Online Appendix for:
Behavioral Responses to Wealth Taxes: Evidence from Sweden

David Seim[*]

---

[*]Department of Economics, Stockholm University; CEPR and Research Institute of Industrial Economics, Stockholm, david.seim@ne.su.se.

A Tax-Filing Forms

FIGURE A.1: PREPOPULATED TAX RETURN.

Notes: Section 5 in the figure displays the filing of the wealth tax. The section *Tillgångar* refers to assets. Taxpayers were supposed to fill in the total value of taxable assets in field 66 if their taxable net wealth exceeded the threshold. The section *Skulder* refers to liabilities. Taxpayers filled in the total value of liabilities in field 67.

FIGURE A.2: FORMULA FOR COMPUTATION OF TAX LIABILITY.

Så här går det till att deklarera

⑤ Förmögenhet (forts.)

Förmögenhetsuppgifter
På specifikationen som du har fått tillsammans med deklarationsblanketten finns de tillgångar och skulder som skattemyndigheten har uppgift om. Om du enligt dessa uppgifter har en beskattningsbar förmögenhet har skattemyndigheten fyllt i beloppet på din deklarationsblankett. Förmögenhetsuppgifter har också fyllts i om du hade beskattningsbar förmögenhet i föregående års deklaration.

Vissa tillgångar är inte med på specifikationen, t.ex. bilar och fritidsbåtar. Om du har sådana tillgångar måste du alltså lägga till värdet av dessa när du beräknar din förmögenhet.

Du får endast dra av skulder som avser anskaffning av tillgångar som ska tas med när du beräknar din skattepliktiga förmögenhet. Om en skuld avser skattepliktiga aktier får du göra avdrag för den, men däremot inte om den avser aktier som inte är skattepliktiga, t.ex. aktier på O-listan. Skattemyndigheten har inte uppgift om vad dina skulder avser. Detta innebär att om det på specifikationen finns skulder som du inte kan göra avdrag för måste du dra ifrån värdet av dessa när du beräknar din förmögenhet.

Observera!
Förmögenhetsskatten ingår inte i den preliminära skatteuträkning som du fått tillsammans med deklarationsblanketten.

Här kan du räkna ut din förmögenhet

	Person 1	Person 2	Hemmavarande barn som lämnar egen deklaration
Fastighet/Bostadsrätt			
Taxeringsvärde på småhus	+	+	+
Förmögenhetsvärde på bostadsrätt	+	+	+
Bankmedel/Värdepapper			
Bankmedel	+	+	+
Aktier	+	+	+
Värdepappersfonder	+	+	+
Obligationer	+	+	+
Övriga värdepapper	+	+	+
Övriga tillgångar			
Bilar och andra yttre inventarier	+	+	+
Kapitalförsäkringar	+	+	+
Förmögenhet för hemmavarande barn som inte lämnar deklaration	+	+	Lämnar egen deklaration
Summa tillgångar den 31 december 2002	=	=	=
Skulder den 31 december 2002	–	–	–
Skattepliktig förmögenhet (+) Underskott (-)	=	=	=

Läs mer om förmögenhet i broschyren "Skatteregler för privatpersoner" (RSV 330). **Beställ på direktval 7103.** 11

Notes: This formula was appended to the prepopulated tax return. Households were supposed to use this to compute wealth tax liabilities.

B What Does Bunching Capture?

The purpose of this section is to show what bunching can capture in the short vs long run. I start by laying out a static model, assuming that in the short run, agents can only respond through avoidance/evasion-technologies. The long-run response, however, is a function of both avoidance/evasion-technologies and savings responses.

Consider an agent who faces the following short-run maximization problem:

$$\max_{e \leq s} (1 - \tau)(s - e) + e - C(e, s),\tag{1}$$

where s denotes exogenous savings or wealth, e tax sheltering activities and τ a linear tax. The cost of evasion, $C(e, s)$ is assumed to be increasing in evasion and decreasing in savings.[1] In the application to the Swedish wealth tax, s denotes true wealth and $s - e$ taxable (net) wealth. Following Slemrod (2001), I parametrize the cost function as follows:

$$C(e, s) = \left(\frac{e}{s}\right)^{\frac{1}{\gamma}} \frac{pe}{1 + \frac{1}{\gamma}},\tag{2}$$

where γ is the constant tax elasticity of evasion, and p is a parameter which can be interpreted as a linear penalty including fines, costs of going to court and other transfers.[2] The cost function is not specific to the wealth tax and Slemrod (2001) employs it in a labor-income tax setting. The agent's solution to problem (1) is given by

$$e^* = \left(\frac{\tau}{p}\right)^{\gamma} s,\tag{3}$$

and the agent's taxable net wealth by

$$s - e^* = \left(1 - \left(\frac{\tau}{p}\right)^{\gamma}\right) s.\tag{4}$$

The cost function implies that evaded amounts are proportional to true wealth. Assuming that s is distributed according to some continuous and differentiable CDF $F(s)$, the choices of e^* under a linear tax imply that the distribution of taxable net wealth is also described by a continuous and differentiable function $H(s - e)$.

If a kink is introduced in the budget set at taxable net wealth z^*, so that $\tau = \tau_0$ below the kink and $\tau = \tau_1 > \tau_0$ above the kink, agents who chose a taxable net wealth level in $[z^*, z^* + \Delta z]$ will now choose to locate exactly at the kink point. The agent with the highest savings level s, who is bunching, had a taxable net wealth under the linear tax rate given by $(s - e)^U = z^* (1 - (\tau_0/p)^{\gamma}) / (1 - (\tau_1/p)^{\gamma})$. Hence, the number of households that bunch at the kink point is given by $H\left((s - e)^U\right) - H(z^*)$. Using the counterfactual density together with the fact that $\log(1 + x) \approx x$ for small x, I obtain:

$$\frac{B}{h(z^*) z^*} \approx \log\left(\frac{1 - \left(\frac{\tau_0}{p}\right)^{\gamma}}{1 - \left(\frac{\tau_1}{p}\right)^{\gamma}}\right),\tag{5}$$

where B denotes the number of households that bunch at the kink point, and $h(z^*)$ is the density at the kink point under the linear tax scheme. Thus, $B/h(z^*)$ is the mass at the

---

[1] Allingham and Sandmo (1972) formulate this problem as a gamble. Mayshar (1991) shows that the gamble can be represented with a monetary cost of evasion, where the cost is the certainty-equivalent of the gamble that causes extra utility loss as the risk of audit is increased.

[2] Yitzhaki (1974) shows that if p is linear in τ, the tax rate has no effect on sheltering behavior.

kink point in excess of the counterfactual density. If $\tau_0 = 0$, I approximate equation (5) as

$$\frac{B}{h\left(z^*\right) z^*} \approx \left(\frac{\tau_1}{p}\right)^\gamma. \tag{6}$$

Equation (6) interprets bunching as the (constant) fraction of savings/wealth that is evaded. Notice that as the tax rate increases, the fraction of evaded savings goes up and more households bunch at the kink.

Now, augment that model to allow agents to also respond to tax changes by way of savings.[3] In the dynamic model, agents trade off consumption in two periods. Agents pay a tax on their accumulated assets (which in this framework are equal to savings). However, they can shelter money from the government through the same technology as above.

The agent faces the following maximization problem:

$$\max_{s,e \leq s} U\left(c_1, c_2\right) = \max_{s,e \leq s} \frac{c_1^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} + \beta \frac{c_2^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} \tag{7}$$

subject to

$$c_1 = y - s$$

$$c_2 = (1 - \tau)(s - e) + e - \left(\frac{e}{s}\right)^{\frac{1}{\gamma}} \frac{pe}{1 + \frac{1}{\gamma}},$$

where c_t is consumption in period t, β the discount factor, σ the elasticity of intertemporal substitution and y is heterogenous income distributed according to a continuous and differentiable CDF $G\left(y\right)$.[4] The first-order condition that governs the tax sheltering response is given by equation (3), restated here for convenience: $e^* = (\tau/p)^\gamma s$. Substituting this into the Euler equation which determines the savings response to the tax, I obtain:

$$c_1^{-\frac{1}{\sigma}} = \beta \left(1 - \tau \left(1 - \left(\frac{\tau}{p}\right)^\gamma \frac{1}{1 + \gamma}\right)\right)^{1 - \frac{1}{\sigma}} c_2^{-\frac{1}{\sigma}}. \tag{8}$$

An increase in the tax rate has three effects. First, the fraction of savings evaded from tax goes up. The magnitude of this response is given by the structural parameter γ and the penalty cost, p. Second, the return to saving is negatively affected by a tax increase and parameter σ determines the relative importance of the income and substitution effects associated with a tax increase. With $\sigma < 1$, the income effect dominates the substitution effect.[5] An increase in the tax rate actually raises savings. When $\sigma > 1$, the substitution effect dominates the income effect and an increase in the tax rate lowers savings. Third, the cost function possesses the feature that higher savings lower the marginal cost of evasion. Slemrod (2001) refers to this as the *avoidance-facilitating* effect. The distortionary effect

---

[3]The model abstracts from labor supply responses to the wealth tax.

