WEALTH TAXATION AND WEALTH ACCUMULATION: THEORY AND EVIDENCE FROM DENMARK[*]

Katrine Jakobsen
Kristian Jakobsen
Henrik Kleven
Gabriel Zucman

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 greatly reduced starting in 1989 and later abolished. Due to the specific design of the wealth tax, the 1989 reform provides 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 taxable wealth with respect to the net-of-tax return is sizeable at the top of distribution. *JEL* Codes: H20, H31, E21, D31

I 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 and 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. 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 long-run effects, 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, corresponding to a very high rate on the *return* to wealth.[1] The Danish government implemented large changes to the wealth tax starting in 1989 — cutting the marginal rate to 1% and doubling the exemption threshold for married couples — before eventually abolishing the tax in 1997. These policy changes represent some of the largest natural experiments with wealth taxation ever conducted. In addition, 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 reform on wealth accumulation. We consider two different empirical strategies and samples. One 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 singles in the same range or to couples in other ranges. The other 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 very wealthy using a difference-in-differences design comparing bound and unbound taxpayers within the top 1%.[2]

The quasi-experimental analysis shows that wealth taxes have sizeable effects on taxable wealth,

---

[1]For example, assuming a rate of return on wealth equal to 4.4%, a marginal wealth tax of 2.2% corresponds to a 50% tax on the flow of capital income.

[2]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.

with the effects being considerably 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.[3] The effect on wealth builds up over time and is equal to about 19% after 8 years for the moderately wealthy (couples DD) and 31% after 8 years for the very wealthy (ceiling 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 effect accounts for about one-tenth of the effect for the couples DD and about one-fifth of the effect for the ceiling DD.

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 relatively old — most of those in the top 1% are above 50 years of age — 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 Substitution (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 above. 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 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 to investigate 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 19% after 8 years — and simulating the model forward, we obtain a 30-year effect of 30%. When matching the model to the very wealthy in the ceiling DD — an empirical effect on taxable wealth of 31% after 8 years — the long-run effect is considerably larger, 65% after 30 years. While these effects may seem large, note that the underlying tax incentives driving them are also large. The implied long-run elasticity of taxable wealth with respect to the after-tax rate of return equals 0.77 for the moderately wealthy and 1.15 for the very wealthy.

Given our estimates rely on tax records, they capture both real responses and evasion/avoidance responses. Most assets were third-party reported under the Danish wealth tax — thus limiting

---

[3]As we clarify below, the pre-trends are not always parallel in the raw data, but they are parallel after adjusting for linear, group-specific pre-trends.

the scope for evasion — but some were self-reported and therefore susceptible to evasion. If the amount of evasion responds to changes in marginal tax rates, this will be picked up by our elasticity estimates. Wealthy taxpayers may have access to relatively effective evasion vehicles such as offshore accounts.[4] It is worth noting, however, that repatriation of offshore wealth is unlikely to be part of our estimated responses. The wealth tax cuts did not come with an amnesty for previously unpaid taxes, implying that such repatriation would trigger back taxes and potential penalties. But our estimates may reflect other forms evasion and avoidance, and it would be difficult to separate out those responses.[5]

Our paper can be viewed in two ways. One view is that it contributes to a nascent literature studying the effects of wealth taxes on taxable wealth (Zoutman 2015; Brülhart, Gruber, Krapf, and Schmidheiny 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. Unlike earlier work, our paper provides a tractable dynamic framework to shed light on the theoretical mechanisms driving the reduced-form impacts, and it structurally estimates the model 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 and Stantcheva (2018) show that the long-run elasticity of capital supply 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 and 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.[6]

---

[4]Using leaked data from HSBC Switzerland and Mossack Fonseca ("Panama Papers"), Alstadsæter, Johannesen, and Zucman (2019) 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%.

[5]One possible strategy is to consider bunching at the kink point created by the exemption threshold. In the context of wealth taxation, bunching almost surely reflect evasion/avoidance responses rather than real responses. We show that there is very little bunching at the kink, consistent with modest evasion/avoidance responses. However, bunching at the kink may understate responsiveness within brackets, which prevents us from using the bunching evidence to rule out significant evasion and avoidance.

[6]To be clear, our quasi-experimental estimates do not represent "all-inclusive" long-run elasticities of capital supply with respect to capital taxes. As we discuss below, our estimates capture the effect of wealth taxes on the already-wealthy

In the process of producing the findings described above, we provide a number of bonus contributions. It is worth highlighting some of those here. First, our structural approach yields an estimate of the bequest elasticity with respect to the net-of-tax rate on capital.[7] 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. Piketty and Saez (2013) highlight that the bequest elasticity is a key parameter for optimal inheritance taxation. Reviews by Kopczuk (2009, 2013a,b) summarize the few existing estimates of this parameter 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. Our findings suggest that bequest elasticities are large at the top of the wealth distribution.[8]

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 particularly 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 just 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 of the paper contributes to an empirical literature documenting age-wealth profiles among the elderly (see e.g., Love, Palumbo, and Smith 2009 and Poterba, Venti, and Wise 2011, 2018). While these studies relied on small survey datasets, the precision and power of our data allow for providing clean graphical evidence even for the extreme top of the wealth distribution. We also contribute to a literature trying to explain wealth concentration and the lifecycle saving behavior of the rich (see e.g., Carroll 2002; De Nardi 2004; Kaplow 2011; Benhabib and Bisin 2018).

The paper is organized as follows. Section II describes the data and documents the evolution of wealth inequality, section III presents quasi-experimental evidence on the effects of wealth taxes, section IV develops the theoretical model, section V combines the model and quasi-experimental evidence to structurally estimate long-run effects of wealth taxes, and section VI concludes. All

---

rather than the forward-looking effect on those who aspire to become wealthy. The potential aspiration effect of wealth taxes — including career decisions, entrepreneurial risk-taking, and early-life saving decisions made in anticipation of future taxes — would be extremely difficult, if not impossible, to identify convincingly.

[7]Or more precisely, we estimate the elasticity of end-of-life wealth. While we refer to this as a "bequest elasticity", we are not able to distinguish between actual bequest motives and other motives driving utility of residual wealth (e.g., the "capitalist spirit" as discussed in Carroll 2002).

[8]These bequest elasticities (i.e., residual-wealth elasticities) will include any changes in inter-vivo transfers and gifts. Such gift responses are most naturally interpreted as tax evasion. There is a tax exemption threshold for gifts, but it is very small relative to the wealth levels of our population of interest. Gifts above the exemption threshold are not desirable from a tax minimization perspective. However, unreported gifts (in cash or in kind) are difficult to detect for tax authorities and may respond to wealth taxes. The Danish data does not allow us to provide direct evidence on such responses, but they will be part of our estimates of taxable wealth responses.

appendix material is presented in the online appendices.

II Danish Household Wealth: Data and Distribution

II.A Wealth Data

We base our analysis 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.

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 assets and liabilities are reported by third-parties to the Danish government, which makes the data very reliable (see Boserup, Kopczuk, and Kreiner 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 financial institutions (banks, mutual funds, and insurance companies), and the value of mortgages is reported by mortgage lenders. Non-financial assets are recorded using land and real estate registries. 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, and interpersonal debts.

The Danish wealth data are considered to be of a very high quality, and they have been used to study retirement savings (Chetty, Friedman, Leth-Petersen, Nielsen, and Olsen 2014), intergenerational wealth mobility (Boserup, Kopczuk, and Kreiner 2014), and the accuracy of survey responses (Kreiner, Lassen, and Leth-Petersen 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. Moreover, 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 wealth series in 1997, the year in which 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

---

[9]Moreover, there were changes in the coverage of third-party wealth reporting in 1997, implying that even third-party reported wealth by itself suffers from a data break.

for which we have consistently measured taxable wealth both before and after the reform.

II.B   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 aggregate wealth at market value recorded in Denmark's household balance sheet. This implies that the wealth levels and wealth shares for Denmark are comparable to existing series for other countries, including those estimated for the United States by Saez and Zucman (2016).[10] In keeping with standard national account concepts, our definition of wealth includes all financial and non-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. Average wealth per adult person was $237,000 in 2012 (using the market exchange rate to convert Danish kroner to U.S dollars), a level similar 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 2019 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. 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 that individual-level pension wealth was added to the administrative data from 2012 onwards.[11] 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 equals the aggregate stock of business assets from the national accounts divided by the aggregate flow of business income from individual income tax returns), while we impute unlisted equities by capitalizing dividend income. 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.

---

[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, Chancel, Piketty, Saez, and Zucman 2017).

[11]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 and Zucman (2016) use a similar imputation procedure for the United States.

II.C   Trends in Wealth Concentration

Figure I shows wealth shares in three broad classes: the bottom 50%, the next 40%, and the top 10%. These wealth shares have 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 II 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 U.S. In 2012, the top 1% accounts for about 20% of total wealth in Denmark, while it accounts for almost 40% in the U.S. Average wealth in the population is similar in the two countries, but the top 1% are twice as wealthy in the U.S as they are in Denmark.[12] 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 U.S. 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 II, the top 0.1% wealth share in Denmark was only 2-3 percentage points lower than in the U.S 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 U.S. 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.[13]

---

[12]The average person in the top 1% of the U.S distribution owns net wealth of $9.3 million in 2012 (roughly 40 times average wealth), as opposed to $4.8 million in Denmark (roughly 20 times average wealth).

[13]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.

III The Effect of Wealth Taxes: Evidence

III.A Tax Variation and Empirical Strategies

Denmark taxed wealth on an annual basis until 1997. 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 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 III.

This setting offers two sources of exogenous variation: the kink point at the exemption threshold and the tax reforms. Let us first consider the former. The kink point is very sharp 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. As a result, taxpayers have strong incentives to bunch at the kink, allowing for a bunching approach to estimating taxable wealth responses.[14] 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. 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 present a bunching analysis in section A of the online appendix, but we view this evidence primarily as informative of avoidance responses.[15]

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. To estimate behavioral responses to the 1989 reform, we consider difference-

---

[14]Kleven (2016) provides a review of bunching approaches. In the context of wealth taxation, Seim (2017) presents bunching evidence using a kink point in the Swedish wealth tax. He finds clear bunching at the kink, but the implied elasticity of taxable wealth is small.

[15]There is some bunching at the Danish wealth tax kink, but it is small — even smaller than in the Swedish context analyzed by Seim (2017) — and the implied elasticity of taxable wealth is tiny. The presence of bunching is useful for rejecting the null of no avoidance responses, but bunching is unlikely to capture the global responsiveness of avoidance to changes in tax rates. In a world with fixed avoidance costs and/or lumpy assets ("optimization frictions"), the amount of bunching understates the responsiveness of avoidance as taxpayers may overshoot or undershoot the threshold.

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.[16]

The first 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 III. The implication of the reform is that couples in a certain range of the household wealth distribution — 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 below the range (controls). We refer to this strategy as the "couples DD".

The second 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 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.[17] 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. 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. 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 behavioral responses by the very wealthy. We will compare taxpayers unbound by the ceiling (treatments) to taxpayers bound by the ceiling (controls) within the top 1% of the distribution. We refer to this strategy as the "ceiling DD".[18]

---

[16]A compelling aspect of estimating wealth responses using the 1989 tax reform is that there was essentially no room for anticipation effects. The tax bill was first proposed on November 3, 1988, it was passed in parliament on December 21, 1988, and it took effect on January 1, 1989. Given that wealth accumulation is forward-looking variable, the absence of anticipation is important for the validity of our empirical strategies.

[17]The definition of taxable income used to assess the ceiling rule included labor income, pension income, and capital income (interest income, dividends, capital gains, etc.). The exact income definition was quite complicated and underwent some changes over time. See Skatteministeriet (2002) for details on the history of the personal income tax code in Denmark.

[18]It is useful to define the treatment and control groups more formally. If we denote taxable income by z, taxable wealth by W, and the wealth exemption threshold by $\bar{W}$, then the tax ceiling binds iff

$$\frac{t \cdot z + \tau \cdot (W - \bar{W})}{z} \geq 0.78 \qquad \Leftrightarrow \qquad \frac{W - \bar{W}}{z} \geq \frac{0.78 - t}{\tau},$$

where t is the average income tax rate (including social security taxes) and τ is the marginal wealth tax rate. With a top marginal income tax rate of 68% at the time of this reform (the upper bound on the *average* income tax rate) and a wealth tax rate of 2.2%, the ceiling starts binding for those with a taxable wealth-income ratio $(W - \bar{W})/z$ of at least five and typically much higher. In the data, the average household bound by the ceiling has a wealth-income ratio well

The empirical analysis is based on a standard DD event study specification, i.e.

$$\log W_{it} = \sum_{j \neq 1988} \beta_j \cdot Year_{j=t} \cdot Treat_{it} + \gamma_i + \eta_t + \nu_{it}, \tag{1}$$

where W_{it} denotes the wealth of household i in year t, $Year_{j=t}$ is a dummy equal to one when the year equals t, $Treat_{it}$ is a dummy equal to one when household i is in the treatment group at time t, γ_i is a household fixed effect, η_t is a year fixed effect, and ν_{it} is an error term. The DD coefficient β_t captures the effect of the tax reform in year t relative to the pre-reform year, 1988.

The assignment of treatment status depends on the empirical strategy, either being a couple in the exempted wealth range (couples DD) or being unbound by the tax ceiling (ceiling DD). A basic issue with specification (1) is that concurrent treatment status $Treat_{it}$ is endogenous to the outcome variable. To avoid bias from such endogeneity, we construct instruments based on pre-reform variables. Specifically, defining $Treat_i^{pre}$ as an indicator for being in the treatment group based on pre-reform behavior, we instrument $Year_{j=t} \cdot Treat_{it}$ in equation (1) using $Year_{j=t} \cdot Treat_i^{pre}$. For pre-reform treatment status to be a strong predictor of post-reform treatment status, household behavior must be sufficiently persistent over time. To increase persistence, we focus on households with the same status in several consecutive pre-years. As a baseline, treatment status is assigned based on six pre-reform years (1982-88), but we show that results are robust to shorter and longer treatment windows.[19]

The IV coefficients $\hat{\beta}_t$ estimated based on (1) represent treatment-on-the-treated (TOT) effects. We also consider reduced-form specifications, i.e. regressing log wealth directly on the instruments, $Year_{j=t} \cdot Treat_i^{pre}$. The reduced-form coefficients represent intention-to-treat (ITT) effects. The discrepancy between ITT and TOT effects depends on the persistence of treatment status. Even though we consider households with the same treatment status over several pre-reform years, their status is not perfectly persistent over the post-reform period. This attenuates the ITT effects relative to the TOT effects.

As always, the difference-in-differences approach relies on the assumption of parallel trends. We assess the validity of this assumption by inspecting the pre-trends of the comparison groups, and where these are not parallel, we adjust the series for differential pre-trends. Specifically, using n pre-reform years to estimate the trend, we consider the following extension of the baseline event study specification

$$\log W_{it} = \sum_{j \notin (1988-n, 1988)} \beta_j \cdot Year_{j=t} \cdot Treat_{it} + \theta^T \cdot t \cdot Treat_i^{pre} + \gamma_i + \eta_t + \nu_{it}, \tag{2}$$

where θ^T is a linear differential pre-trend identified based on n pre-reform years (i.e., the omitted

---

above ten. Since taxable income z includes capital income, the households affected by the ceiling tend to be those with large assets in a form that do not generate a correspondingly large flow income. Examples include valuable real estate or equity with unrealized capital gains. Such households are the controls in the ceiling DD.

[19]Households who switch status during the specified treatment window (1982-88 in the baseline) are dropped from the estimation sample.

years in the first term on the right-hand side). In the implementation below, we estimate the pre-trend using four pre-reform years.[20]

These specifications give the effect of "treatment" on log wealth, where treatment results either from the exemption of couples in a certain wealth range or from the tax rate cut on households unbound by the ceiling. It is useful to convert these effects into elasticities with respect to the net-of-tax rate on wealth. This allows for comparing magnitudes across different specifications, and for comparing the quasi-experimental estimations to the structural elasticity estimations presented later. To calculate elasticities, we relate the ITT effect on wealth to the *expected* change in the net-of-tax rate on wealth. The expected tax change accounts for the variation in tax treatment driven by the churn in and out of treatment and control groups after the reform. Specifically, we define

$$\varepsilon = \frac{\mathsf{E}\left[\hat{\beta}_t^{ITT}\right]}{\mathsf{E}\left[\Delta \log\left(1 - \tau_{it}\right) \mid T\right] - \mathsf{E}\left[\Delta \log\left(1 - \tau_{it}\right) \mid C\right]}, \tag{3}$$

where $\mathsf{E}\left[.\right]$ is the expectations operator, $\hat{\beta}_t^{ITT}$ is the ITT coefficient in year t, and τ_{it} is the wealth tax rate on household i at time t. The denominator represents the expected log-change in the net-of-tax rate for the treatment group T relative to the control group C. The elasticity defined in (3) is based on the average effect over the post-reform window.

It is useful to walk through the elasticity calculation by way of a concrete example. Consider the couples DD. Households in the treatment group — couples located in the exempted wealth range before the reform — may experience one of three treatment states after the reform: they may stay in the exempted range for couples (thus experiencing a tax cut of 2.2pp), they may move above the new exemption threshold for couples or become singles above the threshold applying to singles (thus experiencing a tax cut of 1.2pp), or their wealth may fall below the old exemption threshold (where the tax cut is zero). The expected tax change for the treatment group accounts for this churn and is therefore smaller than the mechanical tax change of 2.2pp. Similarly, the expected change for the control group is different from 1.2pp as they may fall below the exemption threshold or become couples in the exempted range. By scaling the ITT effect using the expected tax change, the resulting parameter represents a TOT elasticity. An alternative approach would be to divide the TOT effect by the mechanical, reform-induced tax change for the treatment group relative to the control group (i.e., a tax cut of 2.2pp relative to 1.2pp). These two approaches are not equivalent in our setting, but in practice they give almost the same result.[21]

The elasticity in equation (3) is defined with respect to the net-of-tax rate on wealth, $1 - \tau$. An

---

[20]In general, the presence of pre-trends in the raw data suggests that the unadjusted difference-in-differences estimates will be biased by non-reform confounders. Controlling for such confounders by extrapolating the pre-event trend is valid only under the assumption that the post-event behavior of the confound can be inferred from the pre-event trend (see e.g., Freyaldenhoven, Hansen, and Shapiro 2019; Roth 2019). We discuss the validity of this assumption when presenting the results from the specifications where a pre-trend adjustment is necessary.

[21]The reason why these approaches are not equivalent is that the treatment category (a tax cut of 2.2pp) and control category (a tax cut of 1.2pp) are not mutually exclusive, because some of the churn is driven by households in either group falling below the old exemption threshold and thus receiving zero treatment. These movements create bias in the elasticity calculated directly from the TOT effects. In practice, this has a small effect on the elasticity as the movements from the treatment and control groups create offsetting biases of similar magnitudes.

alternative elasticity can be defined with respect to the net-of-tax rate rate of return to wealth, i.e. $(1 - \tau) R - 1$ where R is the gross rate of return on wealth. This is a more meaningful elasticity concept in the context of wealth taxation, but it requires us to make an assumption about the (unobserved) rate of return R. We consider both types of elasticity calculations.

III.B   Descriptive Statistics

Before investigating behavioral responses to the wealth tax, we present descriptive statistics in Table I. The table shows means of wealth, income and demographics for households in the full population (column 1) and for households in our treatment and control groups (columns 2-6). As discussed above, the assignment of treatment status is based on pre-reform variables and restricts attention to households whose status stays constant during 1982-88. The statistics in the table are based on pooled data between 1982-1988. The table reports both taxable wealth and total market value wealth, the latter being computed as described in section II.B.

The following points are worth highlighting. First, our population of interest is very different from the general population. The treatment and control groups consist of households who are wealthier, older, and more self-employed than the average household in the population. They also hold a larger share of their wealth in equities and a somewhat smaller share in housing.[22] Second, market value wealth is in general larger than taxable wealth, but less so in our estimation sample of wealthy taxpayers. This is primarily because pension wealth (which is not part of taxable wealth) weigh less heavily in the portfolio of the wealthy. Third, the difference between labor income and total income (including capital income) is relatively small in the full population, but large among the wealthy who receive most of their income in the form of asset returns.

Finally, there are some noticeable differences in pre-reform means for the treatment and control groups. This is to be expected given how these groups are defined. The couples DD — especially when we compare couples and singles within the exempted range — is much more balanced than the ceiling DD where we compare households who are bound and unbound by the tax ceiling. Those bound by the ceiling are much wealthier, hold more of their wealth in equities and less in housing, and are more self-employed than the treated group of unbound taxpayers. This lack of balance could be a concern for the ceiling DD approach, but only insofar as it affects the credibility of the parallel trends assumption.

---

[22]Like a number of countries, Denmark subsidizes homeownership through a mortgage interest deduction. The value of this deduction used to be larger at higher incomes, but an income tax reform enacted in 1987 lowered the deduction and made it uniform across income tax brackets. Gruber, Jensen, and Kleven (2018) investigate behavioral responses to this change in the mortgage interest deduction and find zero effect at the extensive margin of housing investments. In any case, the mortgage interest deduction is not very important among the wealthy population studied here (as they have little or no mortgage debt) and the empirical strategies used in this paper rely on different tax variation (within the group of wealthy people) than the variation created by the 1987 income tax reform.

III.C Couples DD: Responses by the Moderately Wealthy

We first consider behavioral responses by the moderately wealthy using the couples DD strategy described in section III.A. This strategy exploits 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) singles located in the exempted range and (ii) couples located below the exempted range (within the top 5%). The advantage of the first specification is that the comparison groups have the same level of household wealth, but the disadvantage is that both groups are treated to some degree. While couples in the exempted range have their tax rate cut to zero, singles in this range have their tax rate cut to 1%. The second specification is based on an untreated comparison group and therefore larger identifying variation, but it will require us to deal with differential trends in different parts of the wealth distribution. Under both specifications, the assignment of treatment status (from marital status and wealth bracket) is based on the observed values in pre-reform years (1982-88 in the baseline).

Figure IV provides evidence based on using singles as the comparison group. Panel A shows the time series of log taxable wealth for couples in the exempted wealth range (red dots) and singles in the same range (black squares) between 1980-1996, with both series normalized to zero in the year before the 1989 reform. Panel B shows the differences between these two series. Two key insights emerge from the figure. First, the two groups are on similar trends prior to the reform. While there are some differences in the early 1980s, the trends are almost perfectly parallel in the five years leading up to the reform. Second, the two series begin to diverge immediately after the reform. The difference in wealth levels between the two groups is gradually increasing over time, consistent with a change in the savings rate. Overall, this figure provides clear evidence of behavioral responses to the reduction in wealth taxation.

The results in Figure IV correspond to reduced-form or intention-to-treat (ITT) effects. The comparison groups are based on pre-reform treatment status, which is not perfectly persistent over time and this attenuates the observed effects.[23] Figure V investigates the persistence of treatment status and converts the intention-to-treat (ITT) effects into treatment-on-the-treated (TOT) effects. Panel A documents the degree of persistence by showing the fraction treated in each of the two groups over time. By construction, couples within the exempted range are 100% treated in the six pre-reform years, while singles within this range are 0% treated in those years. After the reform, taxpayers may switch status due to changes in relationship status (through marriage, divorce or widowhood) or changes in their wealth bracket. The figure shows that the "control group" (singles) is very persistent, reflecting the fact that it is unusual in this predominantly older sample to become married and at the same time stay within the same wealth bracket. The "treatment group" (couples) is less persistent due to wealth changes and spousal death or separation. Eight years after the reform, the difference in treatment intensity is about 50%. Panel B converts the

---

[23]In terms of the regression framework in section III.A, the estimates in Panel B correspond to the reduced-form version of (1) where log wealth is regressed directly on the instruments, $Year_{j=t} \cdot Treat_i^{pre}$.

ITT series into a TOT series by dividing the former with the differences in treatment intensity from Panel A (Wald estimator).[24] This implies that the dynamically growing effect on taxable wealth is enhanced, an implication of the gradual reduction in persistence. The TOT effect on log wealth is equal to 0.186 in the last post-reform year (1996), i.e. an increase of about 18% 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 counter-factual 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. We will assume an annual rate of return equal to 5%. This falls within the range of existing estimates of wealth returns at the top of the distribution (see e.g., Fagereng, Guiso, Malacrino, and Pistaferri 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 wealth tax cuts. The details of this calculation are provided in section B and the results presented in Figure A.IV of the online appendix.

