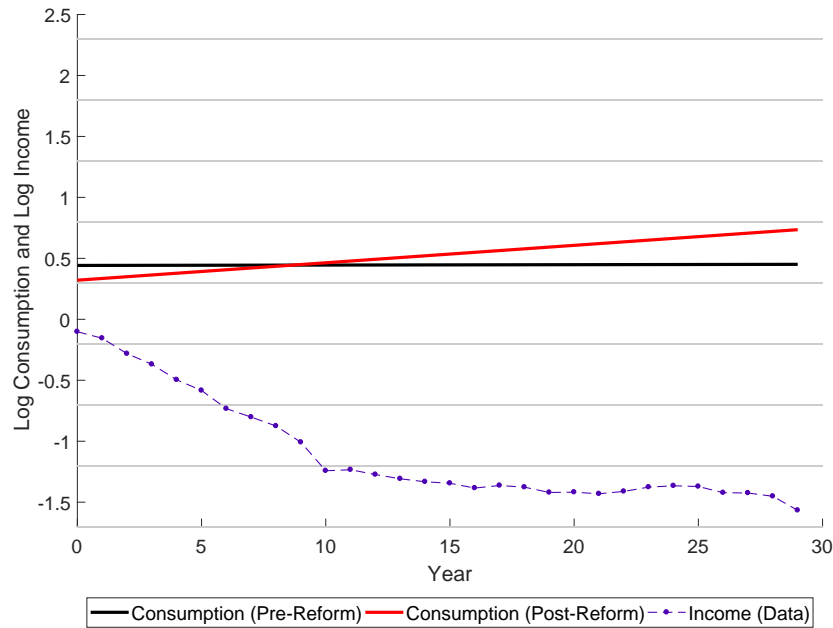

B: Ceiling DD (Very Wealthy)
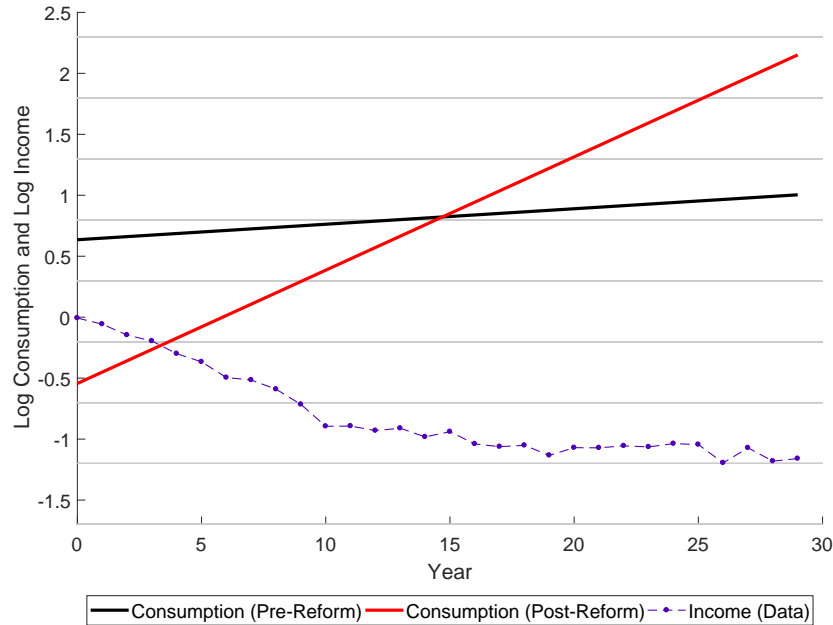
Notes: The figure shows the simulated consumption profiles before and after the wealth tax reform and the observed income profile. Panel A shows the scenario where we calibrate the model to fit the estimates from the couples DD (the moderately wealthy). Panel B shows the scenario where we calibrate the model to fit the estimates from the ceiling DD (the very wealthy). These samples of older, wealthy people dissave in the baseline as consumption is larger than income. When wealth taxes are cut, their dissavings fall in the beginning of the period and rise later in the period.