[4]For simplicity, I assume that the gross interest rate is zero, but this can easily be relaxed. In the estimation procedure, the choice of the interest rate does not have a large effect on estimated entities.

[5]However, both the uncompensated and income-compensated effects on consumption by an increase in the tax rate are negative.

of an increased tax rate on savings is thus attenuated by agents evading a fraction of their savings.

In the general version of this economy, the Euler equation determines the balanced growth path. From a growth-enhancing policy perspective, tax evasion thus weakens the distortionary effects of the tax on long-run growth.

The agent chooses s^* according to

$$s^* = f(\tau) y, \tag{9}$$

where

$$f(\tau) = \frac{\beta^\sigma \left(1 - \tau \left(1 - \left(\frac{\tau}{p}\right)^\gamma \frac{1}{1+\gamma}\right)\right)^{\sigma-1}}{1 + \beta^\sigma \left(1 - \tau \left(1 - \left(\frac{\tau}{p}\right)^\gamma \frac{1}{1+\gamma}\right)\right)^{\sigma-1}}, \tag{10}$$

and taxable net wealth becomes

$$s^* - e^* = f(\tau) \left(1 - \left(\frac{\tau}{p}\right)^\gamma\right) y. \tag{11}$$

Taxable net wealth is proportional to exogenous income y. Therefore, it is again distributed according to some continuous and differentiable CDF denoted by $K(s-e)$ under the linear tax rate. Increasing the marginal tax rate above threshold z^*, such that $\tau = \tau_0$ for taxable net wealth levels below the kink and $\tau = \tau_1 > \tau_0$ above z^*, leads agents close to the kink to adjust their taxable net wealth levels downwards and bunch at the threshold. This could be done either by savings (real response) or by evasion (reporting response), or a combination of the two.

Identifying the interval of bunchers as in the static case, I can relate the bunching at the kink point to the parameters of the model:

$$\frac{B}{k(z^*) z^*} \approx \left(\frac{f(\tau_0) \left(1 - \left(\frac{\tau_0}{p}\right)^\gamma\right)}{f(\tau_1) \left(1 - \left(\frac{\tau_1}{p}\right)^\gamma\right)} - 1\right). \tag{12}$$

In (12), $k(z^*)$ denotes the density of the distribution of taxable net wealth at the kink point with a linear tax rate. Equation (12) is a generalized version of equation (5). The left-hand side is the excess mass at the kink point z^*. The right-hand side is the interval of taxable net wealth values under the linear tax where wealth holders bunch at the kink point when under the progressive tax.

If $\tau_0 = 0$, as in the Swedish wealth-tax case, the following approximation holds:

$$\frac{B}{z^* k(z^*)} \approx \log\left(\frac{\beta^\sigma}{1 + \beta^\sigma}\right) - \log\left(\frac{\beta^\sigma \left(1 - \tau_1 \left(1 - \left(\frac{\tau_1}{p}\right)^\gamma \frac{1}{1+\gamma}\right)\right)^{\sigma-1}}{1 + \beta^\sigma \left(1 - \tau_1 \left(1 - \left(\frac{\tau_1}{p}\right)^\gamma \frac{1}{1+\gamma}\right)\right)^{\sigma-1}}\right) + \left(\frac{\tau_1}{p}\right)^\gamma. \tag{13}$$

The log difference on the right-hand side captures the discrepancy in savings rates between the left and the right side of the threshold. A large positive discrepancy implies more bunching at the kink point. The third term on the right-hand side captures the fraction of

6

evaded savings, where higher evasion adds to bunching.[6] Equation (13) illustrates intuitively how bunching can arise through adjusted savings as well as evasion. In the static model, a higher penalty rate p always implies lower overall bunching. Here, the impact is less clear. Higher penalty rates still lower evasion but raise the difference in savings rates between the two sides of the kink. If $\sigma < 1$, the log-difference is actually negative, and the fraction evaded increases to reconcile the estimated amount of bunching.

According to equation (13), bunching depends on: (i) observable tax parameters (the tax rate and the kink point); (ii) preference parameters determining the real response (the discount factor and the elasticity of intertemporal substitution); and (iii) evasion cost parameters (the convexity of the cost function and the penalty).

The uncompensated tax semi-elasticities of taxable net wealth and evasion now take the following form:

$$\varepsilon_{W,\tau}^{D} = -\left((1-\lambda)(\sigma-1)\left(1-\tau+\tau\frac{1}{1+\gamma}\left(\frac{\tau}{p}\right)^{\gamma}\right)^{-1}(1-f(\tau)) + \lambda\frac{\gamma}{\tau\left(1-\left(\frac{\tau}{p}\right)^{\gamma}\right)}\right) \tag{14}$$

$$\varepsilon_{e,\tau}^{D} = -(1-\lambda)(\sigma-1)\left(1-\tau+\tau\frac{1}{1+\gamma}\left(\frac{\tau}{p}\right)^{\gamma}\right)^{-1}(1-f(\tau)) + \frac{\gamma}{\tau} \tag{15}$$

where $\lambda = (\tau/p)^{\gamma}$ denotes the fraction evaded. These expressions are sums of the real and the evasion response. In fact, the first part on the right-hand side of both equations denotes the tax semi-elasticity of actual wealth, or savings. If the tax rate goes up (or the penalty rate goes down), agents evade more (λ goes up) and the elasticity of taxable net wealth is relatively more affected by evasion than savings. If $\sigma < 1$, the income effect is stronger than the substitution effect and agents save a larger fraction of income upon the tax change. This effect arises as agents are only aiming at consumption smoothing.

---

[6]Since the tax rate is zero to the left of the threshold, there is no evasion behavior among households to the left of the kink. If the tax rate to the left of the kink was positive, the amount of bunching would be increasing in the difference between evasion rates on the two sides of the threshold.

C   Bunching Graphs

FIGURE C.1: SENSITIVITY OF BUNCHING ESTIMATES TO FUNCTIONAL-FORM ASSUMPTIONS
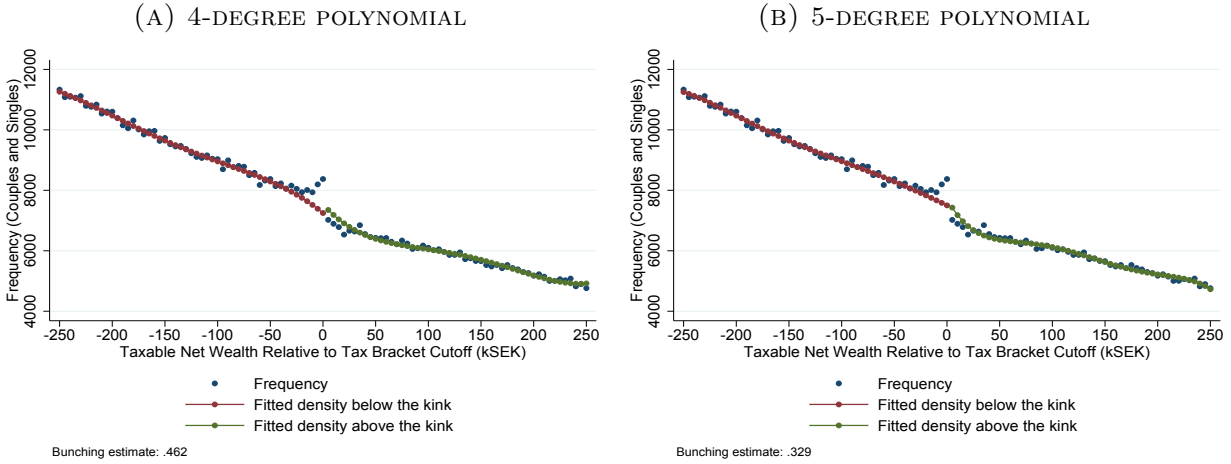
(A) 4-DEGREE POLYNOMIAL   (B) 5-DEGREE POLYNOMIAL

Notes: The figure shows the distribution of taxable net wealth around the shift in the tax brackets, for the years 2000-2006. The dotted series consist of a histogram relative to the normalized kink point. Each bin corresponds to the number of households within SEK 5,000. The estimated counterfactual density, is estimated separately to the left and the right of the threshold. Panel A estimates a 4-degree polynomial on either side of the threshold, while Panel B fits a 5-degree polynomial to the data.