Figure A.IV shows the series of total effects (blue squares) and behavioral effects (green trian-gles), with the differences between the two being the mechanical effects. We see that the mechani-cal effects are small, an effect on log wealth of 0.02 — or 11% of the total effect — after eight years. The reason why the mechanical effects are 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 a number of robustness checks. First, Figure A.V investigates if our results are sensitive to defining comparison groups based on outcomes (marital status and wealth levels) in specific pre-reform years. The figure shows the evolution of log taxable wealth in the two comparison groups — couples and singles within the exempted wealth range — when using different pre-reform windows to define treatment status: 1980-88, 1982-88 (baseline), 1984-88, and 1986-88. The figure shows that the main implication of using a longer treatment window is to make the pre-trends more parallel, especially in the early 1980s. Reassuringly, the results are similar across the alternative specifications. In all four panels, the wealth trends are almost parallel

---

[24] In regression terms, the TOT series in Panel B correspond to the IV estimates from equation (1) where the concurrent year-by-treatment dummies, $Year_{j=t} \cdot Treat_{it}$, are instrumented using $Year_{j=t} \cdot Treat_i^{pre}$.

in the last five years before reform, and the divergence in wealth is about the same after reform. Based on this graph, we conclude that our results are not sensitive to the length of the treatment window.

Second, Figure A.VI provides a set of placebo tests assuming that the reform happened in earlier years: 1983, 1984, 1985, and 1986. The comparison groups are still couples and singles within the exempted wealth range, but the group assignment is based on outcomes prior to the placebo reform rather than the actual reform. When studying placebo reforms in the early 1980s, we have to shorten the window used to assign treatment status to three years (e.g., 1980-82 for the 1983 placebo reform). The figure shows ITT and TOT series for each of the four placebo reforms. The patterns lend further support to our interpretation of the data. In three out of four panels, there is a precisely estimated zero effect of the placebo reform in 1988, the last year before the actual reform starts affecting the patterns. Only the 1984 placebo reform appears to generate an effect. However, this is due to the fact that 1983 is an outlier year (see also Figure IV above), and so normalizing the series to zero in 1983 (as we do when the reform is assumed to happen in 1984) creates an illusory effect.

Third, in Figures A.VII-A.IX we consider the approach in which the comparison group consists of couples below the exempted wealth range. In this case, the comparison group is completely unaffected by the tax cuts and the experiment is therefore larger. The figures are constructed in the same way as the corresponding figures for the previous strategy. Consider first the raw series of log taxable wealth in Panel A of Figure A.VII. The graph shows that couples in the exempted range are on a flatter trend than those below the exempted range in the years before to the 1980 reform, while they are on the same and subsequently steeper trend in the years after the reform. This change in relative trends is consistent with an effect of the tax cuts on wealth accumulation. At the same time, we note that the timing of the trend break does not coincide exactly with the tax reform, but happens a little too early. This points to the possibility of confounding shocks that have different impacts on different parts of the wealth distribution. Such confounders may bias the estimated treatment effect and we therefore have less confidence in this specification.

These concerns notwithstanding, it is useful to turn the raw wealth series in Panel A of Figure A.VII into difference-in-differences estimates that are comparable to those obtained from the previous strategy. This requires us to adjust for the differential pre-trends of the treatment and control groups. This is done in Panel B using specification (2) discussed above. The dashed-grey series show the raw differences between the treatment and control groups, while the solid-red series show the pre-trend adjusted differences between the two groups. The next figure documents persistence and presents the series of TOT effects. We see that the TOT effect builds up gradually and is equal to 0.265 log points after eight years. This is much larger than the treatment effect obtained from the previous strategy, but recall that the underlying tax variation is also much larger here. In fact, as we will show later, the estimates are similar in elasticity terms.

The appendix presents results from one additional specification, a cross between the previous two. Instead of using singles in the exempted range or couples below the exempted range as

16

controls, this strategy uses singles below the exempted range as controls. The rationale behind this strategy is that a couple with household wealth W has the same wealth *per capita* as a single person with wealth $W/2$. Therefore, this strategy compares couples in the exempted range (those between the singles' threshold and twice the singles' threshold) to singles in the same *per capita* range (those between half the singles' threshold and the singles' threshold). The results are presented in Figures A.X-A.XII and they look quite compelling. The raw pre-trends are almost perfectly parallel in the eight years before the reform, followed by a clear divergence in the eight years after the reform. Again, we note that the treatment effect starts a little too early, raising concerns about confounders that are not present in our main strategy of comparing couples and singles in the same range of households wealth.

To conclude, we have presented findings from several difference-in-differences specifications that take advantage of the doubling of the exemption threshold for married couples. We have compared these treated couples to different control groups, either singles or other couples within or below the exempted wealth range. Taken together, these specifications provide evidence of quite sizeable taxable wealth responses to wealth taxation.[25]

III.D   *Ceiling DD: Responses by the Very Wealthy*

We now turn to the behavioral responses of the very wealthiest taxpayers using the ceiling DD. This strategy consists in comparing taxpayers who are unbound by the tax ceiling (treatments) to taxpayers who are bound by the tax ceiling (controls). The treatment group experienced a reduction in the marginal wealth tax rate from 2.2% to 1%, while the control group experienced no change in their marginal tax rate. Because the ceiling starts binding only at the very top of the wealth distribution (as shown in Figure A.I), we compare bound and unbound taxpayers within the top 1% of the wealth distribution. We assign taxpayers to treatment and control groups using six pre-reform years, thus dropping taxpayers who frequently switch ceiling status. To further increase the persistence of treatment status, we also drop observations who are only marginally bound by the tax ceiling. Specifically, the bound group includes those whose wealth tax liability would have to fall by at least 20% for them to become unbound, but the results are robust to

---

[25]While we focus on the impact of the wealth tax reform on wealth accumulation, it is worth noting that the reform also changed the marriage incentives. The fact that the nominal exemption threshold was the same for singles and couples prior to the 1989 reform created a significant marriage penalty for wealthy individuals. By doubling the exemption threshold for couples, the reform eliminated this penalty and strengthened the incentives to marry at the top of the distribution. Figure A.XIII in the appendix provides evidence on the potential marriage responses. It shows the evolution of marriage rates in different wealth quantiles above and below the pre-reform threshold: the top 1% and top 2.5-1% percentiles are above the threshold (where the incentive to marry becomes stronger after the reform), while the top 5-2.5% and top 10-5% percentiles are below the threshold (where the incentive to marry is unchanged). Interestingly, while the four groups are on parallel trends before the reform, the marriage rate increases in the higher percentiles relative to the lower percentiles after the reform. This is consistent with a behavioral response to the marriage penalty. However, the figure also highlights a potential caveat to this interpretation by showing that there is a general fanning-out across the four groups. The top 1% increases relative to the top 2.5-1% (even though they are both treated), and the top 5-2.5% increases relative to the top 10-5% (even though they are both untreated). This suggests the presence of confounding effects on marriage. Therefore, while the patterns in Figure A.XIII are intriguing and consistent with a marriage response, the evidence is not conclusive.

alternative cuts.

Figures VI-A.XIV are constructed in the same way as the preceding figures for the couples DD. In Panel A of Figure VI, we see that the treatment group (unbound) is on a flatter trend than the control group (bound) during the pre-reform period. This pattern reverses just after reform, and the treatment group is on a considerably steeper trend during the entire post-reform period. The switch from a flatter to a steeper trend around the reform provides strong evidence of behavioral responses to the reform. Panel B of Figure VI shows the differenced series, with the raw differences in dashed grey and the pre-trend adjusted differences in solid red. The pre-trend adjustment is based on four pre-reform years using specification (2) described above. The adjusted DD series looks quite compelling: it features almost perfectly parallel pre-trends in the decade leading up to the reform combined with a clear and growing divergence in the eight years following the reform.

Figure VII documents the persistence of ceiling status and converts the ITT estimates into TOT estimates. As shown in Panel A, the fraction treated in the treatment group is 100% during the pre-reform years (by construction) and falls only slightly after the reform, while the fraction treated in the control group starts from 0% and increases gradually after the reform. The control group is less persistent in this case, because it is more common for those bound by the ceiling to become unbound due to wealth and income shocks than it is for unbound taxpayers to become bound. In the last year of the post-reform period, the difference in treatment intensity is a little more than 50%. When converting the ITT effects into TOT effects in Panel B, we estimate a treatment effect on log taxable wealth equal to 0.312, an increase in wealth of about 30%.

Figure A.XIV in the appendix splits the total effect on wealth into behavioral and mechanical effects. The method is the same as for the couples DD: we calculate annual tax savings for the treatment group using a measure of counterfactual wealth and simulate cumulative tax savings assuming an annual return of 5%. The figure shows that the mechanical effects are larger for the ceiling DD than for the couples DD. The main explanation is that the ceiling approach captures responses by wealthier households, i.e. households located farther from the exemption threshold. As a result, the reform-induced change in the average after-tax return is larger in this sample. We find that the mechanical effect on log taxable wealth equals 0.068 after eight years, corresponding to 22% of the total effect.

The appendix provides robustness checks similar to those shown for the couples DD. Figure A.XV explores the implications of using different pre-reform windows to define treatment status. It is apparent that, for the ceiling DD, specifying a longer treatment window ensures more parallel pre-trends. Still, even though the pre-trends differ across specifications, they all show clear evidence of behavioral responses: the treatment group is on a flatter trend before the reform and a steeper trend after the reform. The specifications with shorter treatment windows (in particular, the 1986-88 window in Panel D) imply *larger* behavioral responses than those reported above once we adjust for pre-trends. We prefer the specification with a longer treatment window, because it ensures better pre-trends in the raw data and is relatively conservative.

Figure A.XVI shows placebo tests based on assuming the reform happened in earlier years. The

18

analysis is done in the same way as the corresponding analysis for the couples DD. Overall, the placebo tests look quite compelling. In all four panels, there is no significant effect of the placebo reform in 1988, the last year before the actual reform.

III.E   Summary of DD Estimates

Table II shows DD estimates of taxable wealth responses to the 1989 reform and converts these estimates into elasticities. The columns refer to the different quasi-experimental specifications: the couples DD and the ceiling DD, with and without adjusting for pre-trends. The estimates without pre-trend adjustment in columns (1), (3), and (5) are based on specification (1), while the estimates with pre-trend adjustment in columns (2), (4), and (6) are based on specification (2). For each specification, we show both ITT and TOT effects. As described in section III.A, the ITT effects are obtained from a reduced-form specification in which log wealth is regressed directly on the instruments (constructed from pre-reform behavior), while the TOT effects are obtained from an IV specification. We report both the average effect over the post-reform window (1989-96) and the effect in the last post-reform year (1996). While it is standard to show the average effect, the last-year effect is arguably more informative for a dynamic outcome like the stock of wealth. Still, the "last-year effect" shown here does not correspond to the long-run effect as we show in the structural analysis presented later. Finally, the table converts the average effects on log wealth into elasticities using the definition in (3). We show elasticities with respect to the net-of-tax rate, $1 - \tau$, and with respect to the net-of-tax rate of return, $(1 - \tau) R - 1$, assuming a gross rate of return of $R = 1.05$.

The absolute effects on log wealth vary considerably across the different strategies/samples, but so does the underlying tax variation driving them. For example, if we consider TOT effects adjusted for pre-trends, the average effect equals 0.060 log points when comparing couples and singles within the exempted range, while it equals 0.135 log points when comparing couples within and below the exempted range. The last-year effects equal 0.133 log points and 0.277 log points, respectively. However, because the tax variation in the second strategy is larger than in the first strategy, the implied elasticities of taxable wealth are about the same. In both cases, the elasticity with respect to the after-tax rate of return is just above 0.2. Turning to the ceiling DD, the effects are larger both in absolute terms and in elasticity terms, but here we are considering a different population of very wealthy taxpayers. The TOT effect on their wealth equals 0.171 log points on average and 0.312 log points in the last year. The elasticity with respect to the after-tax rate of return is about 0.4.[26]

---

[26]Tables A.I-A.II in the appendix show heterogeneity in the estimated wealth responses by age (below and above the median age in each estimation sample). It would be natural to also study heterogeneity by the presence and number of children, especially considering that wealth responses by wealthy taxpayers may be partly motivated by bequest motives. Such an analysis is not feasible, however, because parents cannot be linked to children born before 1960 in the Danish data. Most people liable to pay wealth tax (i.e., older people) in the 1980s and 1990s would have had children before 1960.

IV.A   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 regarding the wealthy. First, as mentioned earlier, wealthy people tend to be older people. 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 large amounts of wealth. This will be documented in detail in the next section.

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 and Stantcheva 2018 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

$$\frac{\sigma}{\sigma-1}\sum_{t=0}^{T}\delta^t\left(c_t\right)^{\frac{\sigma-1}{\sigma}}+\delta^T V\left(W_{T+1}\right),\tag{4}$$

where c_t is consumption in period t, W_{T+1} is wealth at the end of life (bequests), σ is the elasticity of intertemporal substitution (EIS), and δ is the discount factor. To capture utility of bequests, we

[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 and 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.

adopt the following parameterization

$$V\left(W_{T+1}\right) = A\frac{\alpha}{\alpha - 1}\left(\frac{W_{T+1}}{A}\right)^{\frac{\alpha-1}{\alpha}}, \tag{5}$$

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 and Werning (2010), Piketty and 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

$$
\begin{aligned}
c_t &= y_t + RW_t - \tau R\left(W_t - \bar{W}\right) - W_{t+1} \\
&= y_t + \left(1 - \tau\right) RW_t + \tau R\bar{W} - W_{t+1},
\end{aligned} \tag{6}
$$

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 (6) is a "virtual income" representation: it writes the budget as if the net-of-tax return equals $\left(1 - \tau\right) R$ 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

$$\sum_{t=0}^{T}\frac{c_t}{\left(\left(1 - \tau\right) R\right)^t} + \frac{W_{T+1}}{\left(\left(1 - \tau\right) R\right)^T} = \sum_{t=0}^{T}\frac{y_t}{\left(\left(1 - \tau\right) R\right)^t} + \sum_{t=1}^{T}\frac{\tau R\bar{W}}{\left(\left(1 - \tau\right) R\right)^t} + W_0^n, \tag{7}$$

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

Households maximize lifetime utility (4)-(5) subject to the lifetime budget constraint (7) with respect to consumption and bequests. The first-order conditions for c_t and c_{t+1} yield the standard Euler equation,

$$c_{t+1} = \left(\delta \left(1 - \tau\right) R\right)^{\sigma} c_t, \tag{8}$$

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

$$W_{T+1} = Ac_T^{\alpha/\sigma}. \tag{9}$$

The solution to the model is described by the lifetime budget (7), the Euler equations (8) for all t, and the bequest condition (9). 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.

$$c_t = (\delta (1-\tau) R)^{t\sigma} c_0, \tag{10}$$

$$W_{T+1} = A (\delta (1-\tau) R)^{T\alpha} c_0^{\alpha/\sigma}. \tag{11}$$

Inserting these conditions into (7), we can express the choice of c_0 as

$$\sum_{t=0}^{T} q_t \cdot c_0 + q_b \cdot c_0^{\alpha/\sigma} = \sum_{t=0}^{T} \frac{y_t}{((1-\tau) R)^t} + \sum_{t=1}^{T} \frac{\tau R \bar{W}}{((1-\tau) R)^t} + W_0^n, \tag{12}$$

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.

IV.B   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 it 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.

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 (7) 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

$$\frac{dW_1}{d\left(1-\tau\right)}\frac{1-\tau}{W_0} = \sigma \cdot \left\{ \frac{\sum_{t=0}^{T} t q_t}{\sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\alpha/\sigma-1}} \frac{c_0}{W_0} \right\} + \alpha \cdot \left\{ \frac{T q_b}{\sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\alpha/\sigma-1}} \frac{c_0^{\alpha/\sigma}}{W_0} \right\}$$
$$+ \frac{dW_0^C}{d\left(1-\tau\right)}\frac{1-\tau}{W_0} \cdot \left\{ \frac{1}{\sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\alpha/\sigma-1}} \right\}, \tag{13}$$

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 C. $\square$

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.[29] Second, there is a substitution effect on bequests. A larger 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*

$$\frac{dW_1}{d\left(1-\tau\right)}\frac{1-\tau}{W_0} = \sigma \cdot \left\{ \frac{\sum_{t=0}^{T} t q_t + T A q_T}{\sum_{t=0}^{T} q_t + A q_T} \frac{c_0}{W_0} \right\} + \frac{dW_0^C}{d\left(1-\tau\right)}\frac{1-\tau}{W_0} \cdot \left\{ \frac{1}{\sum_{t=0}^{T} q_t + A q_T} \right\}, \tag{14}$$

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

---

[29]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.

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 (13). □

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 $\sigma = \alpha$. However, we do not rely on the assumption of $\sigma = \alpha$ in the structural estimation presented below, but identify these two parameters separately.

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 (6), 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

$$W_t = [(1 - \tau) R]^{t-1} W_0^n + \sum_{j=0}^{t-1} (y_j - c_j) [(1 - \tau) R]^{t-1-j} + \sum_{j=1}^{t-1} \tau R \bar{W} [(1 - \tau) R]^{t-1-j}. \tag{15}$$

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*

$$\frac{dW_t / W_t}{d(1 - \tau)/(1 - \tau)} = dM_t + dB_t, \tag{16}$$

where dM_t is a mechanical effect given by

$$
\begin{aligned}
dM_t &= (t-1)\,[(1 - \tau) R]^{t-1} \frac{W_0^n}{W_t} + \sum_{j=0}^{t-1} (t - 1 - j)\,[(1 - \tau) R]^{t-1-j} \frac{y_j - c_j}{W_t} \\
&\quad - \sum_{j=1}^{t-1} \left(1 - \frac{\tau(t - 1 - j)}{1 - \tau} \right) [(1 - \tau) R]^{t-j} \frac{\bar{W}}{W_t},
\end{aligned}
\tag{17}
$$

and dB_t is a behavioral effect given by

$$dB_t = \sum_{j=0}^{t-1} \frac{q_j}{[(1 - \tau) R]^{1-t}} \left\{ \frac{dW_1 / W_0}{d(1 - \tau)/(1 - \tau)} \frac{W_0}{W_t} - j\sigma \frac{c_0}{W_t} \right\}. \tag{18}$$

24

In equation (18), 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 C. □

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 wealth over time, holding consumption and 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 (18). 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 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.

Finally, given our objective of estimating elasticities that can be used to assess optimal capital taxation, it is useful to relate our analysis to classic results calling for zero capital taxation. The result of Chamley (1986) and Judd (1985) that capital should be untaxed in steady state arises because, in their models, the steady state elasticity of capital supply is infinite (see e.g., Auerbach and Hines 2002; Piketty and Saez 2013; Saez and Stantcheva 2018). While we have characterized finite elasticities of capital supply, note that our model nests a version of the Chamley-Judd framework when assuming infinite horizons ($T \to \infty$) and no utility of residual wealth ($A \to 0$). The elasticity in steady state is then infinite and the optimal capital tax rate is zero. It would be next to impossible to evaluate the infinite-elasticity prediction in the data, because the steady state equilibrium is not observed. Moreover, as shown by Straub and Werning (2019), the steady state prescription may not be a very relevant as convergence to the steady state can be extremely slow in the Chamley-Judd framework, potentially taking centuries. We take a different approach, matching our model with a finite horizon and utility of residual wealth to estimated responses over the short-medium

run in order to estimate a (finite) elasticity in the long run. This elasticity can be used to assess optimal capital taxation within this general class of models (including e.g., Saez and Stantcheva 2018), but does not by itself rule out the Chamley-Judd steady state prediction.

V Connecting Theory and Evidence: Long-Run Effects

V.A Empirical Lifecycle Profiles of Wealth

To calibrate our model, we start by studying the empirical lifecycle profiles of wealth at the top of the distribution. Besides informing our calibration exercise, these lifecycle profiles will provide 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 VIII 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).[30] 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 $\omega_{iat} \equiv \log(W_{iat}) - \mathsf{E}[\log(W_{iat}) \mid t]$. 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.[31] 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.[32]

The following key findings emerge from Figure VIII. 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.[33] At the

---

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

[31]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 $\mathsf{E}\{\mathsf{E}[\omega_{iat} \mid a, t] \mid a\}$, 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.

[32]However, we do condition the sample on being in the top $p\%$ for at least three years so as to reduce noise from transitory wealth shocks.

[33]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.

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. There is no draw-down of wealth until after the age of 80, 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.XVII in the online appendix shows the same type of graph, 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) to rationalize these empirical patterns.

Have these lifecycle patterns changed over time? Specifically, have they changed from before to after the 1989 reform? Figure A.XVIII investigates this question by comparing the lifecycle profiles of wealth before the reform (1980-88) and after the reform (1989-96). 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.XIX 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 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.

V.B Calibration

To study the long-run effects of wealth taxes on the wealthy, we calibrate the model from section IV 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 Appendix Figure A.XVII 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 (11) 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. Specifically, when setting $R = 1.05$, the calibrated value of δ is around 0.97.[34] 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 are matched to the treatment-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,

---

[34] 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, $R = 1.05$ falls well within the range of estimated returns at the top of the wealth distribution (see Fagereng, Guiso, Malacrino, and Pistaferri 2016).

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.

$$\mathcal{L}\left(\sigma,\alpha\right) = \sum_{t=1}^{8}\left[\text{TOT}_{t} - \Delta\log W_{t}\left(\sigma,\alpha\right)\right]^{2}, \tag{19}$$

where TOT_{t} is the estimated treatment effect in year t and $\Delta\log W_{t}\left(\sigma,\alpha\right)$ 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 household.

Although we could estimate σ and α based solely on equation (19), 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 (19) 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.

V.C Simulating the Long-Run Effects of Wealth Taxation

We first consider the effects of wealth taxation on the moderately wealthy. That is, we calibrate the model to the estimates from the couples DD in which we compare couples and singles in the exempted range of the wealth distribution (between the 98th and 99th percentiles). The results are

presented in Figure IX. Panel A shows three wealth paths: the observed 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 observed wealth path. The simulated wealth path after the reform is based on a reduction of the wealth tax rate by 1 percentage point, corresponding to the differential tax cut between the comparison 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 in which we compare 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 25 years and then stabilizes. At the end of life, wealth is 30% higher than it would have been absent the reform. Second, while this effect seems large, the underlying tax incentive driving it is also large. Defining the net-of-tax return as $(1 - \tau) R - 1$, the percentage change in the return is equal to $-\Delta\tau \cdot R / ((1 - \tau) R - 1) = 39\%$. Therefore, the long-run elasticity of wealth with respect to the net-of-tax rate of return is 0.77. Third, the mechanical effect becomes increasingly important over time due to the compounding effects of a larger net-of-tax rate of return. After 30 years, the mechanical effect accounts for more than one-third of the total effect.

As a robustness check, Figure A.XX in appendix shows a calibration based on the couples DD in which we compare couples within the exempted range to those below. This comparison captures a larger experiment — a reduction of the wealth tax rate by 2.2 percentage points — and we therefore expect the effects to be larger. Indeed, the graph shows that the long-run effect on the stock of wealth equals 50%, much larger than in the baseline specification. Relating this effect to the change in the net-of-tax rate of return, $-\Delta\tau \cdot R / ((1 - \tau) R - 1) = 86\%$, the long-run elasticity of wealth is close to 0.6. It is reassuring that the two different couples DD strategies imply roughly similar long-run elasticities of wealth, despite being based on very different tax variation and reduced-form effects.

We now turn to the effects of wealth taxation on the very wealthy, calibrating the model to the ceiling DD and taxpayers within the top 1%. In this case, the experiment is a reduction in the marginal wealth tax by 1.45 percentage points, corresponding to the average reduction across treated taxpayers.[35] The results are presented in Figure X. As before, the model does a good job of fitting the baseline wealth path to the observed path (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 65%. Given a percentage change in the net-of-tax return equal to $-\Delta\tau \cdot R / ((1 - \tau) R - 1) = 57\%$, the implied long-run elasticity of wealth equals 1.15. Slightly less than half of this effect is mechanical, while the rest is behavioral.