FIGURE A.XIII: CALIBRATING THE BEQUEST ELASTICITY

A: Couples DD (Moderately Wealthy)
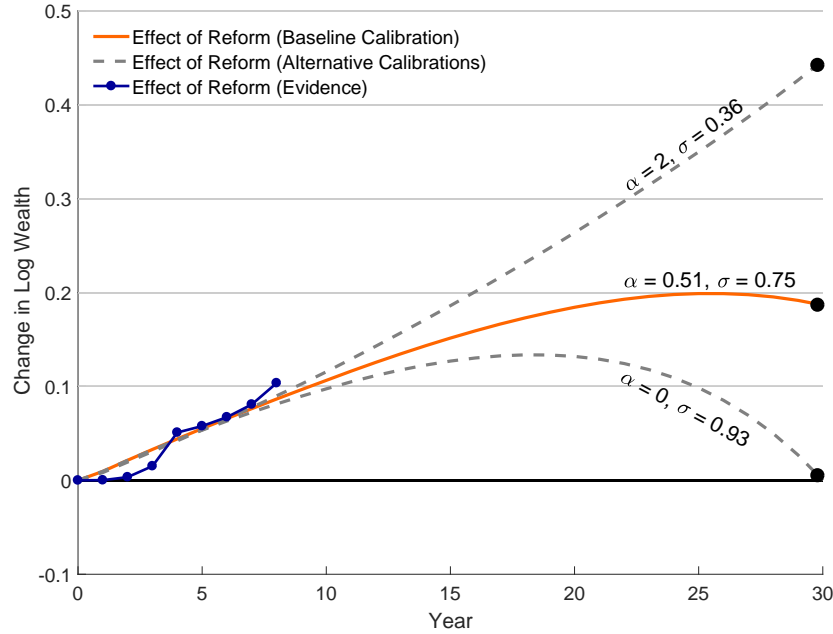
B: Ceiling DD (Very Wealthy)
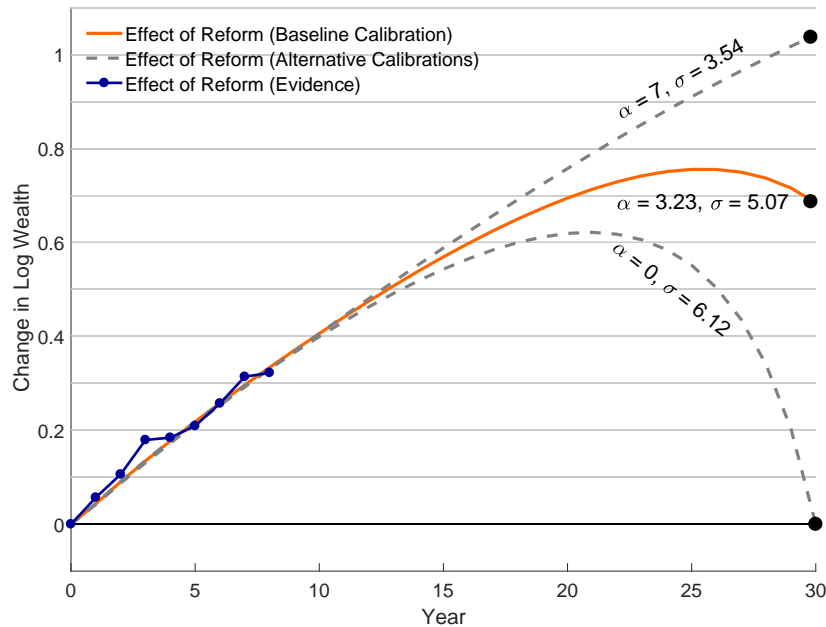
Notes: The figure illustrates how the calibration of the bequest elasticity α (in conjunction with the EIS σ) works. The calibration has to satisfy two requirements: (i) it has to fit the quasi-experimental estimates for years 1,...8 (based on minimizing equation 17), (ii) it has to fit the flatness of the wealth profile over the last 10 years of life. The figure shows that, if one ignored the second requirement and assumed alternative values for α, it is still possible to match the year 1-8 estimates by adjusting σ, but the lifecycle profile at the end of life would look very different from what is observed in the data. We show this for both the couples DD (Panel A) and the ceiling DD (Panel B).