FIGURE C.2: DOES BUNCHING TRACK THE TAX? BUNCHING IN 2001 AND 2006.

(A) 2001   (B) 2006
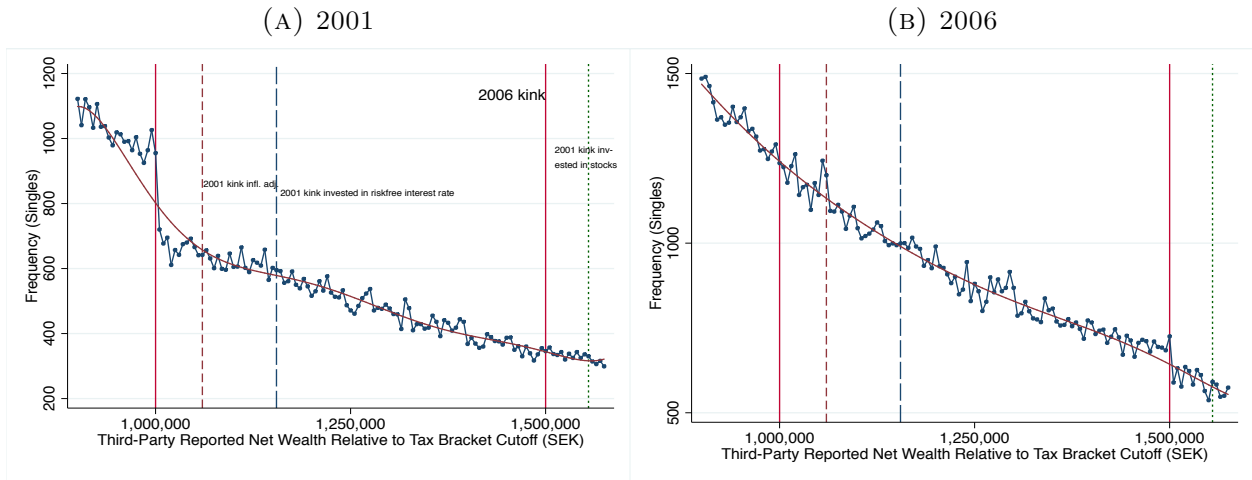
Notes: These figures present the taxable net wealth distribution for singles in 2001 and 2006. The figure shows the kinks in 2001 and 2006, located at SEK 1 million and SEK 1.5 million, respectively. The additional vertical lines represent the position of the 2001-kink if it followed - from the left to the right - inflation, the riskfree interest rate or a stock market index return, respectively. The inflation data was obtained from Statistics Sweden, the riskfree interest rate and the stock market index return from Sveriges Riksbank.

FIGURE C.3: BUNCHING OVER TIME - COUPLES.

(A)
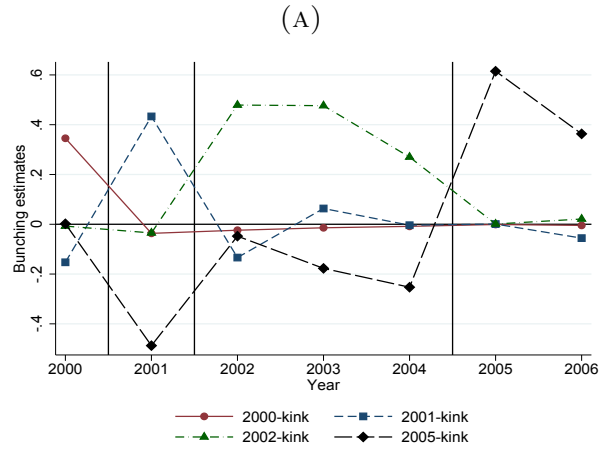
Notes: These graphs show bunching estimates – computed by the non-parametric method – over time at different kinks for couples.

FIGURE C.4: BUNCHING INCLUDING HOUSEHOLDS WITH CHILDREN.

(A)
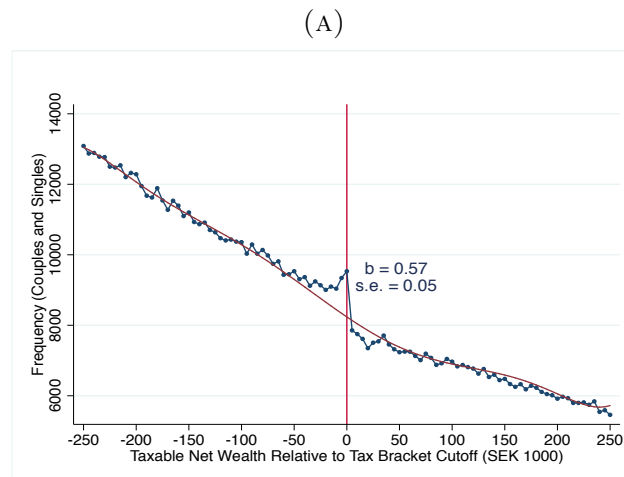
Notes: This figure presents the distribution of taxable net wealth around the exemption threshold for all tax payers, including households with children. The tax rate is 0 below the threshold and 1.5 above. The dotted series consist of a histogram relative to the normalized kink point. Each bin corresponds to the number of households within SEK 5,000. The estimated counterfactual density, displayed by the solid line, was obtained by fitting a seven-degree polynomial to the density, excluding points within SEK 40,000 below the kink. *b* denotes the excess mass and *s.e.* is the estimated standard error.

(A) THIRD-PARTY-REPORTED NET WEALTH   (B) TAXABLE NET WEALTH
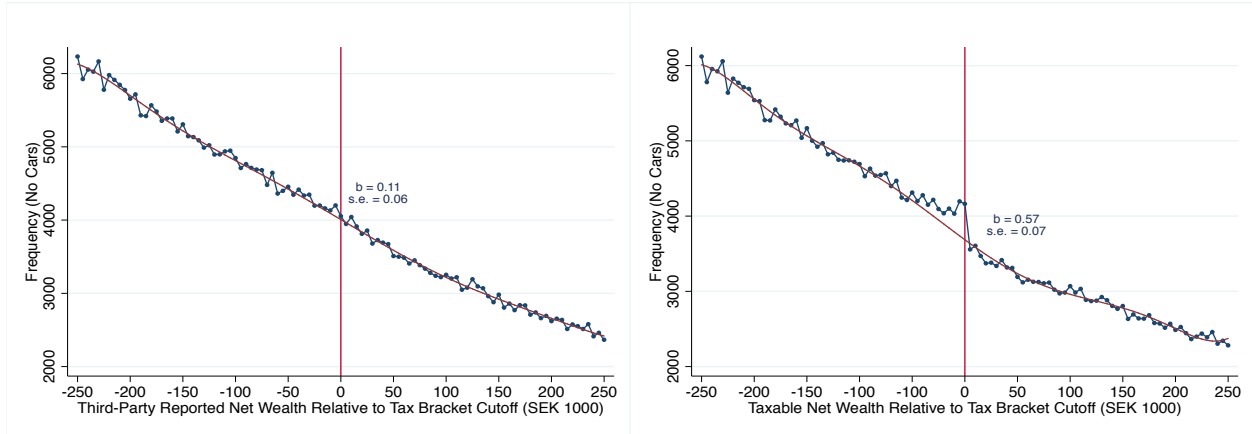
Notes: Panel A shows the distribution of third-party reported net wealth around the shift in the tax brackets, demarcated by the vertical at 0, for the years 2000-2006 for households who own no cars. Panel B shows the distribution of taxable wealth around the threshold for the same sample. The dotted series consist of a histogram relative to the normalized kink point. Each bin corresponds to the number of households within SEK 5,000. The estimated counterfactual density, displayed by the solid line, was obtained by fitting a seven-degree polynomial to the density, excluding points within SEK 40,000 below the kink. *b* denotes the estimated excess mass and *s.e.* is the estimated standard error.