---

[35]While most treated taxpayers in the ceiling DD see their marginal tax rate reduced by 1.2 percentage points, some of them benefit from the increase in the exemption threshold for couples (to the 99.3rd wealth percentile) and therefore see their marginal tax rate reduced by 2.2 percentage points.

Table III summarizes the simulation results and shows the "structural primitives" underlying each calibration. The table also investigates the robustness of the results to the assumed rate of return R. The following points are worth noting. First, the total effect on wealth is very robust to the assumed rate of return R, but it affects the composition into mechanical and behavioral effects. When the return is higher, more of the effect is mechanical and less is behavioral. Second, the rate of return has a big impact on the reduced-form elasticity with respect to $(1-\tau)R-1$, but this is a trivial implication of changing the denominator of the elasticity. In some sense, this highlights an issue with focusing on reduced-form elasticities in this context. Third, while the long-run elasticity of taxable wealth equals 0.58-0.77 for the moderately wealthy and 1.15 for the very wealthy (under $R = 1.05$), the underlying structural elasticities — the EIS σ and the bequest elasticity α — are larger in magnitude. For example, the EIS is around 2 for the moderately wealthy and even larger for the very wealthy. This is much larger than existing estimates of this parameter (see e.g., Best, Cloyne, Ilzetzki, and Kleven 2019).[36] When comparing the structural elasticities obtained here to those in the existing literature, two key aspects must be kept in mind. One is that we are considering behavioral responses by very wealthy households, something the existing literature has not been able to do. The other is that we are considering effects on *taxable wealth*, which will include both real savings responses and any evasion or avoidance responses. The larger elasticities among the very wealthy than among the moderately wealthy are consistent with larger avoidance at the top. Therefore, the elasticities we estimate represent upper bounds on real wealth accumulation responses.

VI 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 and Stantcheva (2018) 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

---

[36]Figure A.XXI 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 shorter-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 quasi-experimental moments is in itself informative of the relative magnitudes of α and σ. In this particular case, however, 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.

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 is growing evidence on migration responses to *labor* income taxes at the top (see Kleven, Landais, Muñoz, and Stantcheva 2019 for a review), there is virtually no evidence on migration responses to capital or wealth taxes. 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.

Finally, 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.

UNIVERSITY OF COPENHAGEN AND CEBI
THE THINK TANK DEA
PRINCETON UNIVERSITY AND NBER
UC BERKELEY AND NBER

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at the Quarterly Journal of Economics online.

REFERENCES

ALSTADSÆTER, ANNETTE, NIELS JOHANNESEN, AND GABRIEL ZUCMAN (2019): "Tax Evasion and Inequality," *American Economic Review*, 109(6), 2073–2103. 4

ALVAREDO, FACUNDO, LUCAS CHANCEL, THOMAS PIKETTY, EMMANUEL SAEZ, AND GABRIEL ZUCMAN (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. 21

——— (1990): "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *Economic Journal*, 100(401), 464–477. 21

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

AUERBACH, ALAN J., AND JAMES R. HINES (2002): "Taxation and Economic Efficiency," in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. S. Feldstein, vol. 3. Elsevier: Amsterdam, NL. 25

BENHABIB, JESS, AND ALBERTO BISIN (2018): "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 56(4), 1261–91. 5

BEST, MICHAEL C., JAMES CLOYNE, ETHAN ILZETZKI, AND HENRIK J. KLEVEN (2019): "Estimating the Elasticity of Intertemporal Substitution Using Mortgage Notches," *Review of Economic Studies*, forthcoming. 31

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

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

CARROLL, CHRISTOPHER D. (2002): "Why Do the Rich Save So Much?," in *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, ed. by J. B. Slemrod. Harvard University Press: Cambridge, MA. 5, 20

CHAMLEY, CHRISTOPHE (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54(3), 607–622. 4, 25

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

CHETTY, RAJ, JOHN N. FRIEDMAN, TORE OLSEN, AND LUIGI PISTAFERRI (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. 50

DE NARDI, MARIACRISTINA (2004): "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 71(3), 743–768. 5, 20

FAGERENG, ANDREAS, LUIGI GUISO, DAVIDE MALACRINO, AND LUIGI PISTAFERRI (2016): "Heterogeneity and Persistence in Returns to Wealth," NBER Working Paper No. 22822. 15, 28

FARHI, EMMANUEL, AND IVÁN WERNING (2010): "Progressive Estate Taxation," *Quarterly Journal of Economics*, 125(2), 635–673. 21

FREYALDENHOVEN, SIMON, CHRISTIAN HANSEN, AND JESSE M. SHAPIRO (2019): "Pre-event Trends in the Panel Event-study Design," *American Economic Review*, forthcoming. 12

GRUBER, JONATHAN, AMALIE S. JENSEN, AND HENRIK J. KLEVEN (2018): "Do People Respond to the Mortgage Interest Deduction? Quasi-Experimental Evidence from Denmark," NBER Working Paper No. 23600. 13

JUDD, KENNETH L. (1985): "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28(1), 59–83. 4, 25

KAPLOW, LOUIS (2011): "Utility from Accumulation," *National Tax Association Proceedings*, pp. 102nd Annual Conference on Taxation, 189–194. 5

KLEVEN, HENRIK J. (2014): "How Can Scandinavians Tax So Much?," *Journal of Economic Perspectives*, 28(4), 77–98. 20

——— (2016): "Bunching," *Annual Review of Economics*, 8(1), 435–464. 9, 50

KLEVEN, HENRIK J., MARTIN KNUDSEN, CLAUS T. KREINER, SØREN PEDERSEN, AND EMMANUEL SAEZ (2011): "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark," *Econometrica*, 79(3), 651–692. 51

KLEVEN, HENRIK J., CAMILLE LANDAIS, MATHILDE MUÑOZ, AND STEFANIE STANTCHEVA (2019): "Taxation and Migration: Evidence and Policy Implications," *Journal of Economic Perspectives*, forthcoming. 32

KLEVEN, HENRIK J., AND ESBEN A. SCHULTZ (2014): "Estimating Taxable Income Responses Using Danish Tax Reforms," *American Economic Journal: Economic Policy*, 6(4), 271–301. 20

KLEVEN, HENRIK J., AND MAZHAR WASEEM (2013): "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics*, 128(2), 669–723. 50

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

——— (2013a): "Incentive Effects of Inheritances and Optimal Estate Taxation," *American Economic Review, Papers and Proceedings*, 103(3), 472–477. 5, 21

——— (2013b): "Taxation of Intergenerational Transfers and Wealth," in *Handbook of Public Economics*, ed. by A. J. Auerbach, R. Chetty, M. S. Feldstein, and E. Saez, vol. 5. Elsevier: Amsterdam, NL. 5

KREINER, CLAUS T., DAVID D. LASSEN, AND SØREN LETH-PETERSEN (2015): "Measuring the Accuracy of Survey Responses Using Administrative Register Data: Evidence from Denmark," in *Improving the Measurement of Household Consumption Expenditures*, ed. by C. D. Carroll, T. F. Crossley, and J. Sabelhaus. 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

LOVE, DAVID A., MICHAEL G. PALUMBO, AND PAUL A. SMITH (2009): "The trajectory of wealth in retirement," *Journal of Public Economics*, 93, 191–208. 5

OECD (1988): *Taxation of Net Wealth, Capital Transfers and Capital Gains of Individuals*. Organization for Economic Cooperation and Development (OECD): Paris, France. 2

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

PIKETTY, THOMAS, AND EMMANUEL SAEZ (2013): "A Theory of Optimal Inheritance Taxation," *Econometrica*, 81(5), 1851–1886. 5, 21, 25

POTERBA, JAMES, STEVEN VENTI, AND DAVID WISE (2011): "The Composition and Drawdown of Wealth in Retirement," *Journal of Economic Perspectives*, 25(4), 95–118. 5

——— (2018): "Longitudinal Determinants of End-of-Life Wealth Inequality," *Journal of Public Economics*, 162, 78–88. 5

ROTH, JONATHAN (2019): "Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends," Working Paper. 12

SAEZ, EMMANUEL (2010): "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2(3), 180–212. 50

SAEZ, EMMANUEL, AND STEFANIE STANTCHEVA (2018): "A Simpler Theory of Optimal Capital Taxation," *Journal of Public Economics*, 162, 120–142. 4, 20, 25, 26, 31

SAEZ, EMMANUEL, AND GABRIEL ZUCMAN (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]

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

SKATTEMINISTERIET (2002): *Skatteberegningsreglerne for personer – før og nu*. Danish Ministry of Taxation: Copenhagen, Denmark. [10]

SLEMROD, JOEL B., AND SHLOMO YITZHAKI (2002): "Tax Avoidance, Evasion, and Administration," in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. S. Feldstein, vol. 3. Elsevier: Amsterdam, NL. [51]

STRAUB, LUDWIG, AND IVÁN WERNING (2019): "Positive Long-Run Capital Taxation: Chamley-Judd Revisited," *American Economic Review*, forthcoming. [25]

ZOUTMAN, FLORIS T. (2015): "The Effect of Capital Taxation on Household Savings," Working Paper. [4]

ZUCMAN, GABRIEL (2019): "Global Wealth Inequality," *Annual Review of Economics*, 11, 109–138. [7]

FIGURE I: Distribution of Wealth in Denmark, 1980-2012
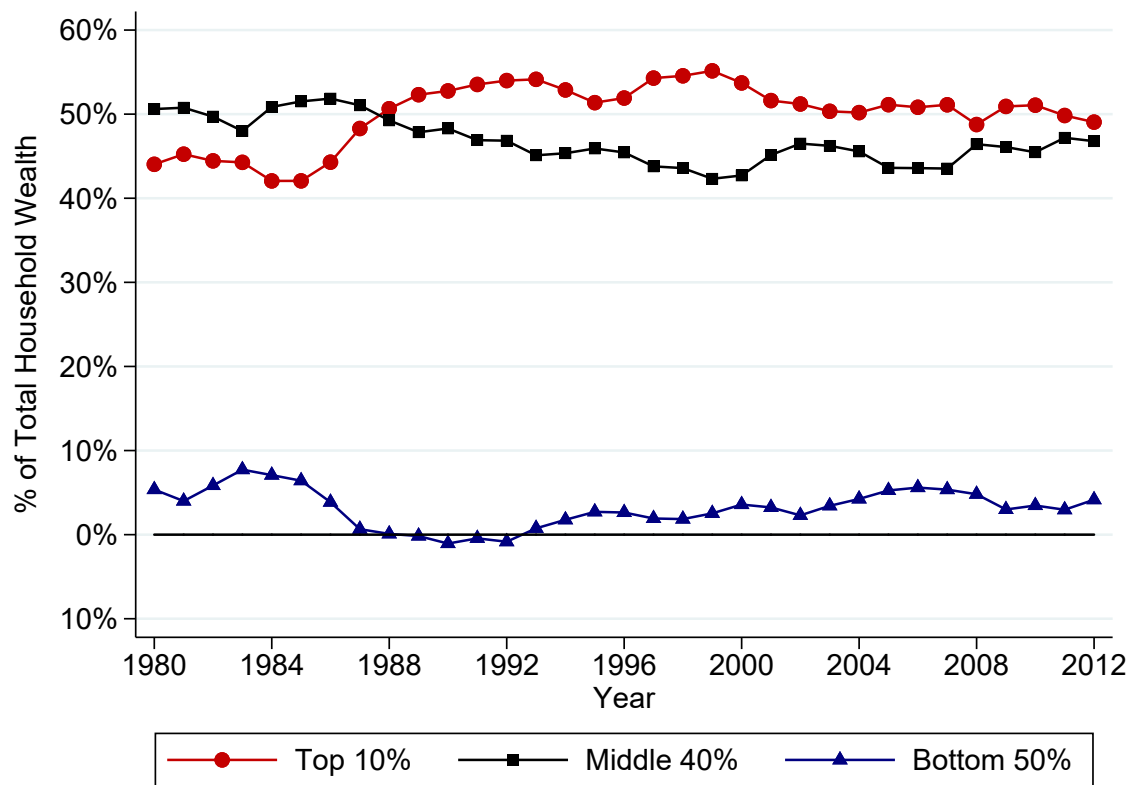
Notes: This figure shows the share of total household wealth in Denmark owned by the bottom 50% of the distribution, the middle 40% (adults between the median and the 90th percentile), and the top 10%. The unit of observation is the adult individual (aged 20 or above), splitting household wealth in married 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 II: Top 1% and Top 0.1% Wealth Shares in Denmark vs the United States

A: Top 1% Wealth Share
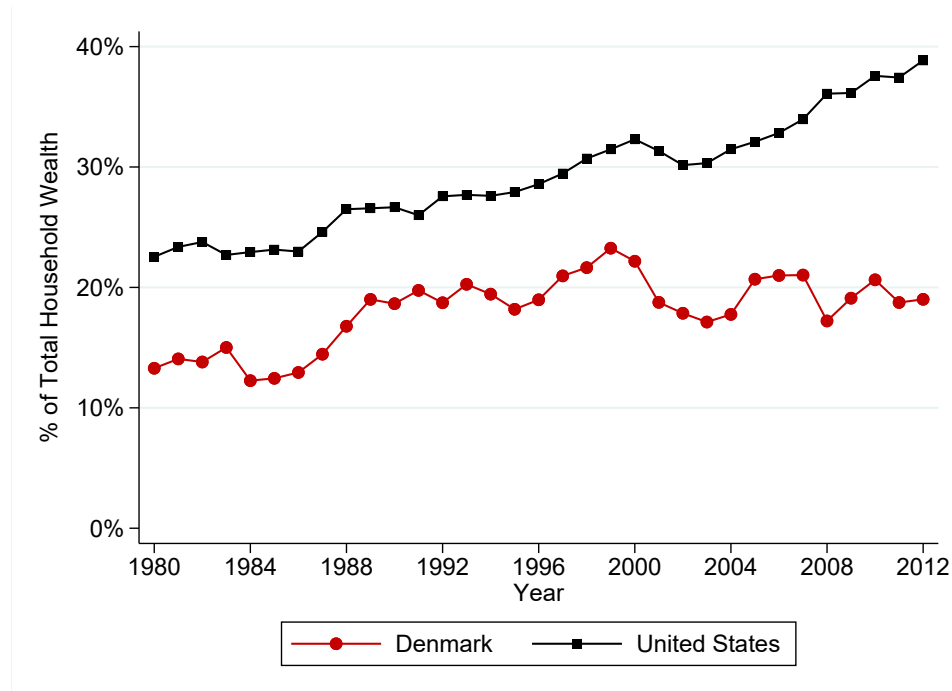
B: Top 0.1% Wealth Share
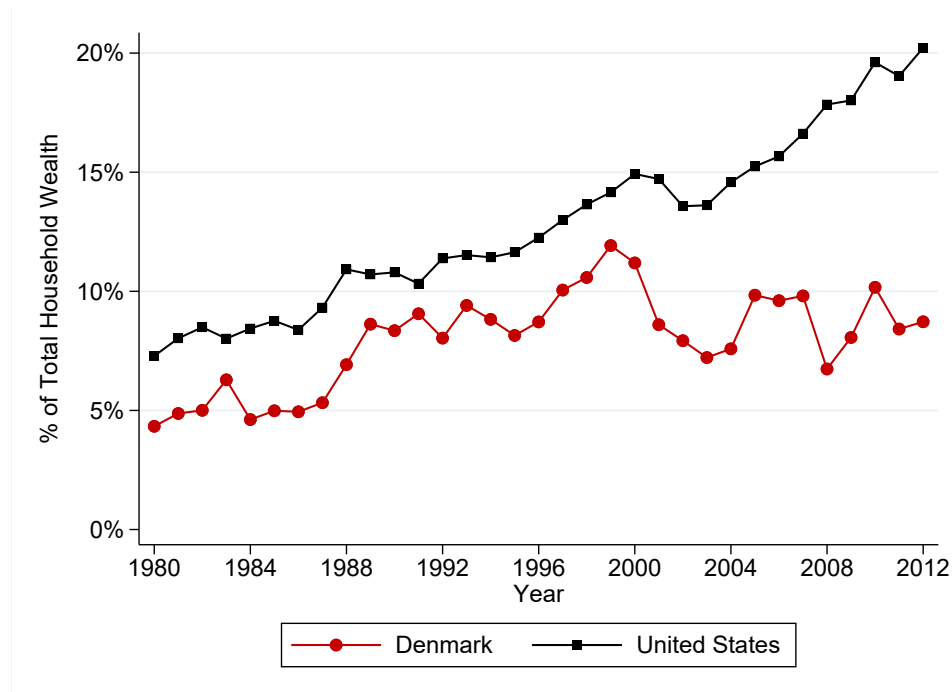
Notes: This figure shows the share of total household wealth owned by the top 1% (Panel A) and the top 0.1% (Panel B) in Denmark vs United States. In both countries, the unit of observation is the adult individual (aged 20 or above), splitting household wealth in married couples equally among the spouses. 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.

FIGURE III: Wealth Tax Variation

A: Marginal Tax Rate

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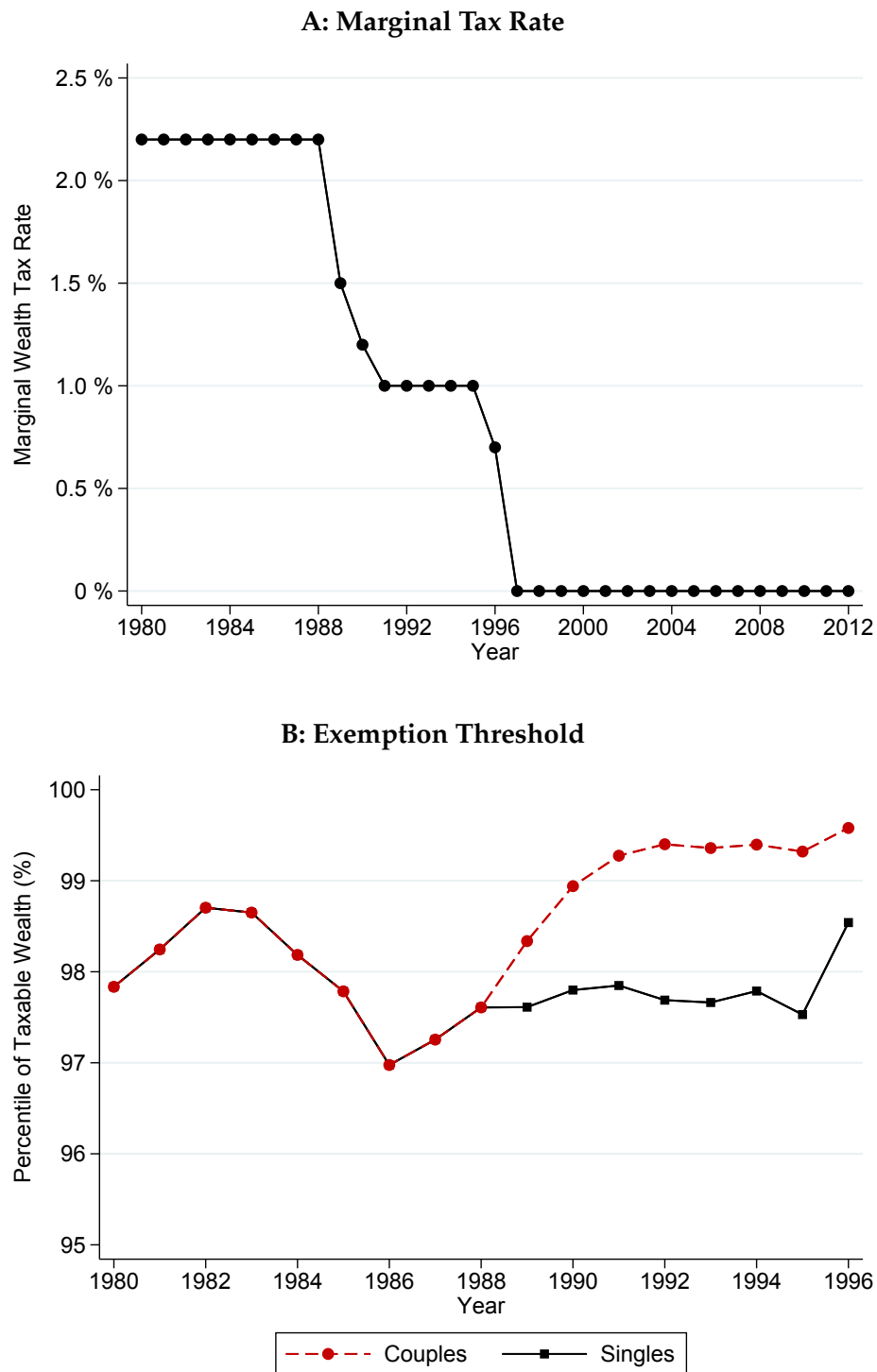
Notes: This figure shows the evolution of the marginal tax rate (Panel A) and the exemption threshold (Panel B) in the Danish wealth tax.

**FIGURE IV: Difference-in-Differences Comparing Couples and Singles within Exempted Range**

A: Time Series of Couples and Singles

B: Difference Between Couples and Singles

Notes: This figure shows the effects of the 1989 wealth tax reform on taxable wealth based on the couples DD in which we compare couples and singles within the exempted wealth range. The assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. The sample is a balanced panel of households observed in all years 1980-1996. Panel A shows the evolution of log taxable wealth in the two comparison groups, normalized to zero in the pre-reform year 1988. Panel B shows the differences between these two series, i.e. our reduced-form or intention-to-treat (ITT) estimates. The 95% confidence intervals are based on robust standard errors clustered at the household level.