FIGURE C.6: TAXABLE WEALTH DISTRIBUTION AROUND THE THRESHOLD FOR PREVIOUS TAX PAYERS.

(A)

Notes: The figure shows the distribution of taxable net wealth around the shift in the tax brackets, demarcated by the vertical at 0, for previous wealth tax payers. The graph consists of couples and singles in 2002 and 2003 who paid wealth taxes in 2001 and couples in 2005 and 2006 who paid wealth taxes in 2004. The dotted series consist of a histogram relative to the normalized kink point. Each bin corresponds to the number of households within SEK 5,000.

FIGURE C.7: IMPUTED TAXABLE WEALTH DISTRIBUTION.

(A)
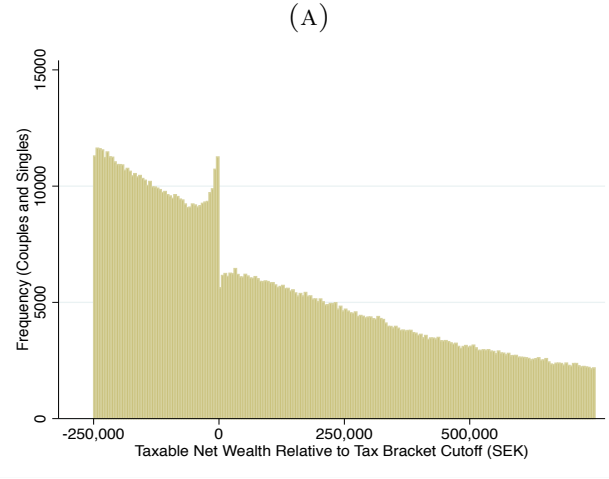
Notes: The figure shows taxable wealth around the threshold, when non-third-party-reported taxable wealth has been imputed and added to households who are not obliged to self-report the value of those assets and liabilities (i.e. for those who do not have third-party-reported wealth above the threshold or who do not make self-reported adjustments). To be restrictive, only 10 % of the imputed car value has been added to the non-third-party-reported wealth. The Swedish Bankers' Association report that collaterlized loans for cars, boats and consumption durables amount to 2.5% of other forms of collateralized debt. 2.5 of the third-party-reported debt are thus subtracted from the imputed non-third-party-reported assets.

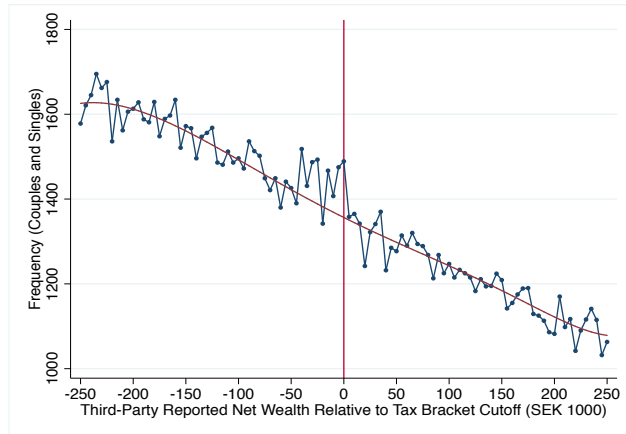FIGURE C.8: ESTIMATED BUNCHING OF THIRD-PARTY-REPORTED NET WEALTH AT THE THRESHOLD.
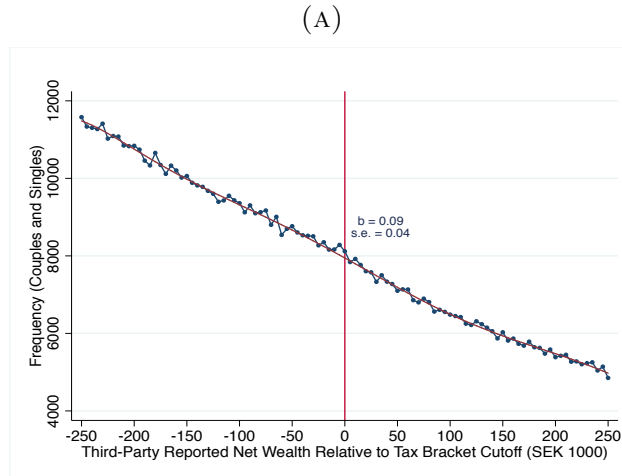
(A)

Notes: The figure shows the distribution of third-party-reported net wealth around the shift in the tax brackets, demarcated by the vertical line at 0, for the years 2000-2006. The dotted series consist of a histogram relative to the normalized kink point. Each bin corresponds to the number of households within SEK 5, 000. The estimated counterfactual density, displayed by the solid line in red, was obtained by fitting a seven-degree polynomial to the actual density, excluding points within SEK 40, 000 below the kink. *b* denotes the estimated excess mass and *s.e.* is the estimated standard error.

D Asymmetric Bunching

Figures 1 of the main text reveals that bunching is asymmetrically distributed around the threshold: there is excess mass also to the left of the kink point. I explore three potential explanations for this finding below.

First, asymmetric bunching is consistent with a *fixed cost* incurred at the kink point, implying a discontinuity in average, rather than marginal, tax rates. As explained in Section I, tax payers were required to report non-third party reported assets and liabilities if they were located above the threshold in terms of third-party-reported wealth or, alternatively, in case total taxable wealth placed them in the positive tax bracket. If there were a fixed cost of self-reporting, one would expect everyone located to the right of the kink in third-party-reported wealth to incur this cost. To avoid this fixed cost, households would then locate just below the threshold in terms of third-party-reported wealth. However, Figure C.8 displays the distribution of third-party-reported net wealth around the threshold and the estimated excess mass is only 0.09.

Moreover, if one believes the asymmetry to be the result of a fixed cost, symmetric bunching would obtain at kinks free of such fixed costs. Gelber et al. (2015) study bunching at the kink in the marginal tax rate induced by the U.S. Social Security Annual Earnings Test and find evidence of asymmetric bunching, although no fixed costs are involved in crossing the threshold. In Sweden, capital income is taxed annually at 30% and capital losses yield a tax credit of 30% of the aggregate losses up to losses of SEK 100,000. Losses above that threshold result in a tax credit of SEK 30,000 plus 21% of losses in excess of the threshold. If the realization of gains and losses were easily monitored intertemporally, one would expect a sharp kink at this salient threshold. However, looking at Figure D.1, there is asymmetric bunching in the same manner as for the wealth tax, although no fixed costs are incurred to the right of the threshold.

Second, asymmetric bunching at the threshold is consistent with the *confusion of marginal and average tax rates*, so that individuals believe that their total tax liability increases discontinuously at the threshold. If this were the case, one should expect similar responses to other marginal tax kinks. However, investigating bunching at thresholds in the Swedish taxable earnings schedule, Bastani and Selin (2014) find symmetric bunching around a kink where the marginal tax rate increases by 20%. This bracket shift affects individuals at the high end of the income distribution who, presumably, also pay wealth taxes. Another way to address this concern would be to investigate bunching among households who have paid wealth taxes previously and therefore should have learned about the workings of the tax schedule. Figure C.6 shows bunching at the threshold for those who paid wealth taxes in the past. Although less precise, the graph displays asymmetric bunching in the same way as in the graphs above. It therefore seems unlikely that individuals confounding MTRs and ATRs lies behind the observed asymmetry.

Third, asymmetric bunching could occur *mechanically* if only households above the threshold self-report. However, the fraction of households who self-report assets or liabilities is essentially continuous around the threshold defined for third-party-reported wealth.[7]

Another way to investigate whether bunching occurs mechanically is to use car values and

---

[7]The smoothness of this fraction depends on the sample studied. Results are available upon request.

estimate non-third-party-reported wealth for households who are not obliged to self-report. To take a conservative approach, I let 10% of the registered car value constitute non-third-party-reported assets. To assess liabilities, I use reports from the Swedish Bankers' Association that the total value of collateralized debt for cars, boats and consumption durables amount to about 2.5% of other collateralized debt in 2014. Non-observable liabilities are thus estimated to be 2.5% of each household's third-party-reported collateralized debt. When using the resulting values for non-third-party-reported wealth, there is even clearer bunching, as displayed in Figure C.7.[8]

FIGURE D.1: CAPITAL INCOME LOSSES AROUND THE THRESHOLD.