40

FIGURE V: Couples DD: Treatment Effect on the Treated
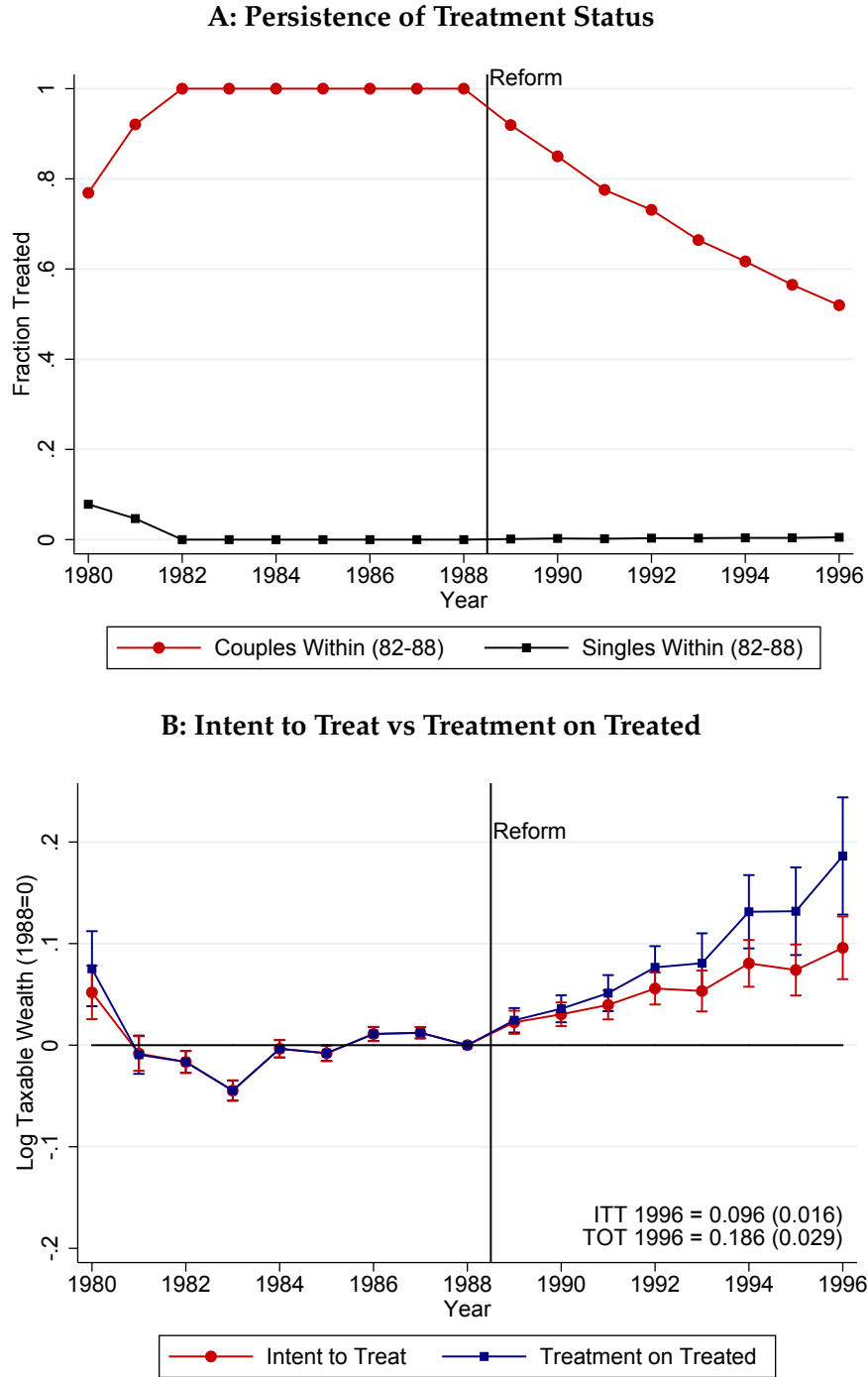
A: Persistence of Treatment Status

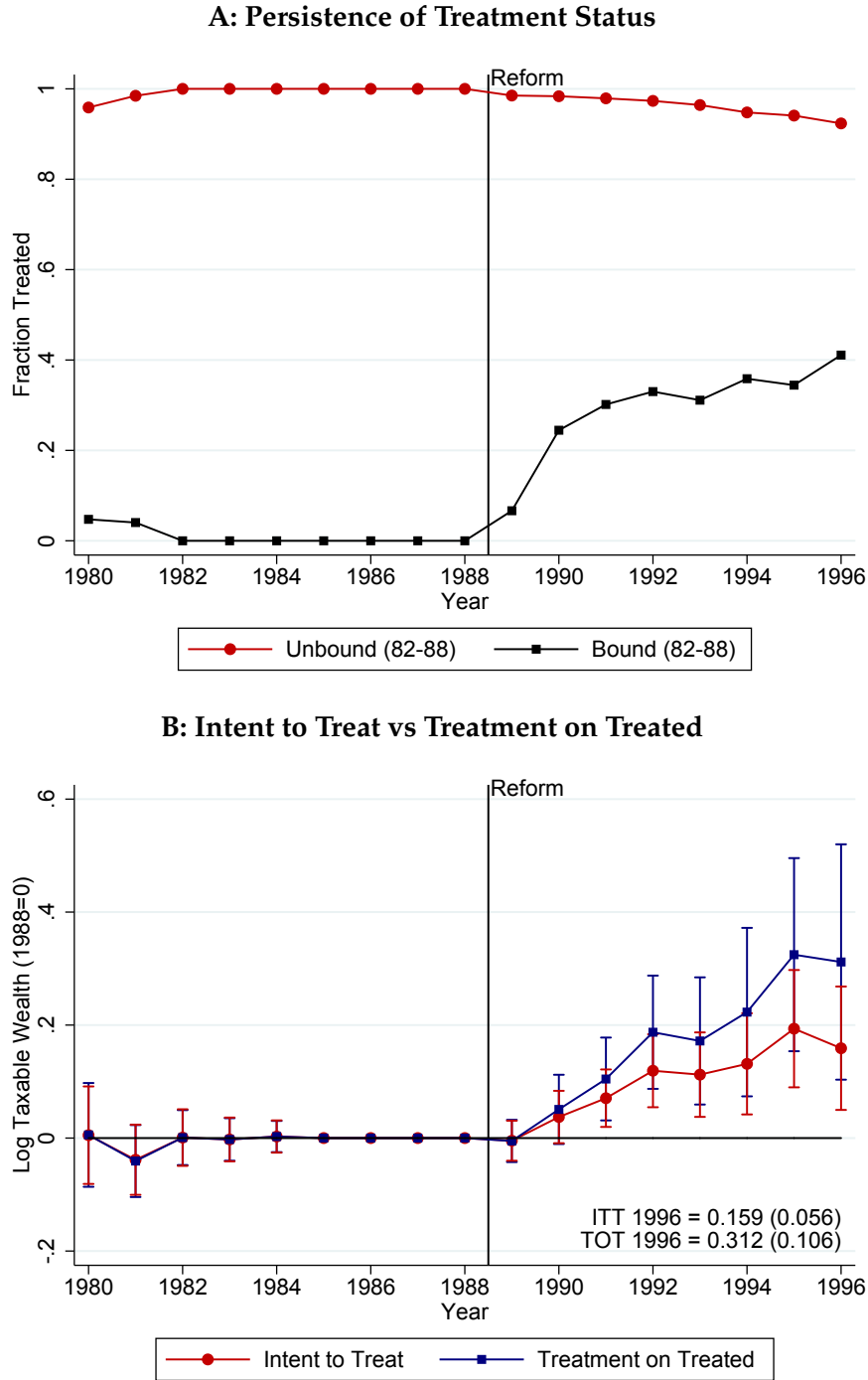## B: Intent to Treat vs Treatment on Treated

Notes: This figure converts intention-to-treat (ITT) estimates into treatment-on-the-treated (TOT) estimates based on the couples DD in which we compare couples and singles within the exempted wealth range. The assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. Panel A shows the persistence in treatment status over time. By construction, couples in the exempted range have a treatment status of 100% before the reform, while singles in this range have a treatment status of 0% before the reform. After the reform, taxpayers may switch status due to changes in relationship status or changes in wealth bracket. Panel B compares the ITT and TOT series, where the TOT estimates are obtained by instrumenting treatment status in equation (1) using treatment status in the pre-reform years 1982-88.

FIGURE VI: Difference-in-Differences Comparing Households Unbound and Bound by Tax Ceiling
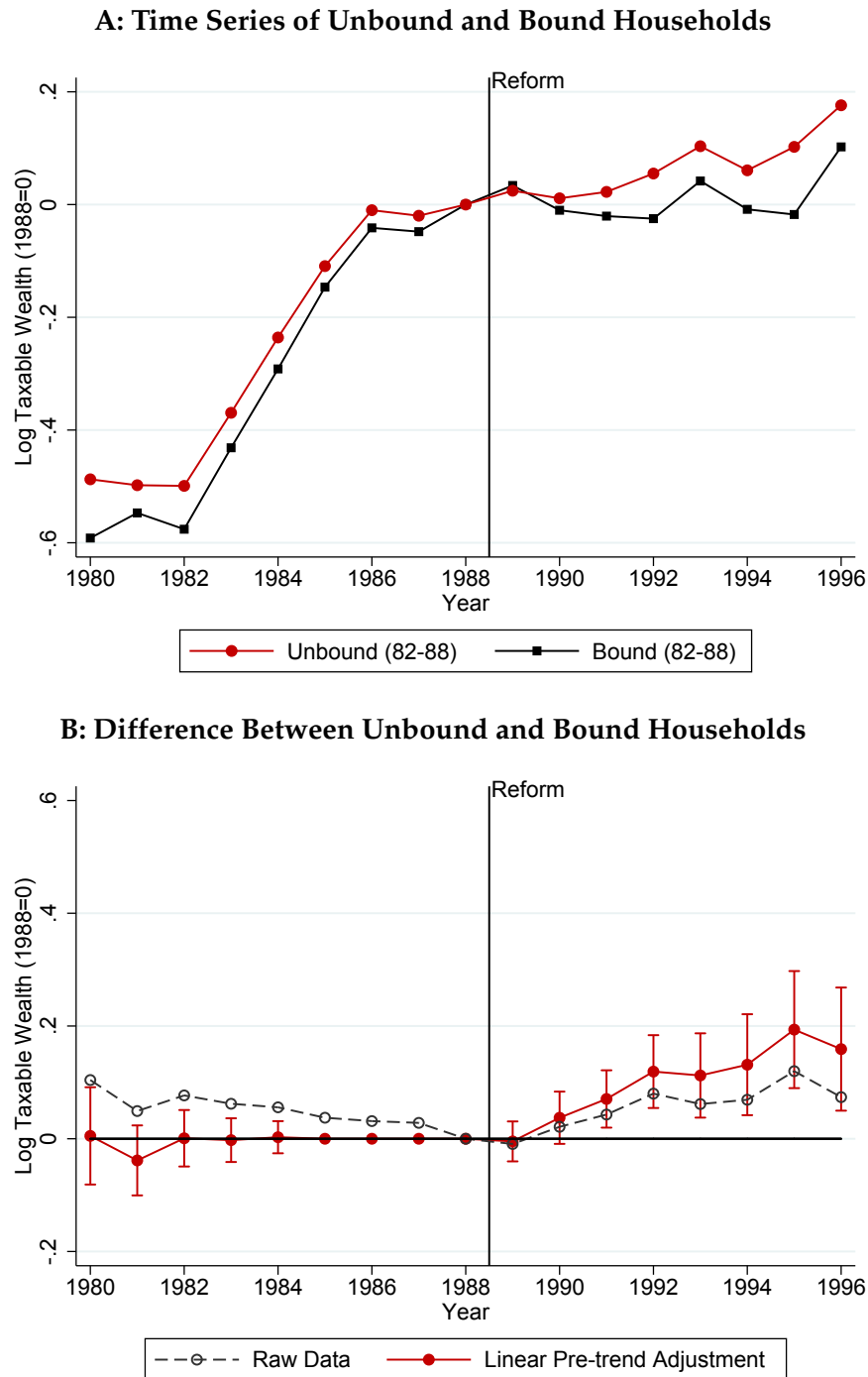
A: Time Series of Unbound and Bound Households

B: Difference Between Unbound and Bound Households

Notes: This figure shows the effects of the 1989 wealth tax reform on taxable wealth based on the ceiling DD in which we compare households who are unbound by the tax ceiling (treatments) to those who are bound by the tax ceiling (controls). The assignment of treatment status is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. The sample is a balanced panel of households observed in all years 1980-1996 and located in the top 1% of the wealth distribution before the reform. Panel A shows the evolution of log taxable wealth in the two comparison groups, normalized to zero in the pre-reform year 1988. Panel B shows the raw differences between these two series and the pre-trend adjusted differences (using equation 2). The 95% confidence intervals are based on robust standard errors clustered at the household level.

FIGURE VII: Ceiling DD: Treatment Effect on the Treated

A: Persistence of Treatment Status

B: Intent to Treat vs Treatment on Treated

ITT 1996 = 0.159 (0.056)
TOT 1996 = 0.312 (0.106)

Notes: This figure converts intention-to-treat (ITT) estimates into treatment-on-the-treated (TOT) estimates based on the ceiling DD in which we compare households bound and unbound by the tax ceiling. The assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. Panel A shows the persistence in treatment status over time. By construction, unbound households have a treatment status of 100% before the reform, while bound households have a treatment status of 0% before the reform. After the reform, households may switch status due to wealth or income shocks that change their tax ceiling status. Panel B compares the ITT and TOT series, where the TOT estimates are obtained by instrumenting treatment status in equation (2) using treatment status in the pre-reform years 1982-88.

FIGURE VIII: Empirical Lifecycle Profiles of Wealth

A: Full Population
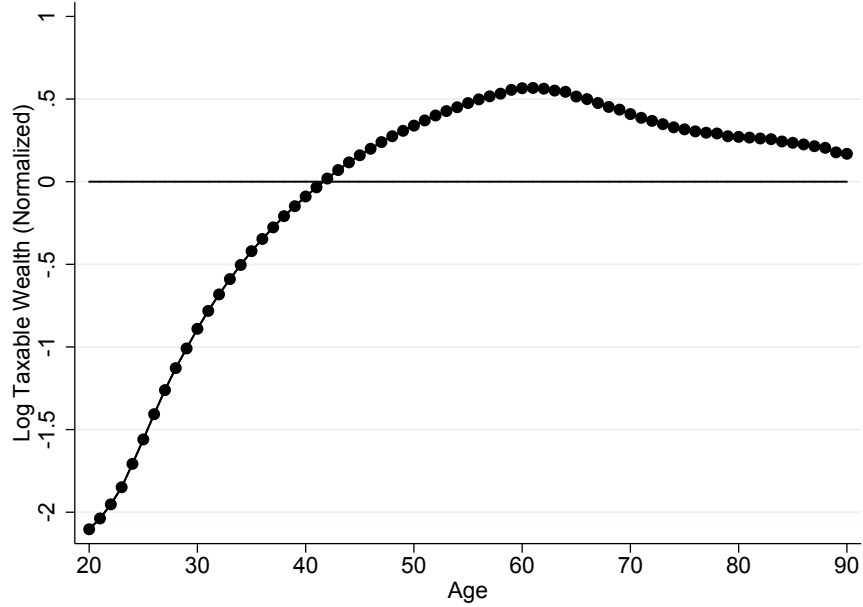
B: Top Percentiles
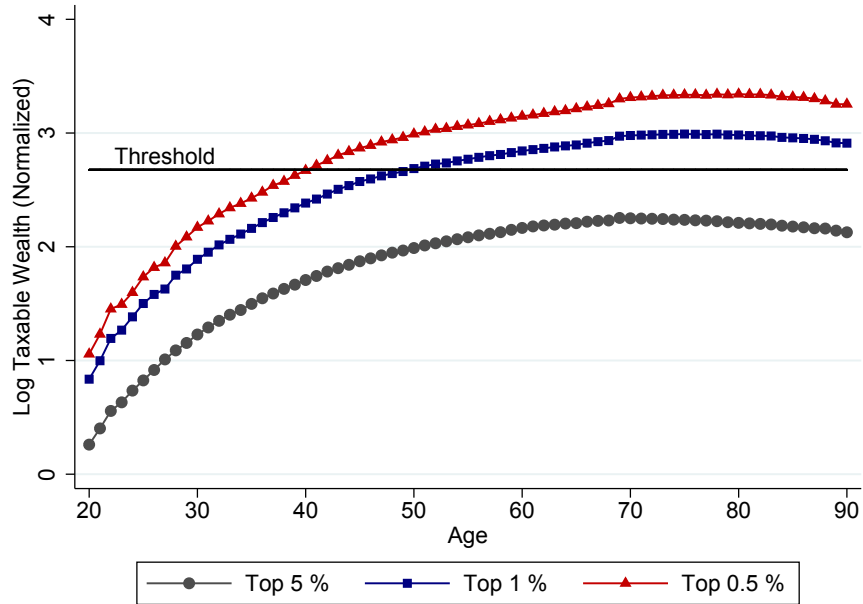
Notes: This figure shows lifecycle profiles of taxable wealth between the ages of 20-90 in an unbalanced panel of individuals over the period 1980-2012. To eliminate the confounding effects of inflation as people grow older, we normalize log wealth for each individual in each year by the average log 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 of individuals with positive net wealth, 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.
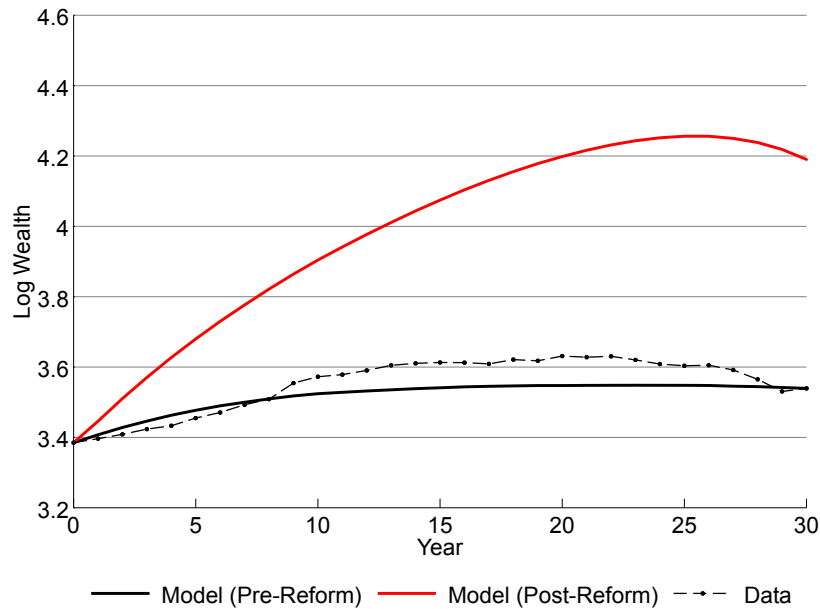
FIGURE IX: Long-Run Effects of Cutting Wealth Taxes: Couples DD (Moderately Wealthy)
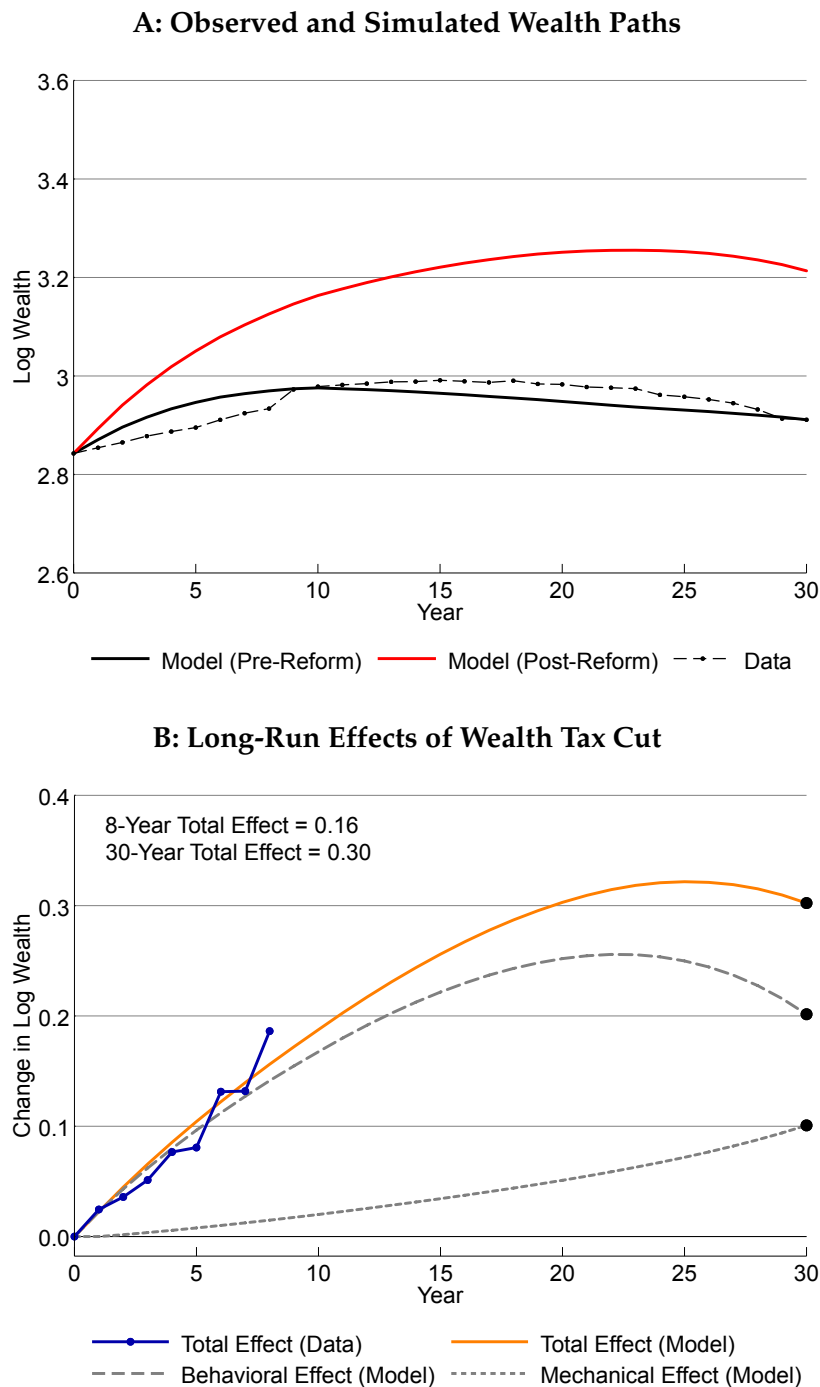
A: Observed and Simulated Wealth Paths

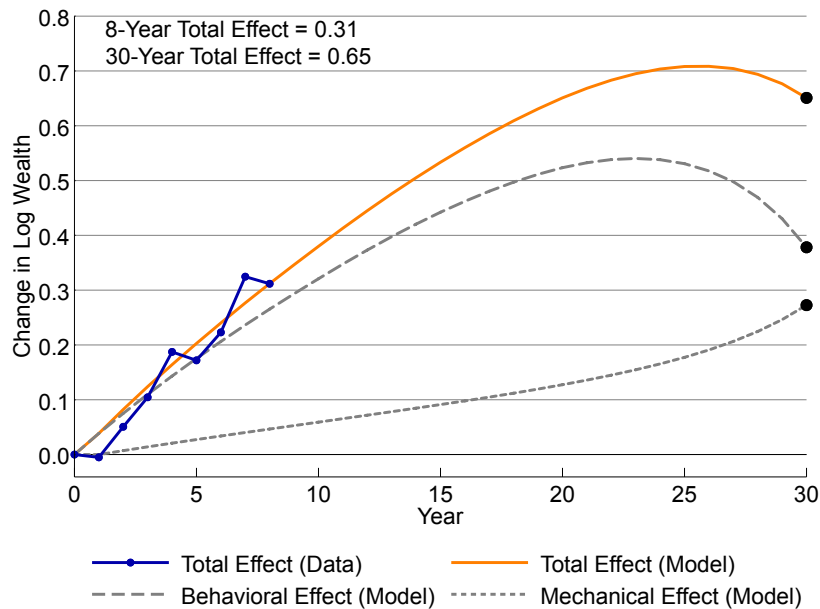**B: Long-Run Effects of Wealth Tax Cut**

Notes: This figure shows the long-run effects of wealth tax cuts when calibrating our model to the couples DD (comparing couples and singles in the exempted range). These are effects for the moderately wealthy (between the 98th and 99th percentile cutoffs). The reform experiment cuts the wealth tax rate by 1 percentage point, corresponding to the differential tax cut between the treatment and control groups. Panel A shows the observed 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 X: Long-Run Effects of Cutting Wealth Taxes: Ceiling DD (Very Wealthy)

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 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 observed 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 I: Descriptive Statistics

| | Full Population | Couples DD | | | Ceiling DD | |
		Couples Within (Treatment)	Singles Within (Control)	Couples Below (Control)	Unbound (Treatment)	Bound (Control)
	(1)	(2)	(3)	(4)	(5)	(6)
Taxable Wealth	336,746	3,155,285	3,144,293	1,788,803	4,680,904	19,859,232
Market Value Wealth	744,734	3,937,661	3,686,328	2,410,357	6,054,538	18,719,788
Gross Labor Income	203,795	134,471	60,741	165,343	206,594	288,987
Gross Total Income	311,219	606,962	420,399	422,403	866,325	1,771,303
Asset Share in Equities	0.01	0.06	0.09	0.03	0.12	0.28
Asset Share in Housing	0.35	0.30	0.30	0.41	0.28	0.19
Fraction Self-Employed	0.14	0.67	0.45	0.48	0.65	0.79
Age	46.25	60.21	63.57	59.49	61.34	57.84
Fraction Married	0.39	1.00	0.00	1.00	0.75	0.74
Years of Schooling	11.03	10.80	11.16	10.27	11.85	12.18
Observations	18,965,710	32,354	10,815	42,924	35,966	2,947

Notes: This table shows means of wealth, income, and demographics for the full population (column (1)) and for our treatment and control groups (columns (2)-(6)). Columns (2)-(4) show the comparison groups in the couples DD: couples and singles located within or below the exempted wealth range (but always within the top 5%). Columns (5)-(6) show the comparison groups in the ceiling DD: households bound or unbound by the tax ceiling (within the top 1%). The assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. The table shows statistics at the household level (for adults aged 20 or above). Individual-specific variables (age, schooling) have been averaged across spouses. The statistics are based on pooled pre-reform data from 1982-1988. Monetary values are reported in Danish Kroner in 2017 terms (The USD/DKK exchange rate was 6.2 as of December 31, 2017). Market value wealth is computed as described in section II.B. Gross total income includes all labor income, pensions, and capital income.

TABLE II: Difference-in-Differences Estimates of Taxable Wealth Responses to the 1989 Wealth Tax Reform

	Couples DD				Ceiling DD	
	Couples vs Singles		Couples Within vs Below		Unbound vs Bound	
	(1)	(2)	(3)	(4)	(5)	(6)
ITT Effects on Log Wealth:						
Average Effect	0.057	0.038	0.014	0.064	0.057	0.102
	(0.008)	(0.009)	(0.005)	(0.006)	(0.021)	(0.033)
1996 Effect	0.096	0.068	0.021	0.100	0.074	0.159
	(0.016)	(0.017)	(0.011)	(0.012)	(0.034)	(0.056)
TOT Effects on Log Wealth:						
Average Effect	0.090	0.060	0.031	0.135	0.095	0.171
	(0.012)	(0.014)	(0.011)	(0.013)	(0.034)	(0.052)
1996 Effect	0.186	0.133	0.058	0.277	0.144	0.312
	(0.029)	(0.033)	(0.029)	(0.032)	(0.066)	(0.106)
Elasticities:						
Elasticity wrt. $1 - \tau$	8.859	5.910	1.092	5.147	6.357	11.325
	(1.189)	(0.813)	(0.288)	(0.311)	(0.910)	(1.138)
Elasticity wrt. $(1 - \tau)R - 1$	0.345	0.232	0.050	0.217	0.219	0.393
	(0.048)	(0.033)	(0.010)	(0.010)	(0.030)	(0.038)
Observations	104,839	104,839	182,835	182,835	94,503	94,503
Household and Year FE	X	X	X	X	X	X
Linear Pre-trends		X		X		X

Notes: This table shows the intent-to-treat (ITT) and treatment-on-the-treated (TOT) effects of the 1989 wealth tax cuts on log wealth as well as the implied elasticities of taxable wealth. The table displays both the average effect over the post-reform period (1989-96) and the last-year effect (1996). The columns refer to the different empirical strategies: the couples DD (either couples vs singles within the exempted range or couples within vs below the exempted range) and the ceiling DD (unbound vs bound taxpayers). In all specifications, the assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. Columns (1), (3), and (5) show raw DD estimates without any pre-trend adjustment (specification 1), while columns (2), (4), and (6) show DD estimates adjusted for linear differential pre-trends (specification 2). For the couples DD using singles as controls, our preferred specification is the one without pre-trend adjustment (column 1). For the couples DD using couples below the exempted range as controls and for the ceiling DD, our preferred specifications are those with pre-trend adjustment (columns 4 and 6). The estimation sample is a balanced panel of households observed in all years 1980-1996. Robust standard errors are clustered at the household level.

TABLE III: Simulation Analysis of Long-Run Effects

	Couples DD				Ceiling DD	
	Couples vs Singles		Couples Within vs Below		Unbound vs Bound	
	Low Return	High Return	Low Return	High Return	Low Return	High Return
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 30-Year Effects on Log Wealth						
Total Effect	0.302	0.313	0.499	0.509	0.651	0.639
Mechanical Effect	0.101	0.142	0.240	0.354	0.273	0.419
Behavioral Effect	0.202	0.170	0.259	0.155	0.378	0.219
Elasticity wrt. $(1-\tau)R-1$	0.774	1.359	0.581	1.005	1.150	1.913
Panel B: Structural Primitives						
R	1.05	1.07	1.05	1.07	1.05	1.07
δ	0.973	0.955	0.973	0.955	0.975	0.957
A	16.51	12.99	16.74	13.56	25.70	19.27
σ	2.62	2.25	2.20	1.83	6.65	4.36
α	2.14	1.86	1.50	1.31	3.81	2.83

Notes: This table summarizes the simulation results shown in Figures IX, X, and A.XX, and reports the structural primitives underlying each calibration. The table also investigates the sensitivity of our results to changing the assumed rate of return R: our baseline assumption of a 5% before-tax return is shown in the columns labelled "Low Return", while the alternative assumption of a 7% before-tax return is shown in the columns labelled "High Return."

ONLINE APPENDIX FOR

WEALTH ACCUMULATION AND WEALTH TAXATION:
THEORY AND EVIDENCE FROM DENMARK

BY KATRINE JAKOBSEN, KRISTIAN JAKOBSEN, HENRIK KLEVEN, AND GABRIEL ZUCMAN

A   BUNCHING ANALYSIS

This section presents evidence of bunching at the kink point created by the exemption threshold of the Danish wealth tax. As discussed in the paper, bunching is useful for detecting evasion and avoidance responses to wealth taxes, but not for estimating real wealth accumulation responses. Figure A.II presents bunching evidence for the full population, pooling data across all the 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 $3,000$ 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.[37] This is the standard approach to estimating counterfactual distributions in the bunching literature (Chetty, Friedman, Olsen, and Pistaferri 2011; Kleven and 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 demonstrated by Saez (2010). We refer to the review by Kleven (2016) for details and discussions of the approach.

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.[38] 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 net-of-tax return. The implied elasticity is therefore tiny.

Figure A.III investigates heterogeneity in bunching across samples with different opportunities to avoid or evade. It compares employees and self-employed taxpayers (Panels A-B) as well as "ordinary" and "non-ordinary" taxpayers (Panels C-D). Non-ordinary taxpayers refer to those 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 greater avoidance opportunities, but that bunching remains modest in all samples.

---

[37]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).

[38]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.

These findings would seem to suggest that evasion and avoidance responses to changes in wealth taxation are modest. A caveat to this interpretation, however, is that the amount of bunching may understate the responsiveness away from the kink. As noted in the paper, optimization frictions such as fixed avoidance costs or lumpy assets attenuate observed bunching even if taxpayers are relatively responsive to taxes. Furthermore, leaving aside frictions, two related points are worth highlighting. First, the finding of small bunching responses applies to the sample of *moderately* wealthy people located around the kink. It does not necessarily extend to the very wealthy located higher up. Second, any finding of small avoidance *responses* does not necessarily imply small avoidance *levels*. 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 and Yitzhaki 2002; Kleven, Knudsen, Kreiner, Pedersen, and Saez 2011).

To compute the mechanical effects of the 1989 reform, we distinguish between single and married households 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 a given 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

$$\widetilde{W}_t^m = \begin{cases} \mathsf{E}\left[W_{it}^m \mid m, t\right] & \text{if } t = 1988 \\ \widetilde{W}_{t-1}^m \times G_t^m & \text{if } t > 1988 \end{cases} \tag{20}$$

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.[39] 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

$$\begin{aligned} \Delta\text{tax}_t^m = \quad (1-m) \quad & \times \quad \mathbf{I}\left[\widetilde{W}_t^0 \geq \overline{W}_t^0\right] \times \left(\Delta\tau_t \cdot \left(\widetilde{W}_t^0 - \overline{W}_t^0\right)\right) \\ +m \quad & \times \quad \left\{\mathbf{I}\left[\overline{W}_t^0 \leq \widetilde{W}_t^1 < \overline{W}_t^1\right] \times \left(0.022 \cdot \left(\widetilde{W}_t^1 - \overline{W}_t^0\right)\right) \right. \\ & \left. + \mathbf{I}\left[\widetilde{W}_t^1 \geq \overline{W}_t^1\right] \times \left(0.022 \cdot \left(\overline{W}_t^1 - \overline{W}_t^0\right) + \Delta\tau_t \cdot \left(\widetilde{W}_t^1 - \overline{W}_t^1\right)\right)\right\}. \end{aligned} \tag{21}$$

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.

Equation (21) 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 $R = 1.05$, broadly in line with empirical estimates of rates of return and consistent with the calibration exercise in section V. The cumulative tax savings for marital status m in year t can be defined as

---

[39] 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.

$$\Delta\text{cumtax}_t^m = \begin{cases} \Delta\text{tax}_t^m & \text{if } t = 1989 \\ \Delta\text{tax}_t^m + \Delta\text{cumtax}_{t-1}^m \times R & \text{if } t > 1989 \end{cases} \tag{22}$$

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

$$\Delta\log W_t = \mathsf{E}\left[\log\left(\widetilde{W}_t^m + \Delta\text{cumtax}_t^m\right) - \log\left(\widetilde{W}_t^m\right) \mid t\right], \tag{23}$$

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 households in the treatment group (couples in the exempted wealth range for the couples DD; households unbound by the tax ceiling for the ceiling DD). In the couples DD specification where we compare couples and singles within the exempted range (between $\overline{W}_t^0$ and $\overline{W}_t^1$), the control group saves some taxes too.[40] We calculate the mechanical tax savings for the control group using the same approach as for the treatment group.[41] The net mechanical effect on wealth (that we subtract from the total estimated effect to obtain the behavioral effect) is equal to the effect for the treatment group minus the effect for the control group.

---

[40]In the other couples DD specifications (comparing to couples or singles below the exempted range) or in the ceiling DD, the control group does not save any taxes.

[41]The calculation of tax savings for the control group using equations (21)-(22) 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.

C Proofs

To prove Proposition 1, we totally differentiate equation (12) with respect to $1 - \tau$ and c_0. After rearranging terms, this gives

$$\frac{dW_1}{d(1-\tau)} \left\{ \sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\frac{\alpha}{\sigma}-1} \right\} = \sum_{t=0}^{T} \frac{\partial q_t}{\partial(1-\tau)} \cdot c_0 + \frac{\partial q_b}{\partial(1-\tau)} \cdot c_0^{\frac{\alpha}{\sigma}}$$
$$+ \sum_{t=0}^{T} \frac{tRy_t}{((1-\tau)R)^{t+1}} + \sum_{t=1}^{T} \left\{ \frac{[(1-\tau)R + t\tau R] R\bar{W}}{((1-\tau)R)^{t+1}} \right\}, \quad (24)$$

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

$$\frac{\partial q_t}{\partial(1-\tau)} = \frac{\sigma t q_t}{1-\tau} - \frac{tR(\delta(1-\tau)R)^{t\sigma}}{((1-\tau)R)^{t+1}}, \quad (25)$$

$$\frac{\partial q_b}{\partial(1-\tau)} = \frac{\alpha T q_b}{1-\tau} - \frac{TRA(\delta(1-\tau)R)^{T\alpha}}{((1-\tau)R)^{T+1}}, \quad (26)$$

which allows us to rewrite equation (24) as follows

$$\frac{dW_1}{d(1-\tau)} \left\{ \sum_{t=0}^{T} q_t + q_b \frac{\alpha}{\sigma} c_0^{\frac{\alpha}{\sigma}-1} \right\} = \sum_{t=0}^{T} \frac{\sigma t q_t}{1-\tau} \cdot c_0 + \frac{\alpha T q_b}{1-\tau} \cdot c_0^{\frac{\alpha}{\sigma}}$$
$$- \sum_{t=0}^{T} \frac{tR(\delta(1-\tau)R)^{t\sigma}}{((1-\tau)R)^{t+1}} \cdot c_0 - \frac{TRA(\delta(1-\tau)R)^{T\alpha}}{((1-\tau)R)^{T+1}} \cdot c_0^{\frac{\alpha}{\sigma}} \quad (27)$$
$$+ \sum_{t=0}^{T} \frac{tRy_t}{((1-\tau)R)^{t+1}} + \sum_{t=1}^{T} \left\{ \frac{[(1-\tau)R + t\tau R] R\bar{W}}{((1-\tau)R)^{t+1}} \right\}.$$

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 (7) 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

$$\frac{dW_0^C}{d(1-\tau)} = - \sum_{t=0}^{T} \frac{tRc_t}{((1-\tau)R)^{t+1}} - \frac{TRW_{T+1}}{((1-\tau)R)^{T+1}}$$
$$+ \sum_{t=0}^{T} \frac{tRy_t}{((1-\tau)R)^{t+1}} + \sum_{t=1}^{T} \left\{ \frac{[(1-\tau)R + t\tau R] R\bar{W}}{((1-\tau)R)^{t+1}} \right\}. \quad (28)$$

The expression on the RHS of (28) is identical to the second and third lines of (27) once we account for the fact that, from household optimization, $c_t = (\delta(1-\tau)R)^{t\sigma} c_0$ and $W_{T+1} = A(\delta(1-\tau)R)^{T\alpha} c_0^{\frac{\alpha}{\sigma}}$.

54

This allows us to rewrite equation (27) to

$$\frac{dW_1}{d\left(1-\tau\right)}\left\{\sum_{t=0}^{T}q_t+q_b\frac{\alpha}{\sigma}c_0^{\frac{\alpha}{\sigma}-1}\right\}=\sum_{t=0}^{T}\frac{\sigma t q_t}{1-\tau}\cdot c_0+\frac{\alpha T q_b}{1-\tau}\cdot c_0^{\frac{\alpha}{\sigma}}+\frac{dW_0^C}{d\left(1-\tau\right)}, \tag{29}$$

which can be immediately rearranged to give the expression in equation (13) 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 (14) of Proposition 2 follows from setting $\alpha=\sigma$ (which implies $q_b=Aq_T$) in equation (13). 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 (15) with respect to $1-\tau$, W_t, and c_j. This gives

$$
\begin{aligned}
\frac{dW_t/W_t}{d\left(1-\tau\right)/\left(1-\tau\right)} &= \left(t-1\right)\left[\left(1-\tau\right)R\right]^{t-1}\frac{W_0^n}{W_t}+\sum_{j=0}^{t-1}\left(t-1-j\right)\left[\left(1-\tau\right)R\right]^{t-1-j}\frac{y_j-c_j}{W_t} \\
&\quad -\sum_{j=1}^{t-1}\left(1-\frac{\tau\left(t-1-j\right)}{1-\tau}\right)\left[\left(1-\tau\right)R\right]^{t-j}\frac{\bar{W}}{W_t} \\
&\quad -\sum_{j=0}^{t-1}\left[\left(1-\tau\right)R\right]^{t-1-j}\frac{dc_j/W_t}{d\left(1-\tau\right)/\left(1-\tau\right)}.
\end{aligned}
\tag{30}
$$

In this expression, the first three terms correspond to the mechanical effect dM shown in equation (17) of Proposition 3. The last term corresponds to the behavioral effect dB in equation (18). To see this, we start from the Euler equation shown in (10): $c_j=\left(\delta\left(1-\tau\right)R\right)^{j\sigma}c_0$. We totally differentiate this condition and rearrange terms, which gives

$$\frac{dc_j/W_t}{d\left(1-\tau\right)/\left(1-\tau\right)}=j\sigma\left(\delta\left(1-\tau\right)R\right)^{j\sigma}\frac{c_0}{W_t}+\left(\delta\left(1-\tau\right)R\right)^{j\sigma}\frac{dc_0/W_t}{d\left(1-\tau\right)/\left(1-\tau\right)}. \tag{31}$$

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

$$
\begin{aligned}
\frac{dc_j/W_t}{d\left(1-\tau\right)/\left(1-\tau\right)} &= j\sigma\left(\delta\left(1-\tau\right)R\right)^{j\sigma}\frac{c_0}{W_t}-\left(\delta\left(1-\tau\right)R\right)^{j\sigma}\frac{dW_1/W_0}{d\left(1-\tau\right)/\left(1-\tau\right)}\frac{W_0}{W_t} \\
&= \left(\delta\left(1-\tau\right)R\right)^{j\sigma}\left\{j\sigma\frac{c_0}{W_t}-\frac{dW_1/W_0}{d\left(1-\tau\right)/\left(1-\tau\right)}\frac{W_0}{W_t}\right\}.
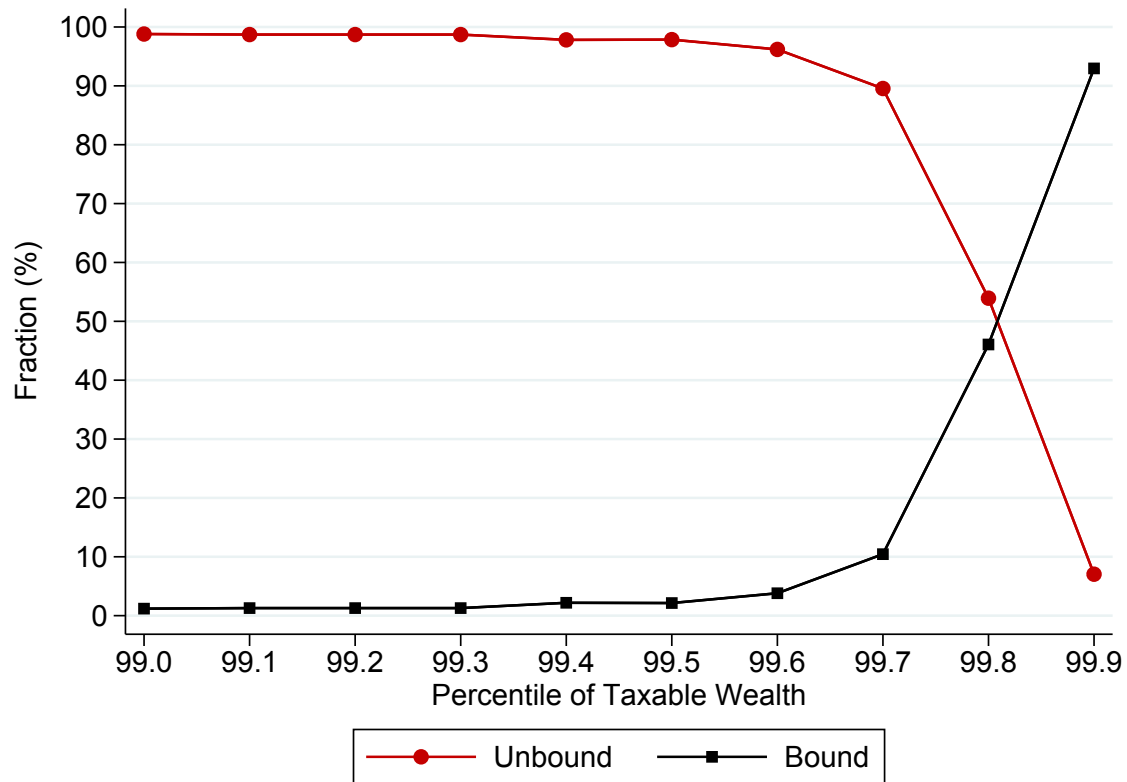\end{aligned}
\tag{32}
$$

Finally, inserting (32) into (30) and using $q_j \equiv \frac{(\delta(1-\tau)R)^{j\sigma}}{((1-\tau)R)^j}$, we obtain

$$
\begin{aligned}
\frac{dW_t/W_t}{d\left(1-\tau\right)/\left(1-\tau\right)} \;=\; & \left(t-1\right)\left[\left(1-\tau\right)R\right]^{t-1}\frac{W_0^n}{W_t} + \sum_{j=0}^{t-1}\left(t-1-j\right)\left[\left(1-\tau\right)R\right]^{t-1-j}\frac{y_j-c_j}{W_t} \\
& -\sum_{j=1}^{t-1}\left(1-\frac{\tau\left(t-1-j\right)}{1-\tau}\right)\left[\left(1-\tau\right)R\right]^{t-j}\frac{\bar{W}}{W_t} \\
& +\sum_{j=0}^{t-1}\frac{q_j}{\left[\left(1-\tau\right)R\right]^{1-t}}\left\{\frac{dW_1/W_0}{d\left(1-\tau\right)/\left(1-\tau\right)}\frac{W_0}{W_t}-j\sigma\frac{c_0}{W_t}\right\},
\end{aligned}
\tag{33}
$$

which is the result in Proposition 3.

FIGURE A.I: Fractions Bound and Unbound by Tax Ceiling

Notes: This 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: Bunching at the Kink

A: Observed Distribution
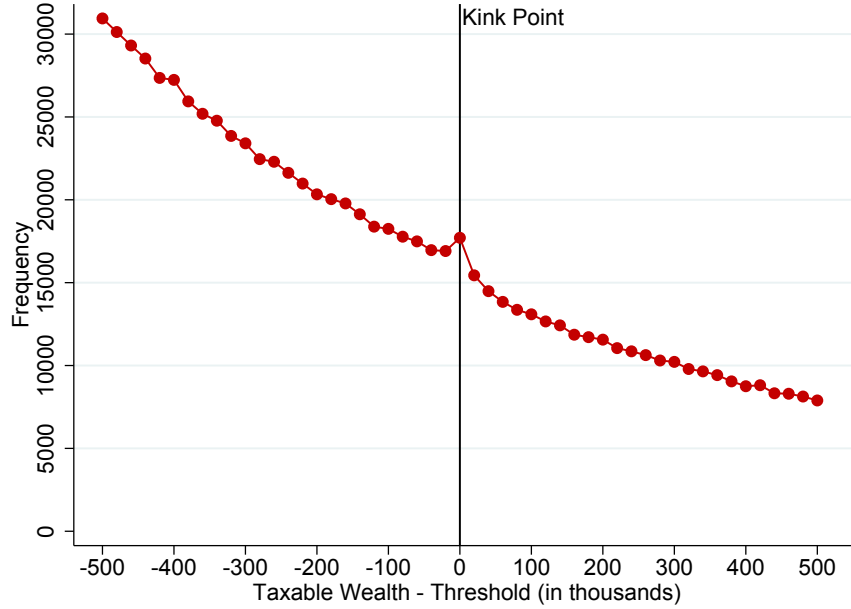
B: Observed vs Counterfactual Distribution
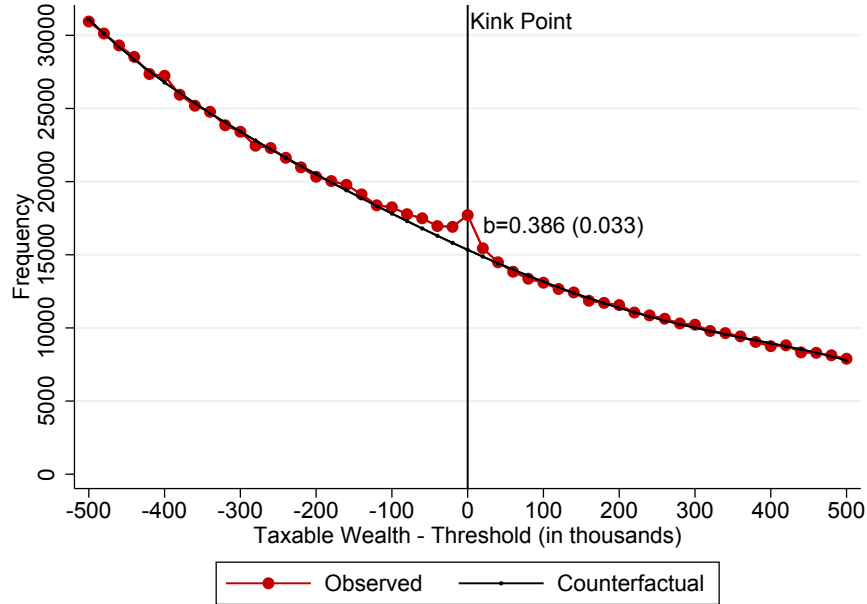
Notes: This figure shows the distribution of taxable wealth in bins of 20,000 DKK (about 3,000 USD) around the exemption threshold (kink point) of the Danish wealth tax, pooling data from 1980-1996. Taxable wealth is measured in terms of its distance to the threshold in each year (putting the kink point 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.