(A)
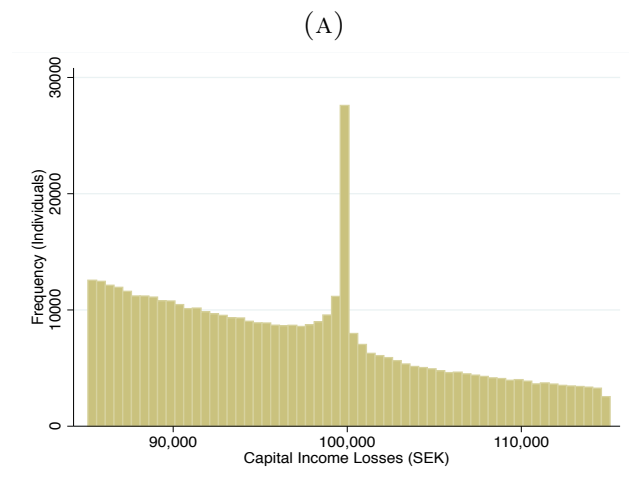
Notes: The figure shows the distribution of taxable capital income losses for 1995-2009 around the threshold of SEK $100,000$ where the induced marginal tax credit decreases from 30% to 21%. Each bin corresponds to the number of households within SEK 500.

---

[8]The restrictively computed W_{it}^{SR} is added to W_{it}^3 if there is no self-reported wealth and $W_{it}^{SR} + W_{it}^3 < W_t^*$.

E Underreporting of Cars, Robustness Analaysis

FIGURE E.1: SELF-REPORTING OF CARS.

(A) REPORTING CARS
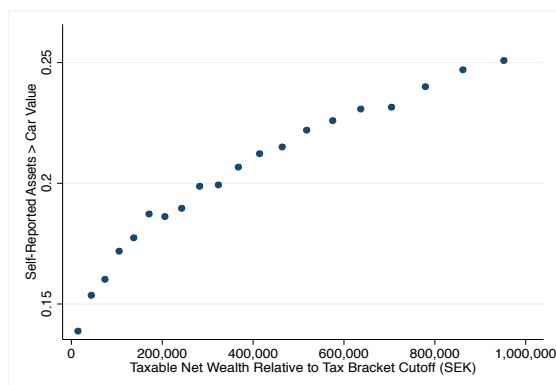
(B) REPORTING NEWLY PURCHASED CARS
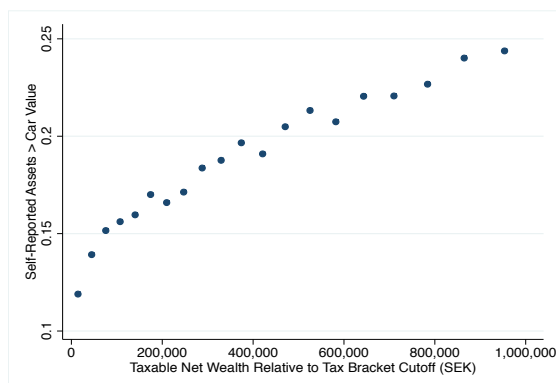
Notes: Panel B shows the fraction of car-owning households who self-report more assets than their cars are worth against taxable wealth close to the tax cutoff for the years 2000-2006. Panel C replicates Panel B but restricts the sample to households who purchased cars during the year. Each bin corresponds to taxable wealth of SEK 25,000. Each dot corresponds to mean of the fraction reporting more within a taxable wealth bin of 25,000.

F   Difference-in-difference Robustness Analaysis

FIGURE F.1: EFFECTS OF THE WEALTH TAX ON VARIOUS OUTCOMES

(A) SAVINGS (BEGINNING-OF-YEAR REBAL-
ANCING)
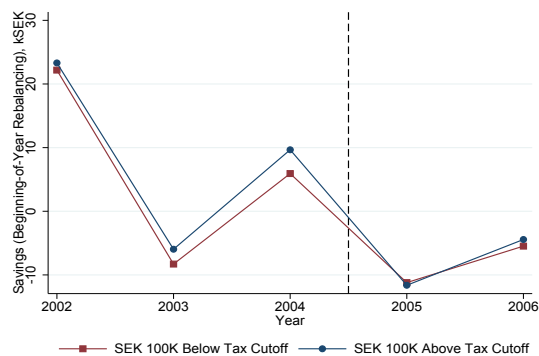
(B) VALUE OF CARS
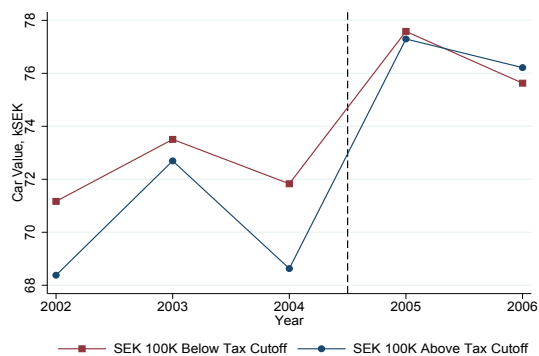
Notes: This graphs show outcome variables for couples divided into two groups year by year: those with wealth within SEK $100,000$ above the threshold that was in place during 2002-2004 (treatment group) and those with wealth within SEK $100,000$ below (control group). Panel A shows active changes in financial wealth (stocks, funds and bonds) assuming beginning-of-year rebalancing while Panel B displays the average value of cars around the kink over time.

FIGURE F.2: EFFECTS OF THE WEALTH TAX ON VARIOUS OUTCOMES
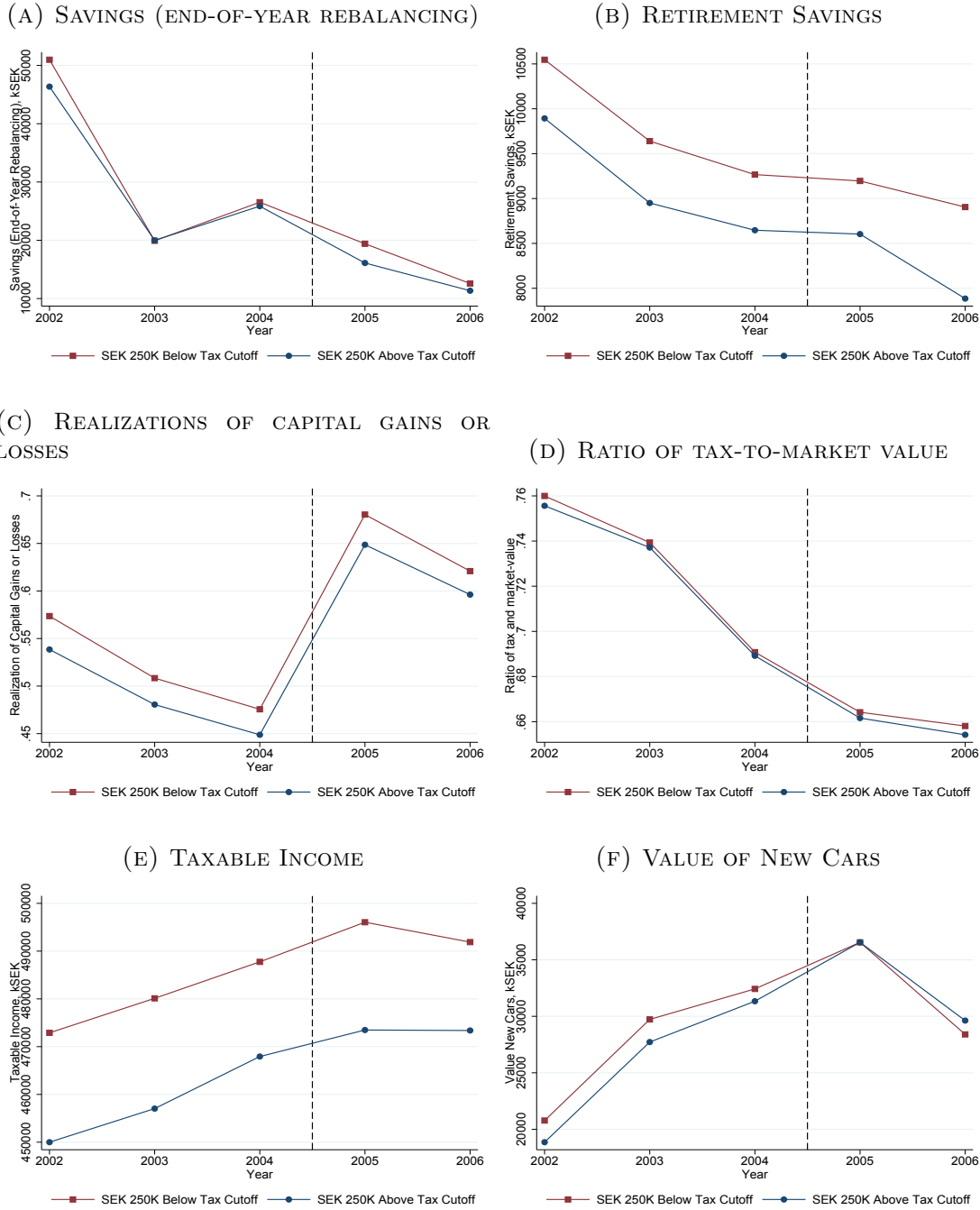
(A) SAVINGS (END-OF-YEAR REBALANCING)