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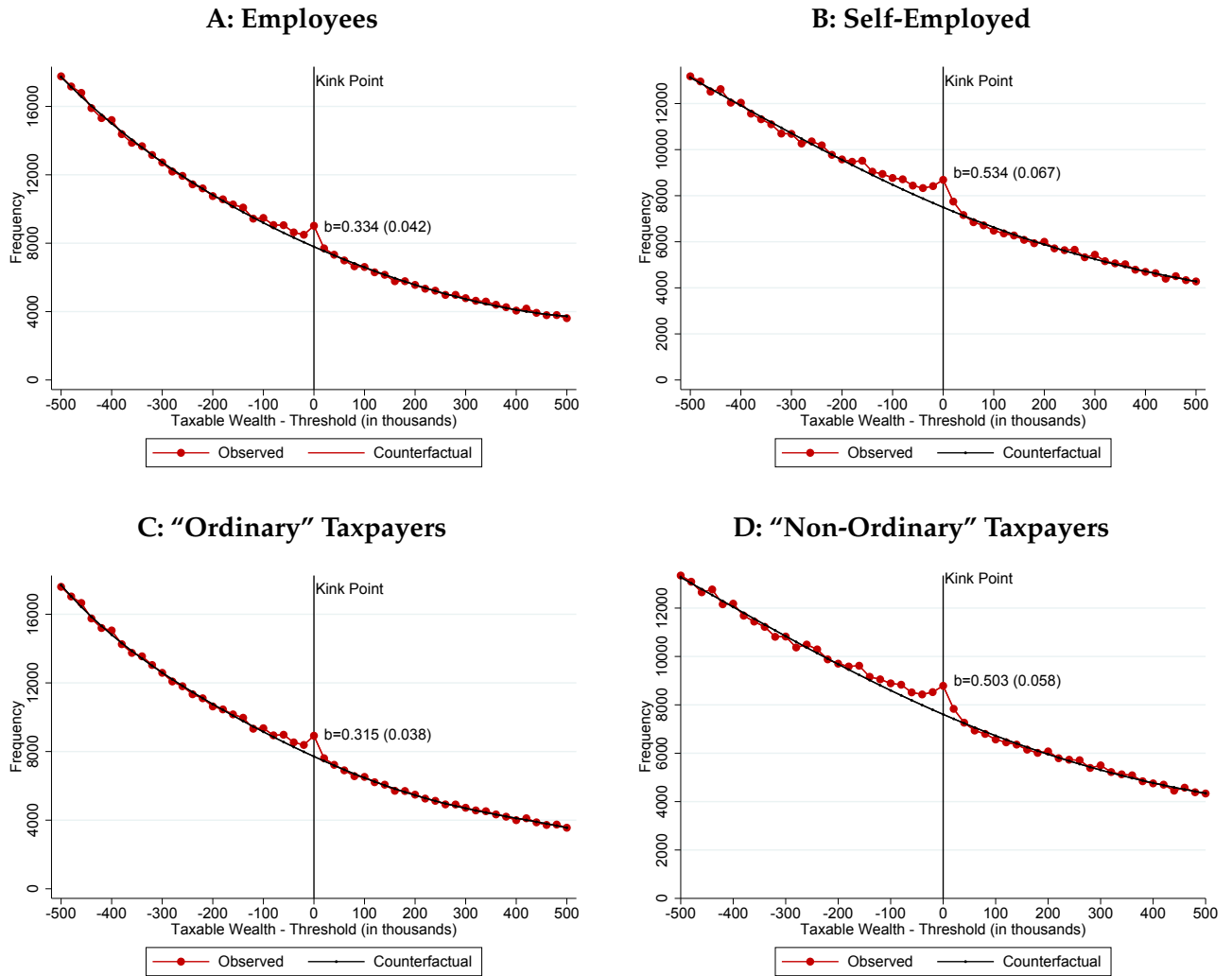
Notes: This figure shows heterogeneity in bunching across samples with different opportunities to avoid or evade taxes. Panels A and B compare employees and self-employed taxpayers, while Panels C and D compare "ordinary" and "non-ordinary" taxpayers. Non-ordinary taxpayers are defined as those who have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

FIGURE A.IV: Couples DD: Behavioral vs Mechanical Effects
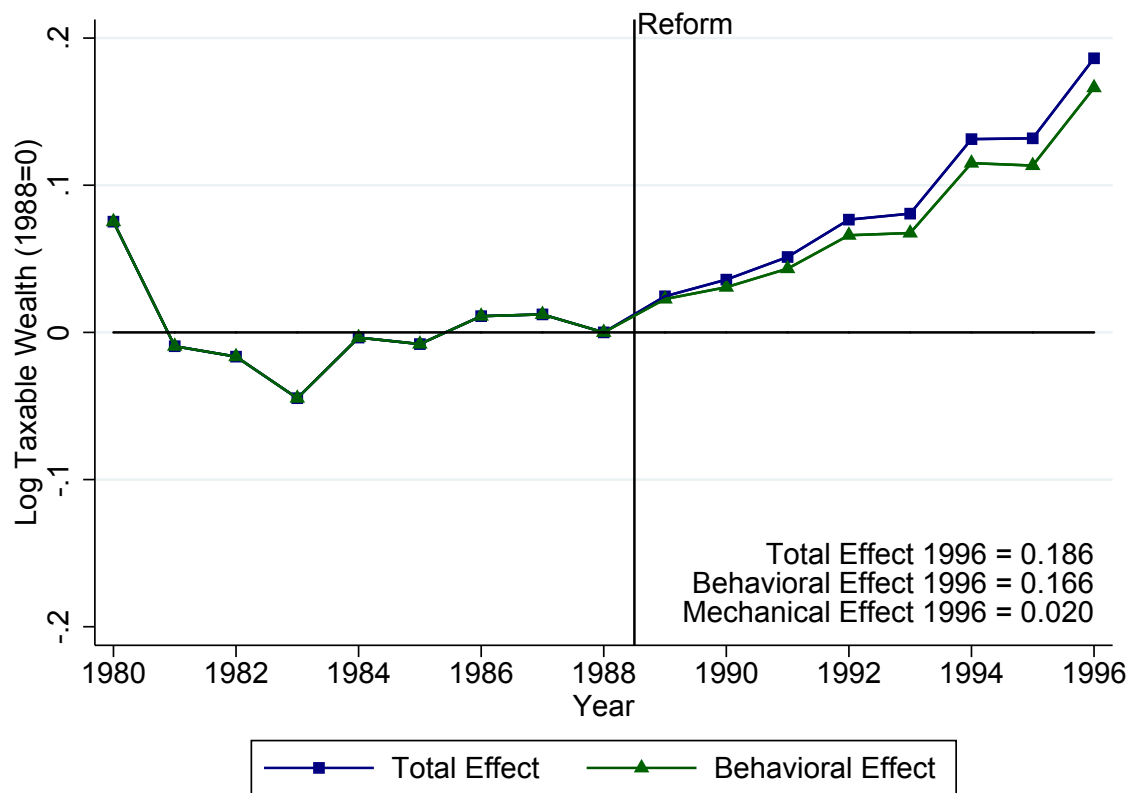
Notes: This figure shows the composition of the total effect of wealth tax reform into behavioral and mechanical effects for the couples DD in which we compare couples and singles within the exempted wealth range. Appendix B provides details on how the mechanical effects are calculated. The behavioral effect equals the total effect minus the mechanical effect.

FIGURE A.V: Couples DD: Robustness to Treatment Window

A: 1980-1988

B: 1982-88 (Baseline)

C: 1984-1988

D: 1986-1988
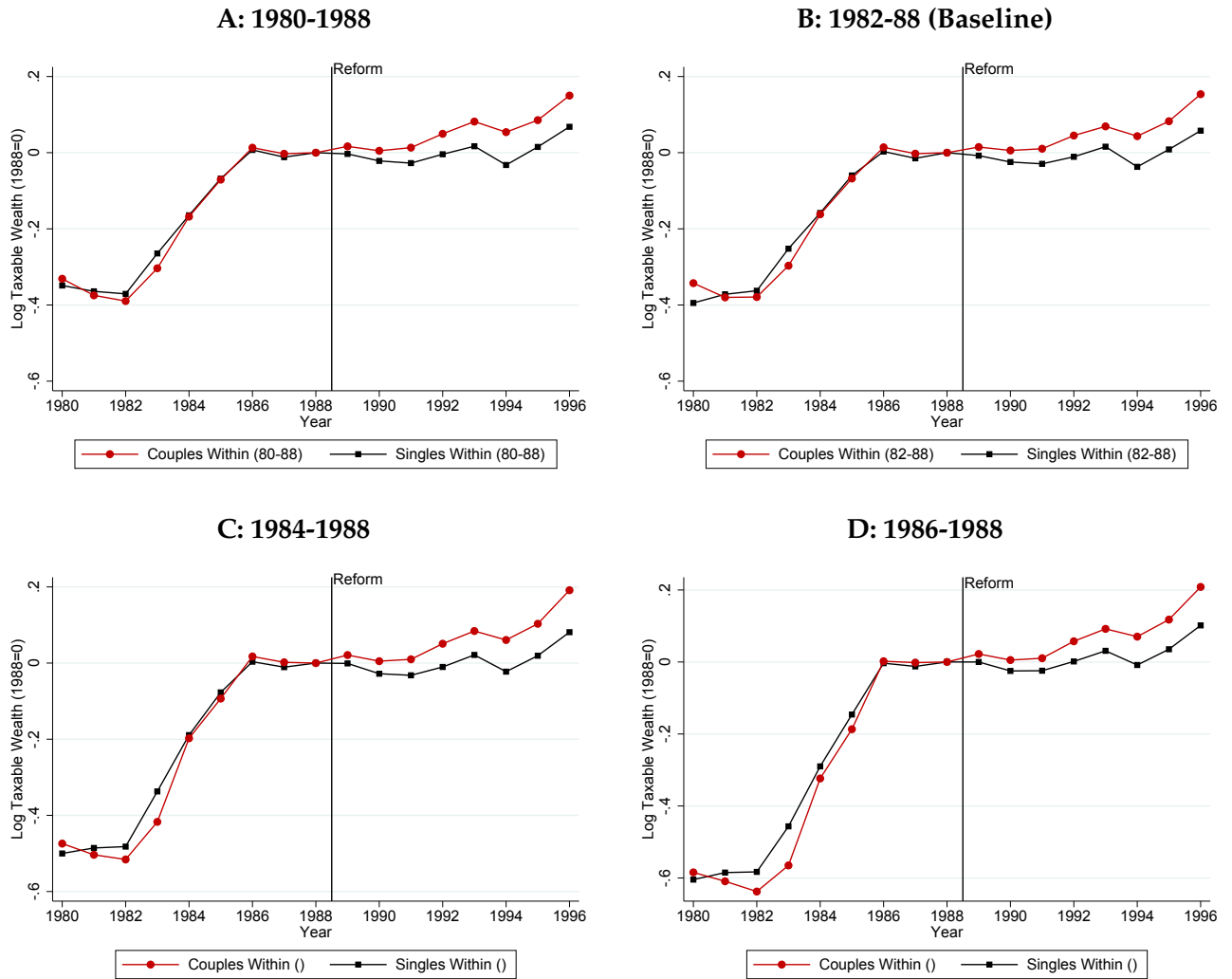
Notes: This figure investigates the robustness of the couples DD results to defining comparison groups (couples and singles in the exempted wealth range) based on outcomes in specific pre-reform years. The four panels show the effects on log taxable wealth when using different pre-reform windows to define treatment status: 1980-88, 1982-88 (baseline specification), 1984-88, and 1986-88. Besides varying the pre-reform treatment window, each panel is constructed in the same way as Panel A of Figure IV. The figure shows that the results are similar across the alternative specifications.

FIGURE A.VI: Couples DD: Placebo Tests

A: Placebo Reform 1983

B: Placebo Reform 1984

C: Placebo Reform 1985

D: Placebo Reform 1986
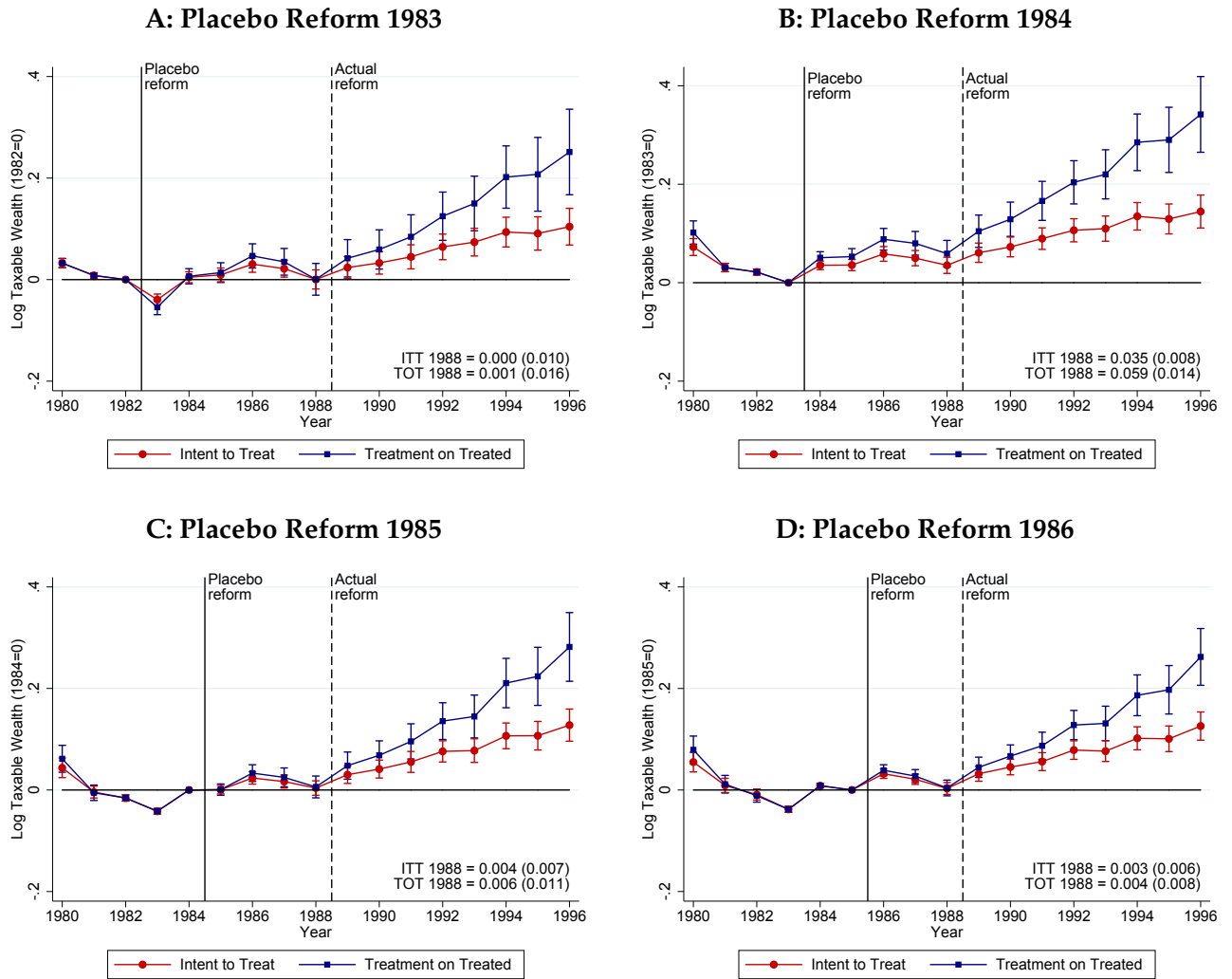
Notes: This figure shows ITT and TOT estimates obtained from the couples DD (couples vs singles in the exempted wealth range) and a set of placebo reforms. We consider placebo reforms implemented in earlier years (1983, 1984, 1985, and 1986) than the actual reform (1989). The comparison groups are couples and singles within the exempted wealth range as in the main specification, but the assignment of treatment status is based outcomes prior to the placebo reforms rather than the actual reform. We use three pre-reform years to assign treatment status (e.g., 1980-82 for the 1983 placebo reform). In all other respects, the figure is constructed in the same way as Panel B of Figure V.

FIGURE A.VII: Couples DD Using Couples Below Threshold as Controls

A: Time Series of Couples Within and Below Range

B: Difference Between Couples Within and Below Range
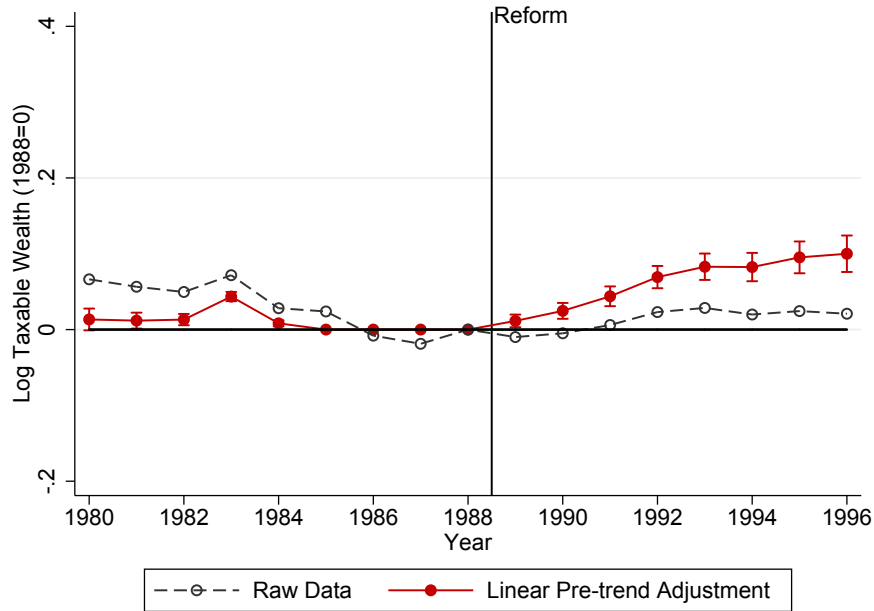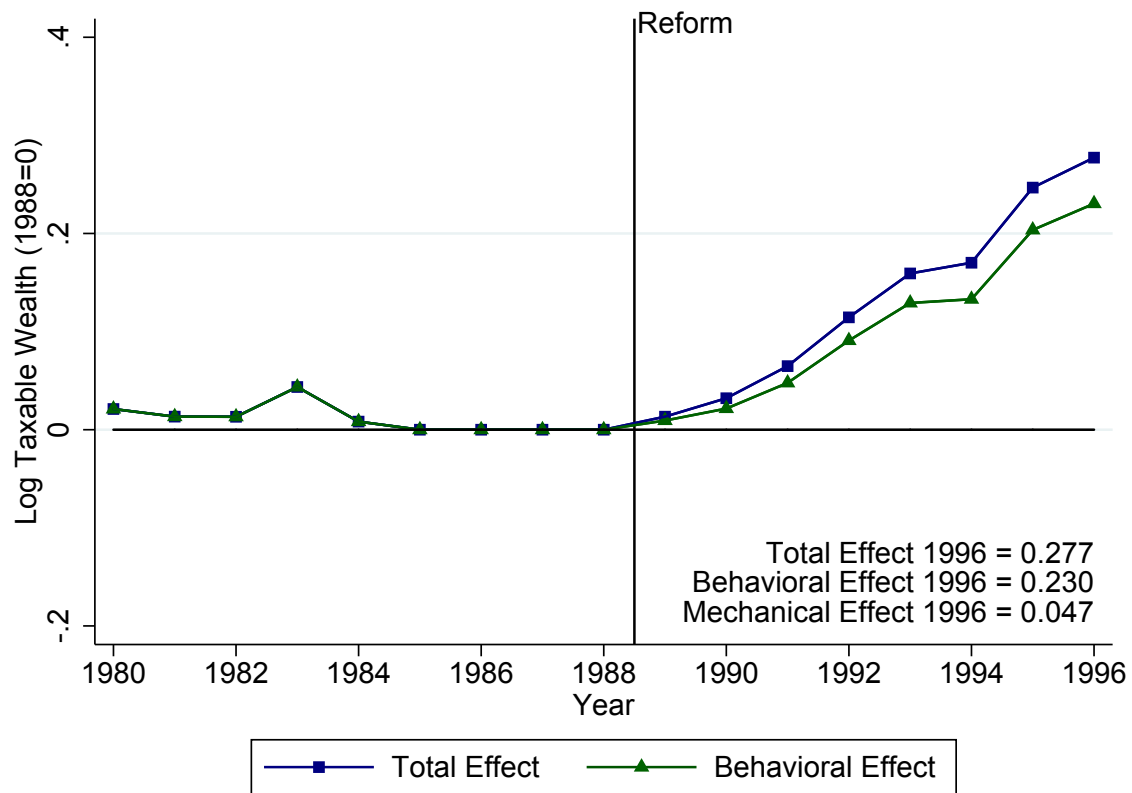

Notes: This figure shows the effects of the 1989 wealth tax reform on taxable wealth based on the couples DD in which we compare couples within and below the exempted wealth range. The assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. The sample is a balanced panel of households observed in all years 1980-1996 and located within the top 5% of the wealth distribution before the reform. Panel A shows the evolution of log taxable wealth in the two comparison groups, normalized to zero in the pre-reform year 1988. Panel A shows the evolution of log taxable wealth in the two comparison groups, normalized to zero in the pre-reform year 1988. Panel B shows the raw differences between these two series and the pre-trend adjusted differences (using equation 2). The 95% confidence intervals are based on robust standard errors clustered at the household level.

FIGURE A.VIII: Couples DD Using Couples Below Threshold as Controls: Treatment Effect on the Treated
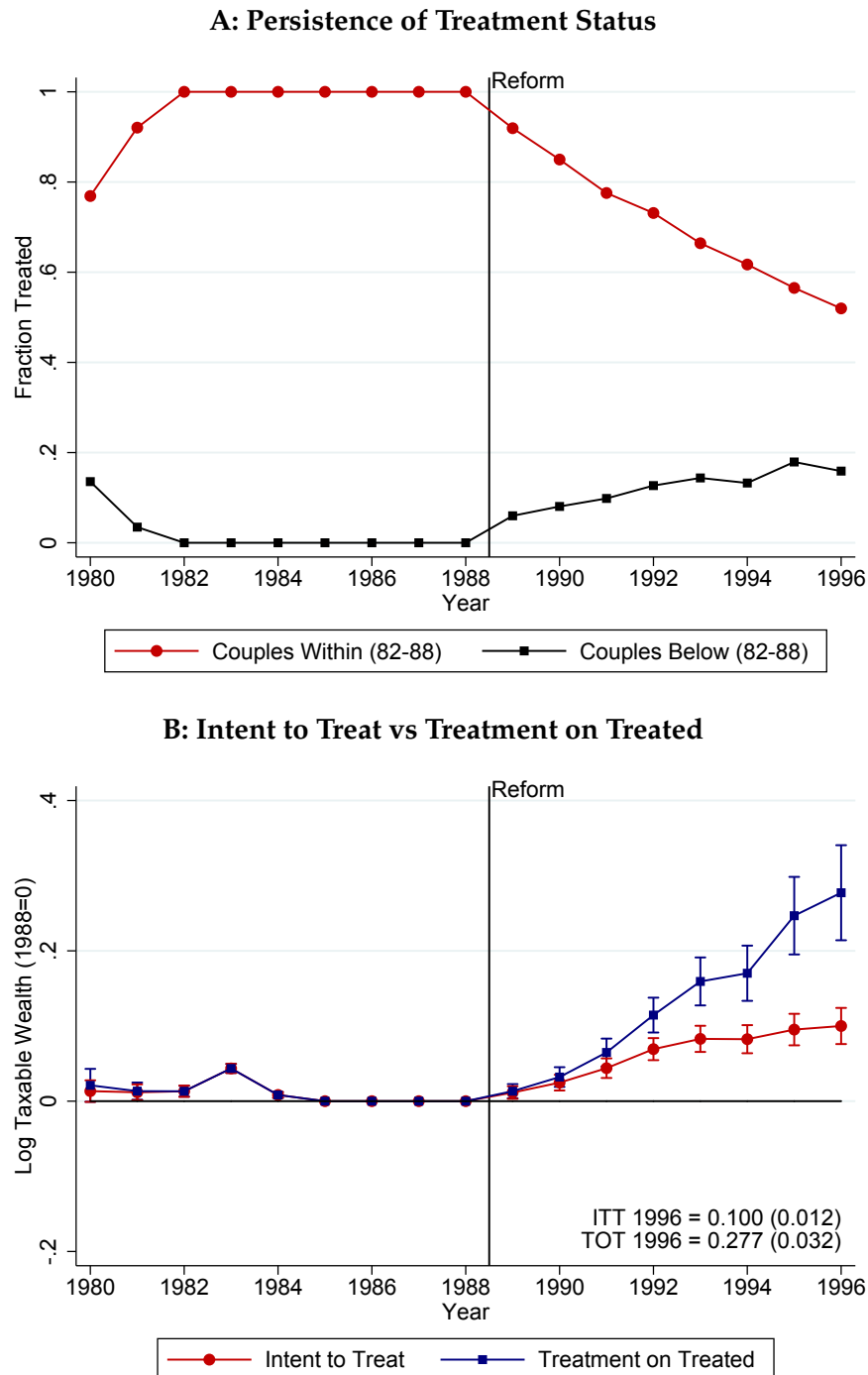
A: Persistence of Treatment Status

B: Intent to Treat vs Treatment on Treated

ITT 1996 = 0.100 (0.012)
TOT 1996 = 0.277 (0.032)

Notes: This figure converts intention-to-treat (ITT) estimates into treatment-on-the-treated (TOT) estimates based on the couples DD in which we compare couples within and below the exempted wealth range. The assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. Panel A shows the persistence in treatment status over time. By construction, couples in the exempted range have a treatment status of 100% before the reform, while couples below this range have a treatment status of 0% before the reform. After the reform, taxpayers may switch status due to changes in relationship status or changes in wealth bracket. Panel B compares the ITT and TOT series, where the TOT estimates are obtained by instrumenting treatment status in equation (2) using treatment status in the pre-reform years 1982-88.

FIGURE A.IX: Couples DD Using Couples Below Threshold as Controls: Behavioral vs Mechanical Effects

Notes: This figure shows the composition of the total effect of wealth tax reform into behavioral and mechanical effects for the couples DD in which we compare couples within and below the exempted wealth range. Appendix B provides details on how the mechanical effects are calculated. The behavioral effect equals the total effect minus the mechanical effect.

FIGURE A.X: Couples DD Using Singles Below Threshold as Controls

A: Time Series of Couples Within and Singles Below Range
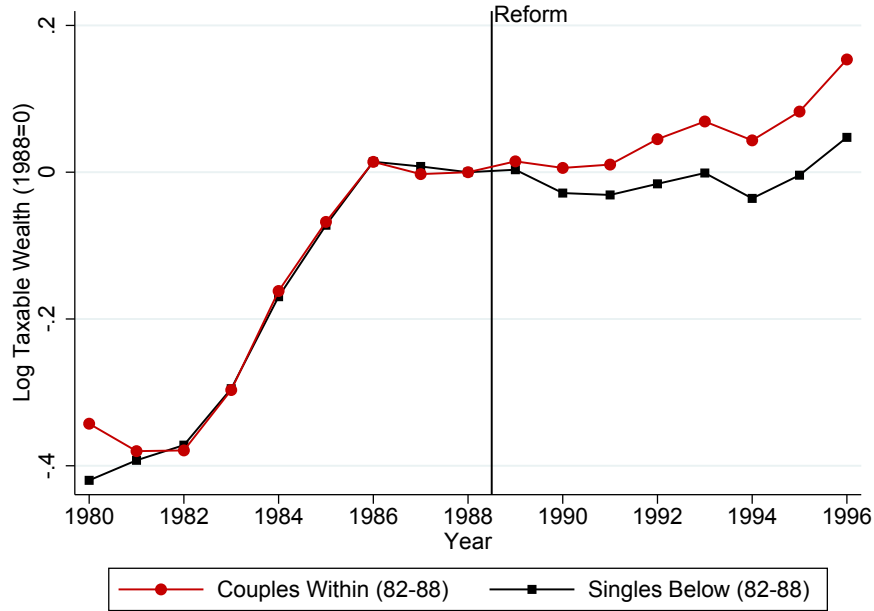
B: Difference Between Couples Within and Singles Below Range

Notes: This figure shows the effects of the 1989 wealth tax reform on taxable wealth based on a couples DD in which we compare couples within the exempted wealth range to singles below the exempted wealth range (specifically, singles located between the exemption threshold and half the exemption threshold). The assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. The sample is a balanced panel of households observed in all years 1980-1996. Panel A shows the evolution of log taxable wealth in the two comparison groups, normalized to zero in the pre-reform year 1988. Panel B shows the differences between these two series, i.e. our reduced-form or intention-to-treat (ITT) estimates. The 95% confidence intervals are based on robust standard errors clustered at the household level.

FIGURE A.XI: Couples DD Using Singles Below Threshold as Controls: Treatment Effect on the Treated
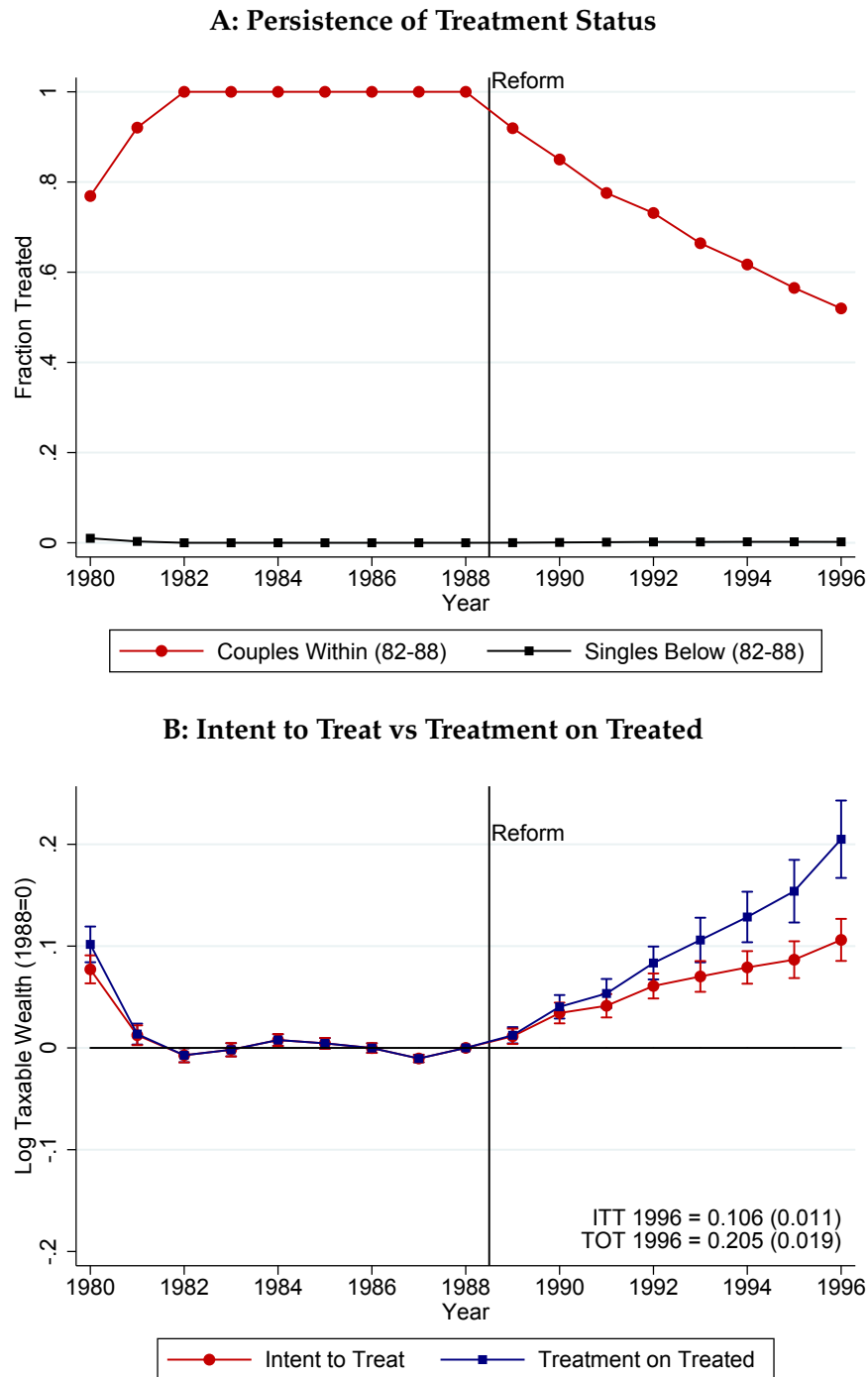
A: Persistence of Treatment Status

B: Intent to Treat vs Treatment on Treated


ITT 1996 = 0.106 (0.011)
TOT 1996 = 0.205 (0.019)

Notes: This figure converts intention-to-treat (ITT) estimates into treatment-on-the-treated (TOT) estimates based on the couples DD in which we compare couples within and singles below the exempted wealth range. The assignment to treatment and control groups is based on pre-reform variables, restricting attention to households whose status stays constant during 1982-88. Panel A shows the persistence in treatment status over time. By construction, couples in the exempted range have a treatment status of 100% before the reform, while singles below this range have a treatment status of 0% before the reform. After the reform, taxpayers may switch status due to changes in relationship status or changes in wealth bracket. Panel B compares the ITT and TOT series, where the TOT estimates are obtained by instrumenting treatment status in equation (1) using treatment status in the pre-reform years 1982-88.

FIGURE A.XII: Couples DD Using Singles Below Threshold as Controls: Behavioral vs Mechanical Effects
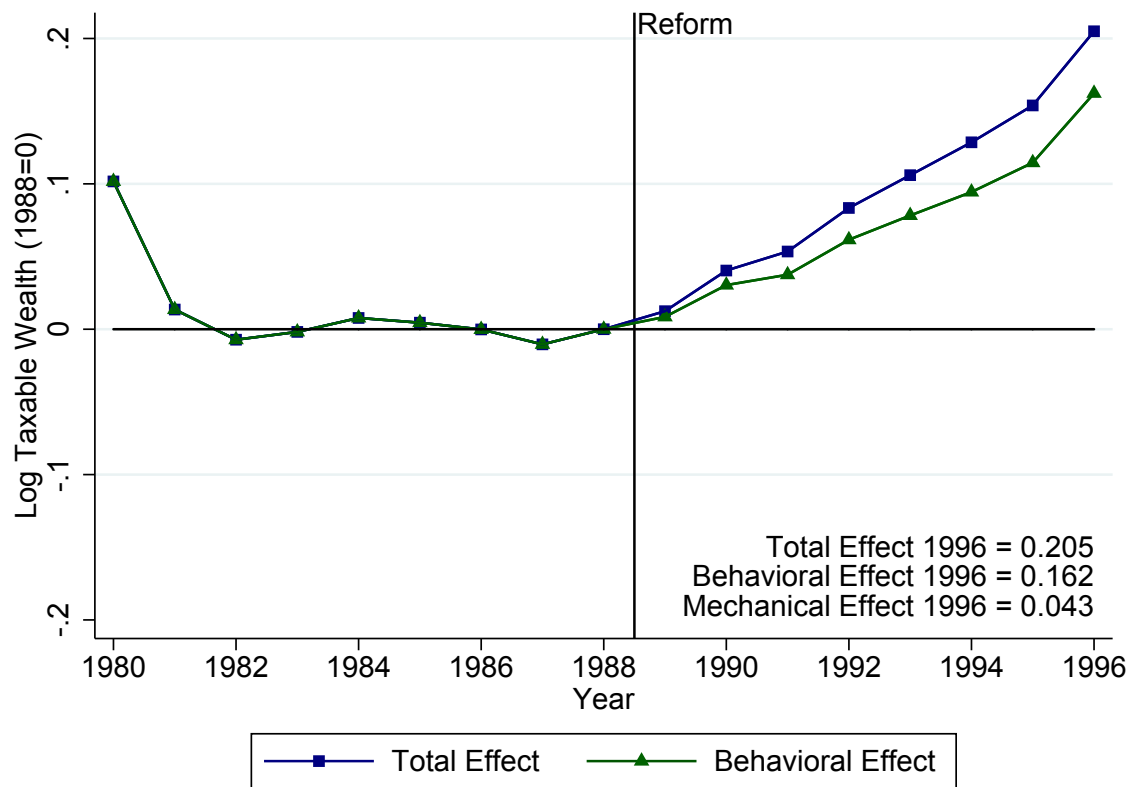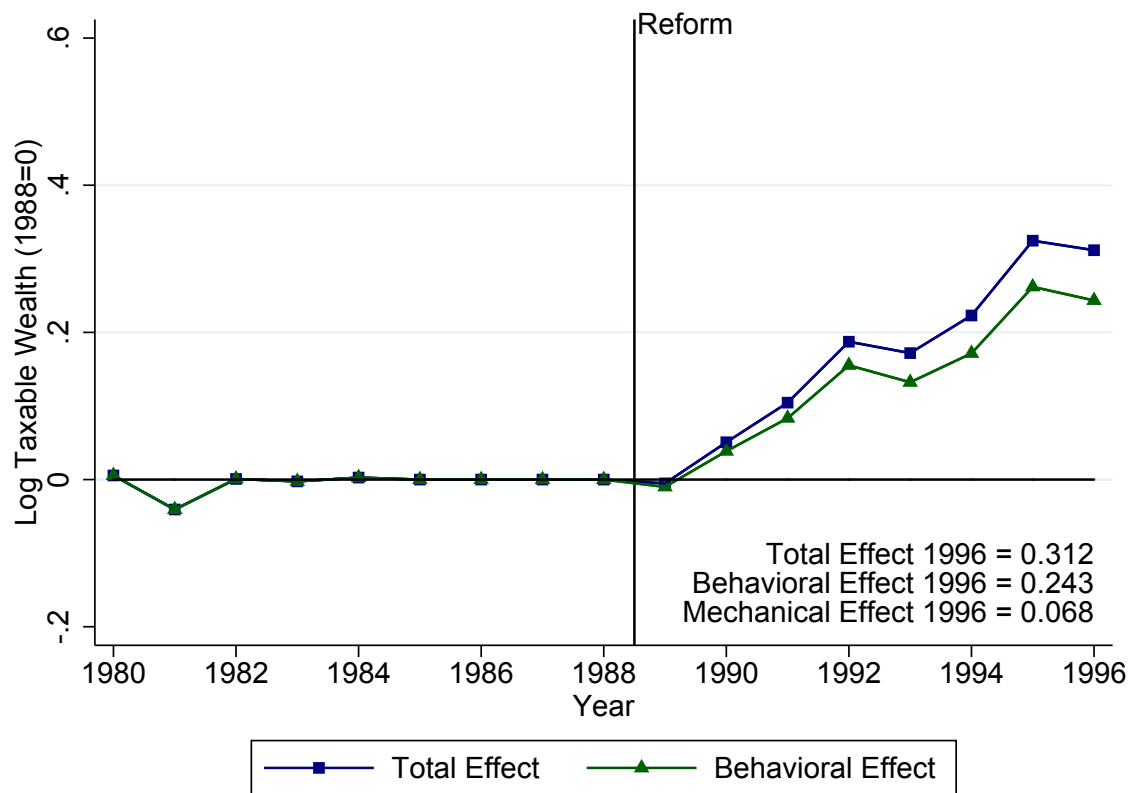

Notes: This figure shows the composition of the total effect of wealth tax reform into behavioral and mechanical effects for the couples DD in which we compare couples within and singles below the exempted wealth range. Appendix B provides details on how the mechanical effects are calculated. The behavioral effect equals the total effect minus the mechanical effect.

# FIGURE A.XIII: Fraction Married Above and Below the Exemption Threshold
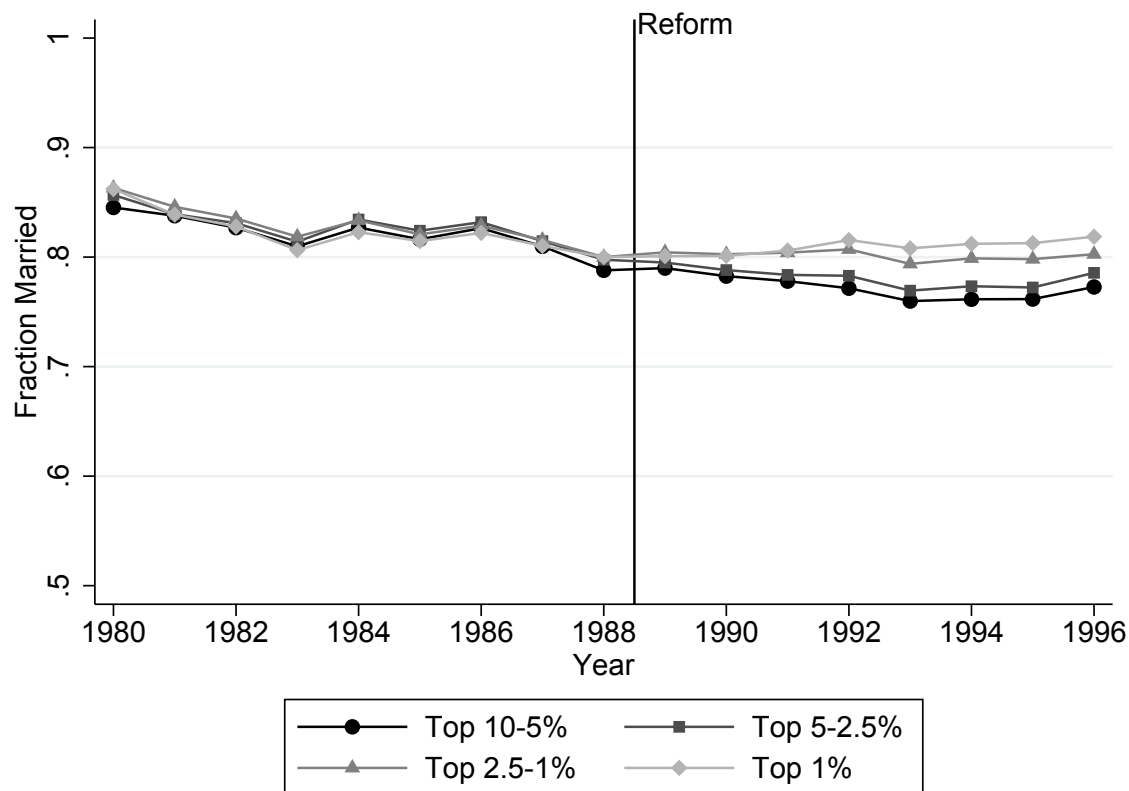
Notes: This figure shows the marriage rate in different quantiles of the wealth distribution. The top 1% and top 2.5-1% are above the exemption threshold (and have a stronger incentive to marry after the wealth tax reform), while the top 5-2.5% and top 10-5% are below the exemption threshold (and have an unchanged incentive to marry after the reform). The figure is based on repeated cross-sections of individuals observed in 1980-1996.

FIGURE A.XIV: Ceiling DD: Behavioral vs Mechanical Effects

Total Effect 1996 = 0.312
Behavioral Effect 1996 = 0.243
Mechanical Effect 1996 = 0.068

Notes: This figure shows the composition of the total effect of wealth tax reform into behavioral and mechanical effects for the ceiling DD in which we compare households bound and unbound by the tax ceiling. Appendix B provides details on how the mechanical effects are calculated. The behavioral effect equals the total effect minus the mechanical effect.

## FIGURE A.XV: Ceiling DD: Robustness to Treatment Window
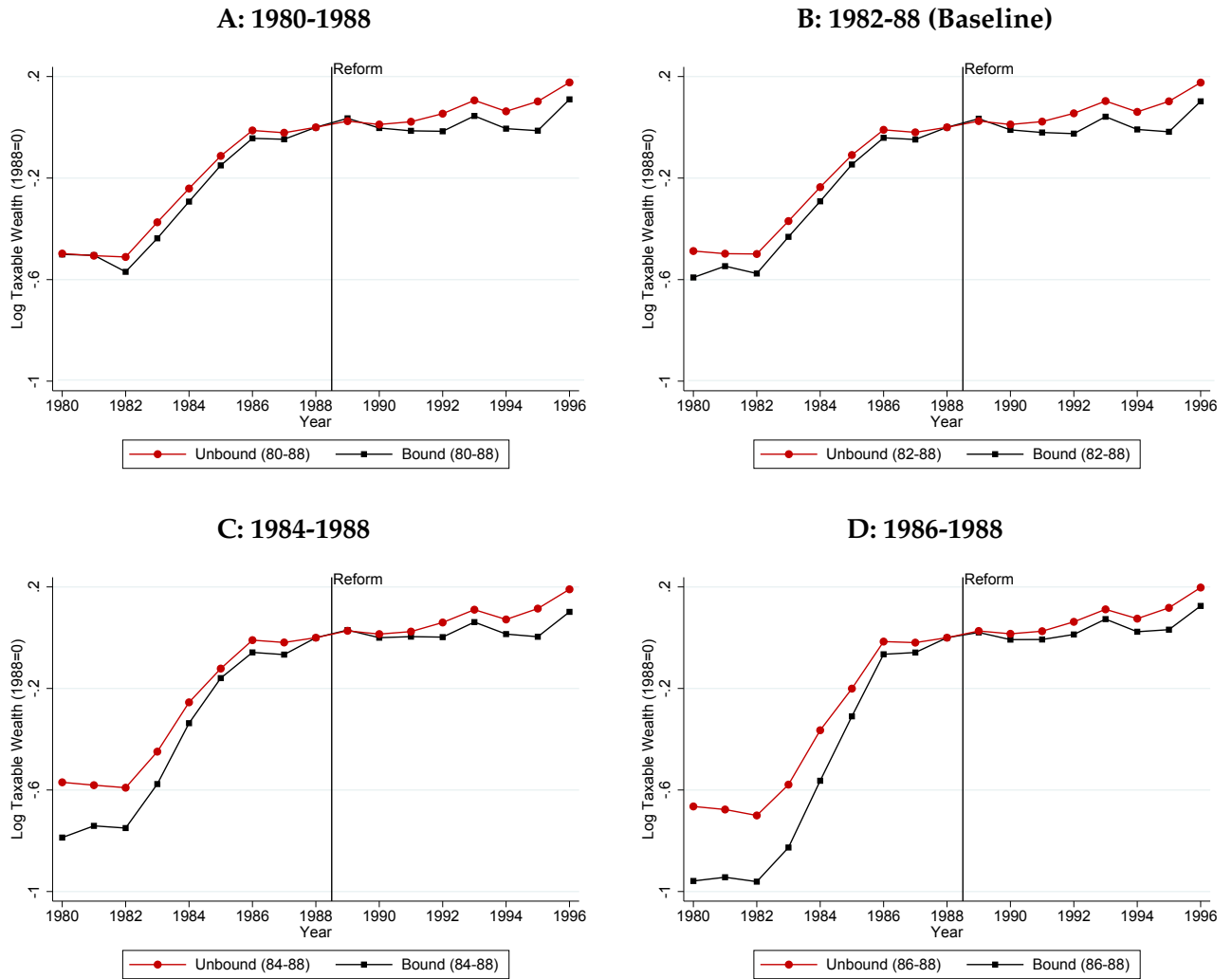
Notes: This figure investigates the robustness of the ceiling DD results to defining comparison groups (households bound and unbound by the tax ceiling) based on outcomes in specific pre-reform years. The four panels show the effects on log taxable wealth when using different pre-reform windows to define treatment status: 1980-88, 1982-88 (baseline specification), 1984-88, and 1986-88. Besides varying the pre-reform treatment window, each panel is constructed in the same way as Panel A of Figure VI. The figure shows that the results are similar across the alternative specifications.