(B) RETIREMENT SAVINGS

(C) REALIZATIONS OF CAPITAL GAINS OR LOSSES

(D) RATIO OF TAX-TO-MARKET VALUE

(E) TAXABLE INCOME

(F) VALUE OF NEW CARS

Notes: This graphs show outcome variables for couples divided into two groups year by year: those with wealth within SEK 250, 000 above the threshold that was in place during 2002-2004 (treatment group) and those with wealth within SEK 250, 000 below (control group). Panel A shows active changes in financial wealth (stocks, funds and bonds), while Panel B and C display retirement savings and the fraction of households who realize capital gains or losses over time. Panel D shows the average ratio of tax value of assets to market values and Panels E and F consider taxable income and the value of new cars, respectively.

G   Dynamic Framework

The purpose of this section is to illustrate how a wealth tax can influence savings and wealth dynamically. I present a canonical overlapping-generations (OLG) framework of intertempo-

ral utility maximization, consistent with the Life Cycle Hypothesis (LCH) (Bernheim, 2002; Modigliani and Brumberg, 1954), in which individuals accumulate wealth both to finance consumption in retirement and as insurance against idiosyncratic shocks in labor earnings.

I consider a dynamic economy with a discrete set of generations $0, 1, \ldots, t$, of measure one. Individuals are identical and live for $T + 1$ years. There is no population growth, so that at each point in time, there are $T + 1$ individuals alive. The first J years of each agent's life are spent working, and the remaining $T + 1 - J$ are spent in retirement. Income while working is stochastic and, for simplicity, assumed to follow a first-order Markov chain. Let $y_t(h_t)$ denote income in period t where h_t is the realization of the state in that period. The transition probability, i.e. the probability that state k is realized next period when the current state is i, is denoted $p_{ik} = P\left(h_{t+1} = h^k | h_t = h^i\right)$ with $\sum_{l=1}^m p_{il} = 1 \ \forall \ i = 1, \ldots, m$. Income in each period of retirement, b, is deterministic and lower than the expected income while working.

Individuals choose savings, s_t, and consumption, c_t, in each period. They face borrowing constraint and a progressive wealth tax τ_W above the threshold W^* and maximize the expected lifetime utility:

$$\max_{\{c_t, s_t\}_{t=0}^T} \mathbb{E} \sum_{t=0}^T \beta^t \frac{c_t^{1 - \frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} \tag{16}$$

subject to

$$c_t + s_t \leq i_t(h_t)$$
$$s_t = W_{t+1} - \min\{W^*, W_t\}(1 + r) - (\max\{W^*, W_t\} - W^*)(1 + r)(1 - \tau_W) \tag{17}$$
$$W_{t+1} \geq 0$$
$$c_t \geq 0$$
$$W_0 \text{ given}$$
$$i_t = \begin{cases} y_t(h_t) & \text{if } t < 40 \\ b & \text{if } t \geq 40. \end{cases}$$

where $\mathbb{E}$ is the expectations-operator, β is the discount factor, σ is the elasticity of intertemporal substitution and r the rate of return on capital. Before solving the problem, I use that the wealth tax is isomorphic to a tax on capital income, τ_c, i.e.

$$(1 - \tau_W)(1 + r) a_t = (1 + (1 - \tau_c) r) a_t \tag{18}$$

$$\tau_c = \frac{1 + r}{r} \tau_W \tag{19}$$

Equation 17 can then be written as:

$$s_t = a_{t+1} - \min\{W^*, a_t\}(1 + r) - (\max\{W^*, a_t\} - W^*)\left(1 + r\left(1 - \frac{1 + r}{r} \tau_W\right)\right). \tag{20}$$

In solving this optimization problem, I rewrite it recursively and apply dynamic programming.[9] Current assets and the realization of the shock are state variables. The optimal

---

[9]This problem could, in principle, be solved by solving a system of equations, one for each history of shocks. However, as the number of periods increase, the number of shock histories quickly increases, rendering such an approach infeasible.

consumption and savings policy functions prescribe no remaining wealth at the time of death, $a_{T+1} = 0$. I solve for optimal consumption and savings as a function of the states numerically, by discretizing the state space and interpolating linearly between grid points.

G.0.1 Quantitative Analysis

The quantitative analysis is admittedly stylized and the objective is simply to show how responses to a wealth tax may materialize, both in steady state and transitorily. I assume that individuals live for 85 years, join the labor force at age 25 and earn labor income until they are 65. Thereafter, they live for 20 years in retirement. Income assumes one of two possible values: $y \in \{y(1)), y(2))\}$ where $y(1) = 1$ and $y(2) = 2$ and $p_{11} = p_{22} = 0.8$ and $p_{12} = p_{21} = 0.2$ are the transition probabilities. Income in retirement is $b = 0.2$. Individuals start without any inherited wealth, so that $W_0 = 0$. The interest rate, r, is set to 8% and $\beta = 0.9$. All individuals start out in the high-income state.

I solve the model separately for two regimes: one where $\tau_W = 0$ and one where $\tau_W = 0.015$ above a wealth threshold $W^* = 1.5$. I then simulate the model for 1000 individuals. Wealth over the lifecycle for one generation in each of the two regimes is presented in Panels A and B of Figure G.1. The results show that households start out without wealth and then, on average, maintain fairly constant wealth until reaching 50-55 years of age, when they start to accumulate wealth quickly. Notice that the distribution of wealth at any age before the asset-accumulation stage is capped close to the threshold in the regime with a tax on wealth, whereas the no-tax regime generates higher wealth levels in almost every stage of life.

When there are 61 equally-sized generations alive at each point in time, Panels C and D of Figure G.1 show the distributions of taxable wealth around the 1.5-threshold in the two regimes. Individuals bunch at the threshold in the progressive tax case, as implied by the life-cycle patterns.

Next, I investigate the dynamics. Suppose that the economy starts out without a wealth tax. When an unexpected and permanent wealth tax is introduced, individuals of different cohorts will react heterogenously depending on where in the life cycle they are. Figure G.2 shows the distribution of taxable wealth around the threshold of 1.5 at different points in time after the tax has been introduced. The first subplot presents the distribution one year after the reform. In this plot, the youngest generation will make all wealth accumulation decisions under the new regime. In subplot (D), the regime has been in place for five years. Here, the youngest five generations will live their entire lives under the wealth tax regime. Notice how the wealth distribution adjusts gradually to the new steady state shown in Panel D of Figure G.1. The speed of adjustment is partly driven by parameter values, but will always be gradual rather than instantaneous.

FIGURE G.1: WEALTH OVER THE LIFE CYCLE.

(A) WITHOUT WEALTH TAX (B) WITH WEALTH TAX

(C) WITHOUT WEALTH TAX (D) WITH WEALTH TAX
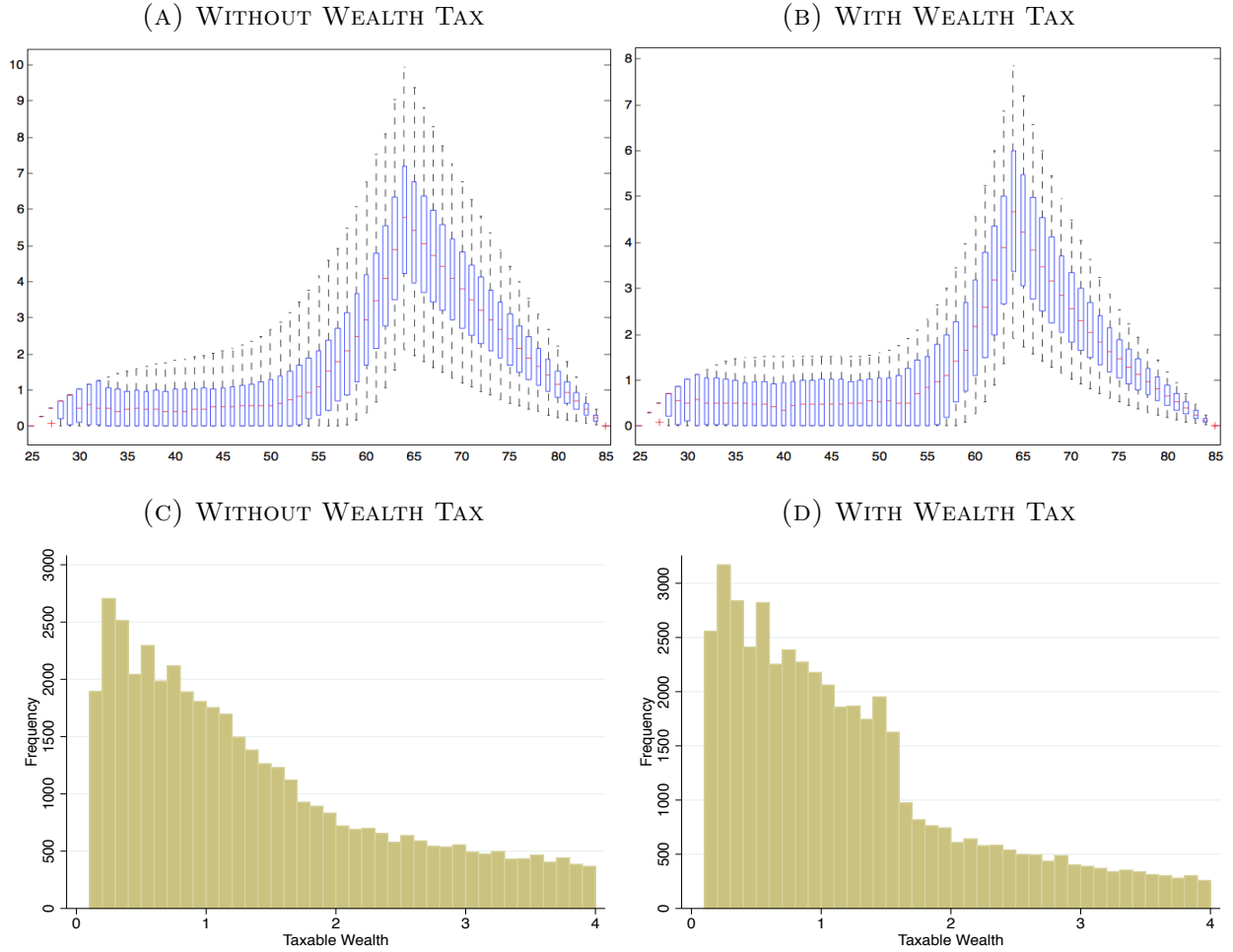
Notes: These figures show simulated wealth over the lifecycle for $1,000$ individuals according to the dynamic model laid out in Equation (16) in the case without a wealth tax (Panel A) and with (Panel B). In these plots, the y-axis represents wealth and the x-axis age. The solid boxes show the wealth interval, for each age, between the 25th and the 75th percentiles. The central mark in each box is the median. Observations not contained in the interval $[q_1 - 1.5\,(q_3 - q_1), q_3 + 1.5\,(q_3 - q_1)]$, where q_1 and q_3 are the 25th and 75th percentiles, are considered outliers and excluded. The dashed interval represents the rest of the wealth support for each age group. In the right plot, a wealth tax of 1.5% is paid on wealth above 1.5. The corresponding distributions of wealth are displayed in Panel C, without a wealth tax and D, with. In constructing these simulated figures, I assume that there are thousand individuals alive in each cohort rendering a total population of the economy of $61,000$.

19

FIGURE G.2: WEALTH DISTRIBUTIONS UPON A REFORM SHIFT.

(A) ONE YEAR AFTER (B) TWO YEARS AFTER (C) THREE YEARS AFTER

(D) FIVE YEAR AFTER (E) TEN YEARS AFTER (F) THIRTY YEARS AFTER
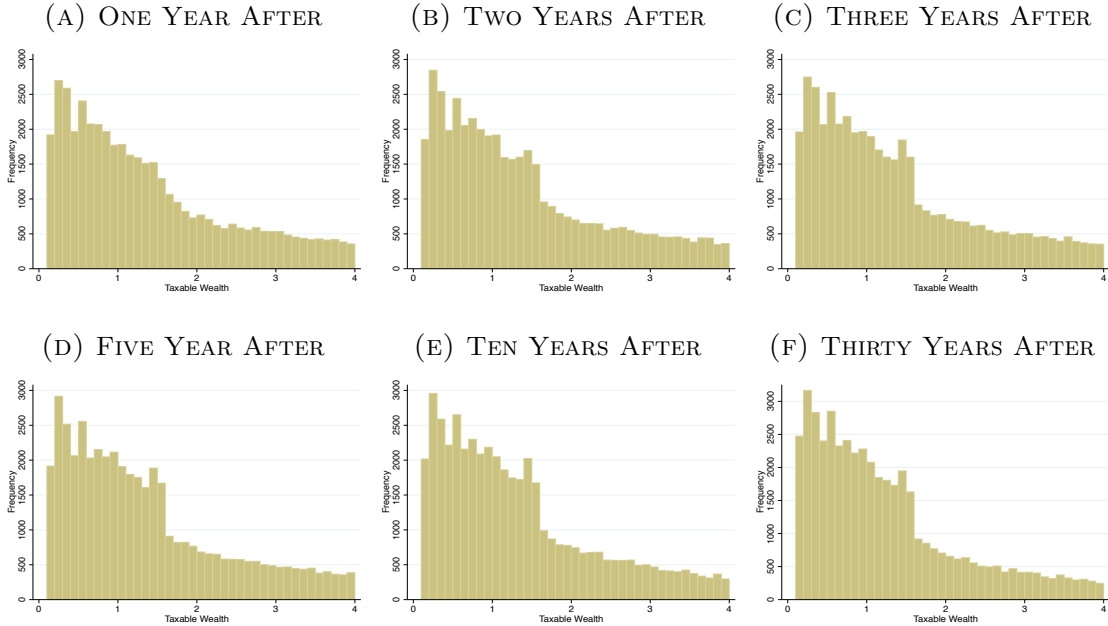
Notes: These graphs show the simulated distributions of wealth after implementing a permanent and un-expected wealth tax. The number of individuals alive in the economy is $61,000$ and is the same in each graph.

H   Estimation of the Inequality Deflator

The purpose of this section is to provide a normative assessment of the role of wealth taxation, based on the empirical results. Instead of approaching this problem from the standpoint of the optimal-taxation literature (e.g. Atkinson and Stiglitz, 1976; Chamley, 1986; Judd, 1985), I ask whether annual wealth taxes are likely to be more or less efficient in redistributing resources than progressive income taxes.[10] In other words, if the aim is to achieve a given level of redistribution, are wealth or income taxes more efficient in achieving that goal?

In the tradition of Kaldor (1939) and Hicks (1940), a common, normative approach to redistributive policies that produce an unequal distribution of gains and losses, is a compensation principle: if the aggregate surplus of a proposed policy is higher than the status quo, then winners can compensate losers and the policy should be implemented.[11] Subsequent papers have extended the Kaldor-Hicks principle by allowing for compensating transfers through the income tax schedule while recognizing the distortionary impacts of such augmentations (Hylland and Zeckhauser, 1979; Kaplow, 1996, 2004). To provide a normative account of wealth taxation, I employ a generalized framework capturing these ideas, following Hendren (2014).

---

[10]The optimal-taxation literature is applicable here because wealth taxes are isomorphic to capital income taxes. As shown in Appendix G, capital taxes are proportional to wealth taxes when the interest rate is positive, i.e. $\tau_c = \frac{1+r}{r}\tau_W$ for $r > 0$, where r is the net return and τ_c denotes the capital income tax.

[11]This approach puts equal value of money in the hands of the rich as the poor (Boadway, 1974; Fleurbaey, 2009). An alternative approach is to specify a social welfare function, with the caveat that the desired redistribution is sensitive to functional-form assumptions.