FIGURE A.XVI: Ceiling DD: Placebo Tests

A: Placebo Reform 1983

B: Placebo Reform 1984

### C: Placebo Reform 1985

D: Placebo Reform 1986
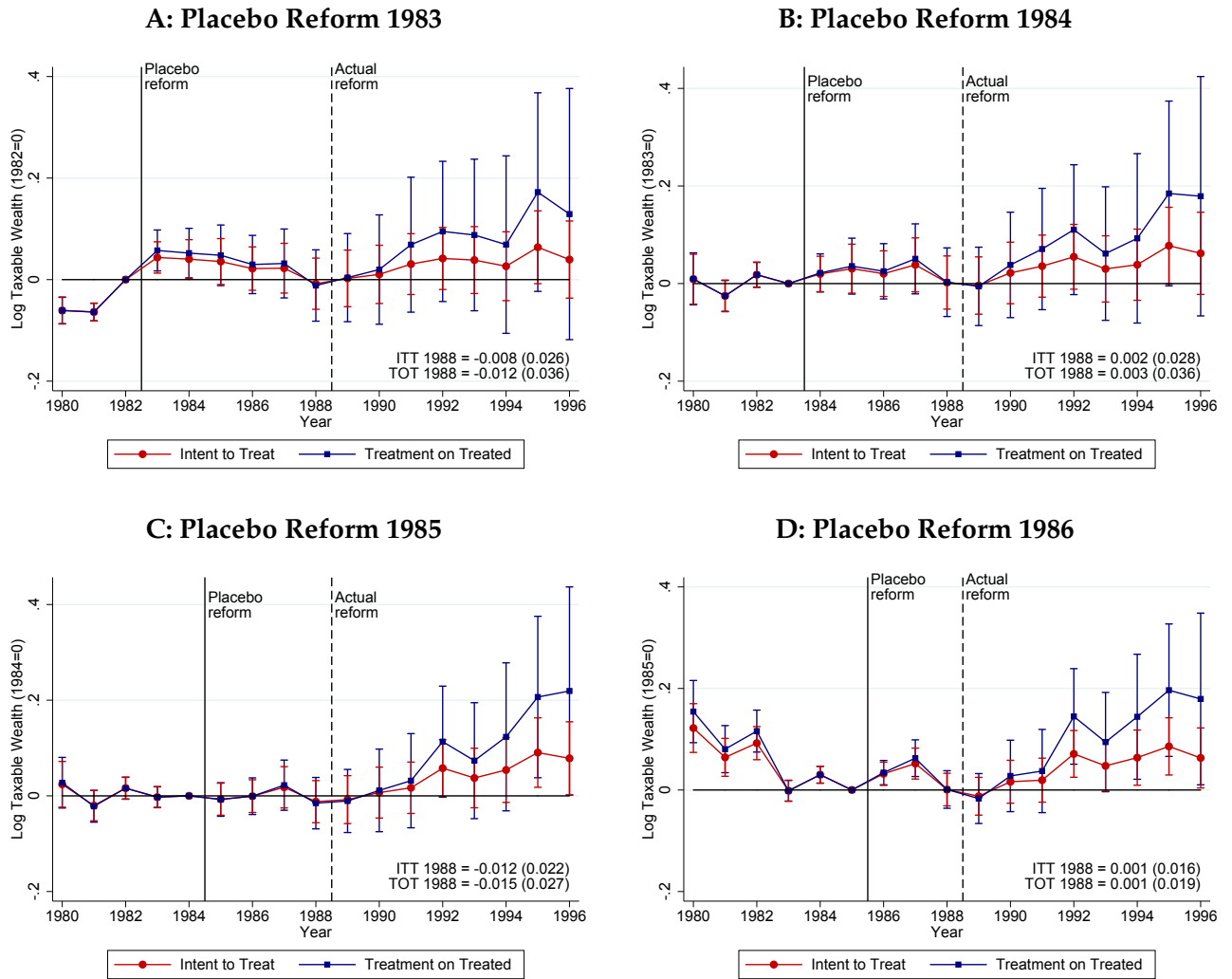
Notes: This figure shows ITT and TOT estimates obtained from the ceiling DD and a set of placebo reforms. We consider placebo reforms implemented in earlier years (1983, 1984, 1985, and 1986) than the actual reform (1989). The comparison groups are households who are bound and unbound by the tax ceiling as in the main specification, but the assignment of treatment status is based outcomes prior to the placebo reforms rather than the actual reform. We use three pre-reform years to assign treatment status (e.g., 1980-82 for the 1983 placebo reform). In all other respects, the figure is constructed in the same way as Panel B of Figure VII.

FIGURE A.XVII: Empirical Lifecycle Profiles of Wealth at the Extreme Top (in Terms of Percentiles)
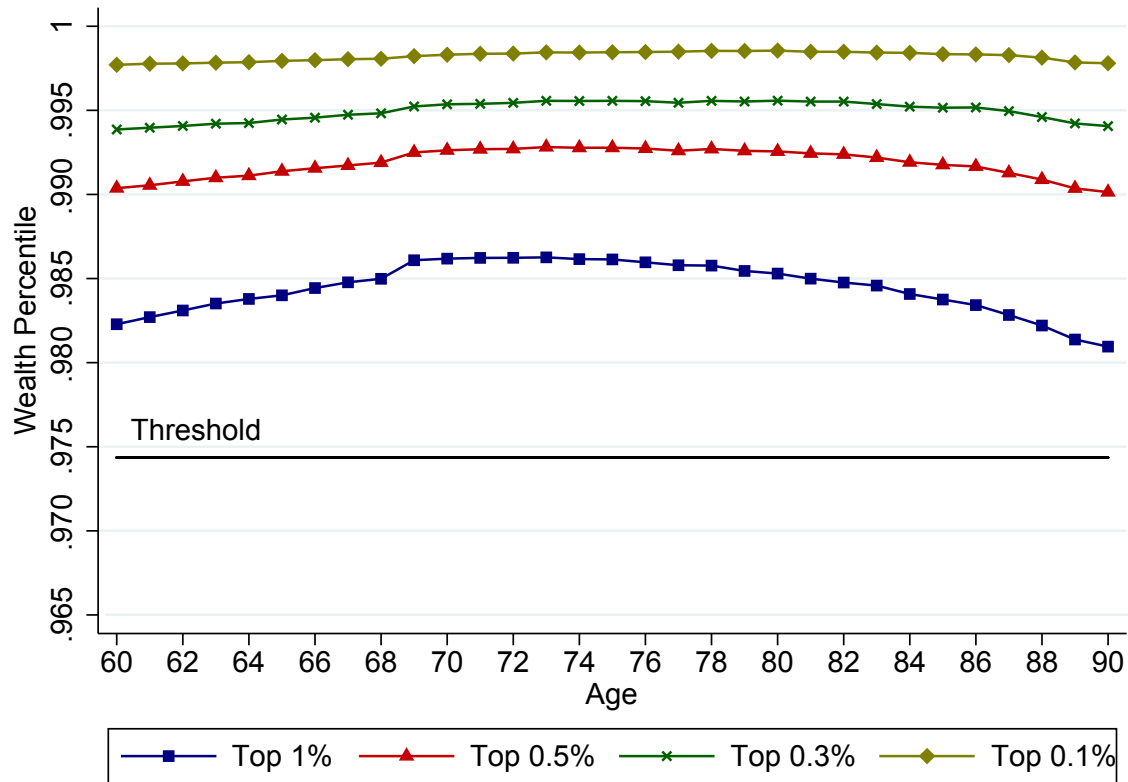
Notes: This figure shows lifecycle profiles of taxable wealth between the ages of 60-90 in an unbalanced panel of individuals over the period 1980-2012. The figure is constructed in the same way as Figure VIII, except that we include the extreme top of the distribution (the top 0.3% and 0.1%) and that 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.
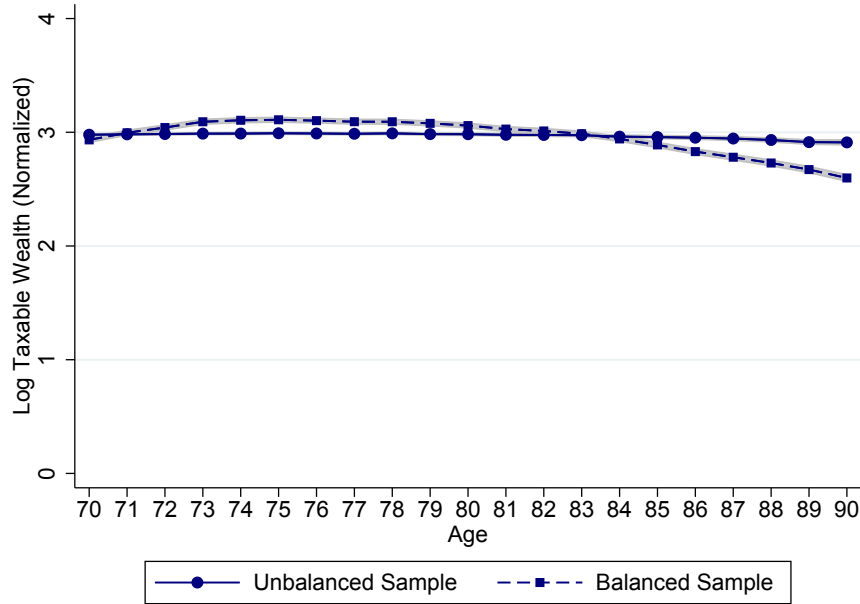
FIGURE A.XVIII: Empirical Lifecycle Profiles of Wealth for the Top 1% Before and After the Reform
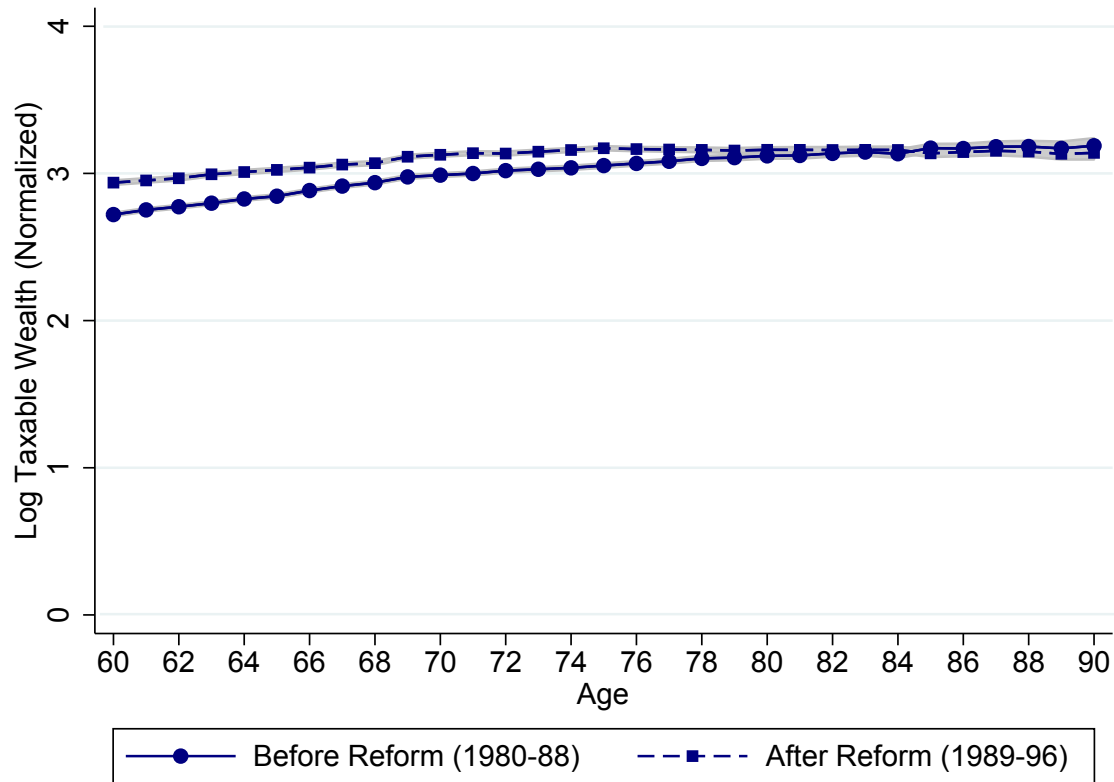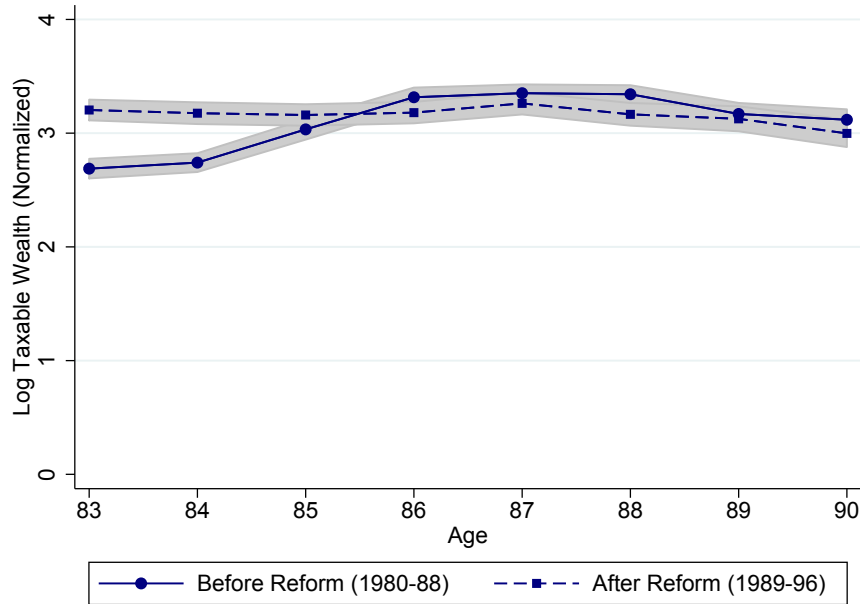

Notes: The figure shows lifecycle profiles of taxable wealth between the ages of 60-90 in an unbalanced panel of individuals before the reform (1980-1988) and after the reform (1989-96), respectively. To eliminate the confounding effects of inflation as people grow older, we normalize log wealth for each individual in each year by the average log 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.

FIGURE A.XIX: Empirical Lifecycle Profiles of Wealth
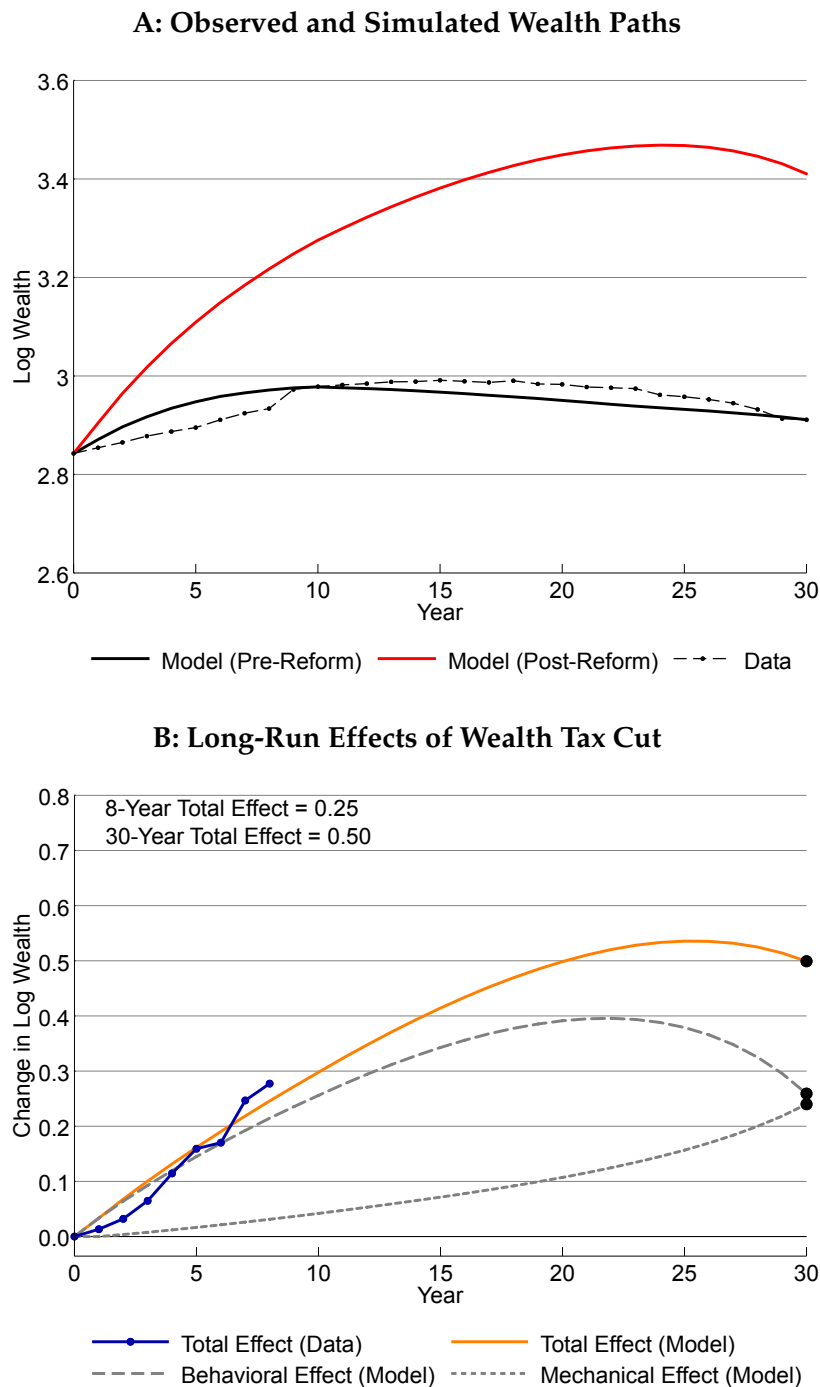
A: Balanced vs Unbalanced

B: Before Reform vs After Reform (Both Balanced)

Notes: This 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.

A: Observed and Simulated Wealth Paths

## B: Long-Run Effects of Wealth Tax Cut

Notes: This figure shows the long-run effects of wealth tax cuts when calibrating our model to the couples DD (comparing couples within and below the exempted range). These are effects for the moderately wealthy (between the 98th and 99th percentile cutoffs). The reform experiment cuts the wealth tax rate by 2.2 percentage points, corresponding to an elimination of the wealth tax for the treatment group with no change for the control group. Panel A shows the observed 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 A.XXI: Calibrating the Bequest Elasticity

A: Couples DD (Moderately Wealthy)
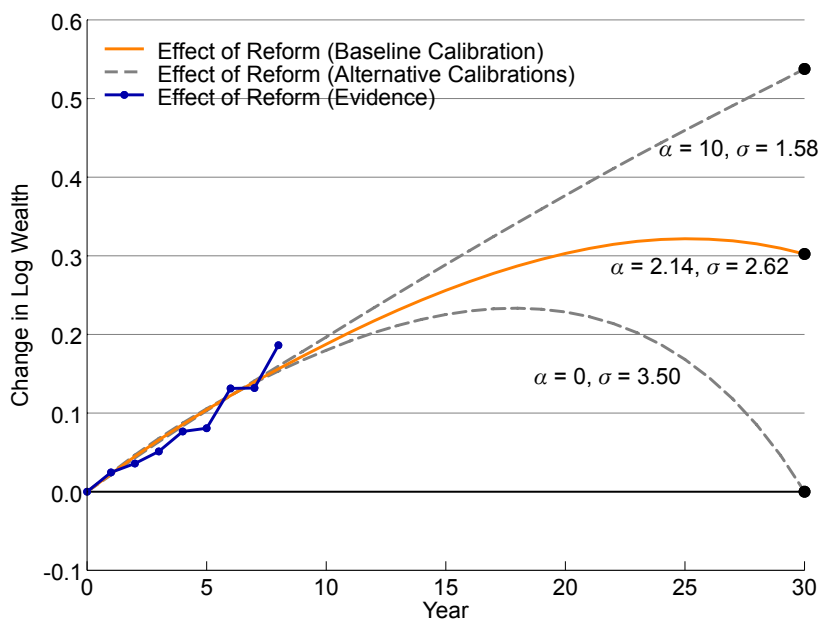
B: Ceiling DD (Very Wealthy)
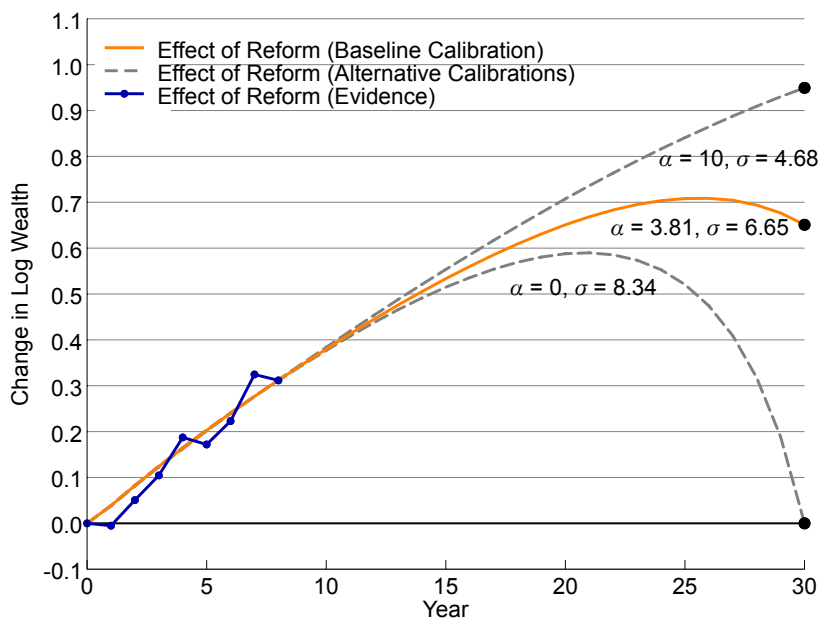
Notes: The figure illustrates how the calibration of the bequest elasticity α (in conjunction with the EIS σ) works. The calibration has to satisfy two requirements: (i) it has to fit the quasi-experimental estimates for years 1,...8 (based on minimizing equation 19), and (ii) it has to fit the flatness 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).

TABLE A.I: Estimated Wealth Responses for Age Below Median

	Couples DD				Ceiling DD	
	Couples vs Singles		Couples Within vs Below		Unbound vs Bound	
	(1)	(2)	(3)	(4)	(5)	(6)
ITT Effects on Log Wealth:						
Average Effect	0.069	0.063	0.026	0.082	0.067	0.085
	(0.012)	(0.014)	(0.007)	(0.009)	(0.027)	(0.043)
1996 Effect	0.113	0.107	0.042	0.132	0.102	0.141
	(0.024)	(0.028)	(0.013)	(0.017)	(0.042)	(0.070)
TOT Effects on Log Wealth:						
Average Effect	0.102	0.093	0.059	0.179	0.095	0.122
	(0.018)	(0.021)	(0.015)	(0.019)	(0.037)	(0.058)
1996 Effect	0.193	0.182	0.119	0.377	0.166	0.229
	(0.040)	(0.046)	(0.036)	(0.046)	(0.068)	(0.111)
Elasticities:						
Elasticity wrt. $1 - \tau$	10.299	9.409	2.180	6.850	6.493	8.313
	(1.810)	(1.184)	(0.408)	(0.459)	(1.091)	(1.379)
Elasticity wrt. $(1 - \tau)R - 1$	0.421	0.384	0.067	0.212	0.214	0.274
	(0.074)	(0.048)	(0.013)	(0.014)	(0.036)	(0.045)
Observations	51,425	51,425	89,148	89,148	45,475	45,475
Household and Year FE	X	X	X	X	X	X
Linear Pre-trends		X		X		X

Notes: This table shows the intent-to-treat (ITT) and treatment-on-the-treated (TOT) effects of the 1989 wealth tax cuts on log wealth as well as the implied elasticities of taxable wealth for those below the median age. Age is defined as the average across spouses within the household. The table is otherwise constructed exactly as Table II.

TABLE A.II: Estimated Wealth Responses for Age Above Median

	Couples DD				Ceiling DD	
	Couples vs Singles		Couples Within vs Below		Unbound vs Bound	
	(1)	(2)	(3)	(4)	(5)	(6)
ITT Effects on Log Wealth:						
Average Effect	0.017	-0.008	0.009	0.053	0.091	0.151
	(0.010)	(0.011)	(0.008)	(0.009)	(0.032)	(0.089)
1996 Effect	0.021	-0.016	0.017	0.084	0.126	0.242
	(0.021)	(0.023)	(0.016)	(0.017)	(0.055)	(0.089)
TOT Effects on Log Wealth:						
Average Effect	0.027	-0.015	0.021	0.108	0.199	0.333
	(0.017)	(0.019)	(0.015)	(0.017)	(0.068)	(0.106)
1996 Effect	0.049	-0.038	0.044	0.222	0.335	0.643
	(0.047)	(0.051)	(0.040)	(0.044)	(0.144)	(0.235)
Elasticities:						
Elasticity wrt. $1 - \tau$	2.874	-1.287	0.679	4.018	12.382	20.663
	(1.690)	(1.184)	(0.381)	(0.401)	(1.631)	(2.049)
Elasticity wrt. $(1 - \tau)R - 1$	0.117	-0.052	0.025	0.149	0.409	0.683
	(0.069)	(0.048)	(0.014)	(0.015)	(0.054)	(0.068)
Observations	53,414	53,414	93,687	93,687	49,028	49,028
Household and Year FE	X	X	X	X	X	X
Linear Pre-trends		X		X		X

Notes: This table shows the intent-to-treat (ITT) and treatment-on-the-treated (TOT) effects of the 1989 wealth tax cuts on log wealth as well as the implied elasticities of taxable wealth for those above the median age. Age is defined as the average across spouses within the household. The table is otherwise constructed exactly as Table II.