I consider the following policy experiment. Fixing the extent of redistribution that the wealth tax accomplished during its last year in place, 2006, I compute the hypothetical welfare effects implied by augmenting the income tax schedule so that it accomplishes that same level of redistribution. The approach acknowledges the potentially different distortions involved in each redistributive tax scheme.

The implementation of the method in Hendren (2014) requires estimating a so-called *inequality deflator*, defined by a function $g(y)$, that weighs individual surplus at each point of the income distribution, y, by the distortionary costs of raising tax revenue. 1 unit surplus in the hands of someone earning y, can be turned into $g(y)$ of government revenue. Typically, the weights are larger in the lower part of the y-distribution, not because a subjectively larger value is placed on poor people, but because redistributing money to the poor is more costly than redistributing to the rich in the presence of behavioral responses to tax rates. Using an estimate of g, I weigh and compute the social surplus of augmenting the income tax schedule to accomplish the same redistribution as under the wealth tax.

I estimate the inequality deflator using the universe of Swedish income tax returns for 2006. The inequality deflator is expressed as a function of the compensated elasticity of taxable income (ETI), the participation elasticity, the income elasticity, the income tax schedule and the local Pareto parameter of the income distributio5n.[12] I define it along the lines of Equations (3) and (6) in Hendren (2014):

$$FE(y) = -\varepsilon_c^P(y) \frac{T(y) - T(0)}{y - T(y)} - \varepsilon^c(y) \frac{\tau(y)}{1 - \tau(y)} \alpha(y),$$

where $\alpha(y) = -\left(1 + \frac{yf'(y)}{f(y)}\right)$ is the elasticity of the income distribution. ε_c^P is the participation elasticity. I follow Hendren (2014) and ascribe the value of 0.09 to this parameter. Similarly, the compensated taxable income elasticity, $\varepsilon^c(y)$. The deflator depends on the ETI because the magnitude of behavioral responses critically determines the effects on the government budget. Saez, Slemrod and Giertz (2012) propose an ETI in the interval $[0.12, 0.4]$, based on earlier studies. I use a conservative point estimate of 0.2 when computing the baseline deflator.[13] Sensitivity analysis shows that a lower elasticity generally reduces the costs of redistribution by income taxation because the effects on the government's budget become smaller.

The sample is defined as the population of Swedish individuals above 15 years of age who file positive taxable income. Income is defined as gross income before deductions and $T(y)$ is the amount of tax paid at income level y. $\tau(y)$ represents the marginal tax rate faced by an individual earning SEK y.

Appendix Figure H.1 presents the estimated inequality deflator along the income distribution. As expected, the lower part of the distribution is ascribed weights above one while the weights in the upper region are below one. Providing marginal resources to the poor is

---

[12]All elasticities are expressed in terms of the keep rate, i.e. net-of-tax. The inequality deflator in Hendren (2014), also takes into account that different individuals face different tax schedules depending on marital status and number of dependents. The Swedish income-tax schedule lacks these features.

[13]0.2 is also the midpoint of the ETIs estimated on Swedish data in Gelber (2014), who reports an estimate of 0.41 for men and 0.49 for women, and in Bastani and Selin (2014), who report a zero estimate for wage earners.

more costly than providing them to the rich, and the surplus accruing to the poor is therefore valued higher than that of the rich.

The social surplus generated by income taxation should be compared to the wealth-tax revenue collected, corrected for fiscal externalities from wealth taxation. A wealth-tax cut forgoes revenue directly, but in addition entails revenue effects because of behavioral responses. The relevant fiscal externalities involved in wealth taxation comprise effects on taxable wealth, effects on revenue from evasion penalties and spillover effects on taxable income.[14] For instance, using the tax elasticity estimate of 0.127 reported in Section III, a wealth tax cut of 50% decreases tax revenue from the wealth tax by 49.9%.

To see this, I apply the notation of the elasticity of taxable income in Saez et al. (2012) to show that the mechanical revenue effect, i.e. the effect in absence of behavioral responses, of a small wealth tax change, $\mathrm{d}\tau$ is given by $\mathrm{d}M = N\left(W^m - W^*\right)\mathrm{d}\tau$, where N is the number of tax payers located in the positive tax bracket, W^m is mean taxable wealth in that sample and W^* is the tax threshold.

The corresponding behavioral effect is, in turn, $\mathrm{d}B = N\mathrm{d}W^m\mathrm{d}\tau$, which is equivalent to $\mathrm{d}B = -N\varepsilon_{W,\tau}W^m\frac{\tau}{1-\tau}\mathrm{d}\tau$.

Exploiting that tax revenue, TR, is given by $TR = N\left(W^m - W^*\right)\tau$, the percentage change in government revenue from a small tax change, $\mathrm{d}\tau$, that yields government revenue $\tilde{TR}$ is then:

$$\frac{\tilde{TR} - TR}{TR} = \left(1 - \varepsilon_{W,\tau}\alpha\frac{\tau}{1-\tau}\right)\frac{\mathrm{d}\tau}{\tau}$$

$$\varepsilon_{G,\tau} = \left(1 - \varepsilon_{W,\tau}\alpha\frac{\tau}{1-\tau}\right),$$

where $\alpha = \frac{W^m - W^*}{W^*}$ and $\varepsilon_{G,\tau}$ denotes the elasticity of government revenue with respect to the wealth tax.

For a quantitative estimate of the elasticity, $W^m \approx 4,251m$ and weighting the thresholds for singles and couples by the number of tax payers of either categories yields a $W^* = 2m$ and an $\alpha = 1.1253$. With an elasticity of 0.127 and a tax rate of 0.015, the elasticity is 0.998, i.e. a one-percent increase in the wealth tax delivers a 0.998-percent increase in government revenue.

If the wealth tax were to trigger dynamic savings responses, the effects of the wealth tax on future government revenue would have to be internalized in the calculations. However, as discussed in Section IV, responses to the wealth tax appear to reflect reporting rather than real responses, leaving savings and future wealth unaffected by a tax cut. The trajectory of future tax revenue will thus be immune to tax adjustments. If the costs of sheltering money reflect transfer costs, a wealth tax cut would moreover influence revenue from penalties for tax evasion. I assume, however, that costs of sheltering are resource costs. This is a conservative assumption: if costs were transfers across agents, these would diminish the costs of wealth taxation, rendering income taxes relatively less attractive. Lastly, the results reported in Section IV suggest that there are no effects of the wealth tax on taxable income, indicating that interaction effects between the taxes can be ignored.

---

[14]For simplicity, I assume away effects of the wealth tax on consumption taxes, even though an absence of effects on savings and labor supply implies that the tax is paid by foregone consumption.

In 2006, the wealth tax generated revenue amounting to SEK 5,739 million and repealing the tax thus led to a loss in government revenue corresponding to that amount. As discussed above, the repeal is not likely to have influenced government revenue through behavioral effects on other tax bases. Augmenting the income tax schedule to generate an equivalent amount of redistribution produces a social, inequality-deflated surplus of SEK 4,362 million when the compensated ETI is 0.2. The difference in surplus thus amounts to SEK 1,377 million, or 24% of the total surplus generated by the wealth tax.

The surplus generated by redistribution under income taxation is lower than that under wealth taxation for two reasons. First, the results in this study suggest that behavioral responses to wealth taxes tend to be relatively smaller than those of income taxes. Second, wealth-tax payers tend to be located in the upper end of the income distribution, where fiscal externalities are greater and the deflator weights the lowest: the lower the taxable income elasticity, the higher is the inequality-deflated surplus.

Moreover, even when the low ETIs reported in Chetty et al. (2011) and Bastani and Selin (2014) are used, wealth taxation is preferable to income taxation under the baseline wealth-tax elasticity. Assuming an ETI of 0.05, the surplus from income taxation is SEK 5,110 million, which is still 11% lower than the surplus from wealth taxation. An even lower value of 0.01 produces a surplus of SEK 5,309 million, still suggesting that wealth taxation is superior to income taxation as a redistributive tool.

A few cautionary remarks are in order. The estimated wealth-tax elasticity is local along two dimensions. First, the elasticity is estimated at a threshold where the tax rises from 0 to 1.5 percent. These estimates may not be representative of larger marginal-tax-rate discontinuities, especially if some of the potential responses involve fixed costs of, for instance, rebalancing the portfolio to minimize tax liabilities. Second, the elasticity is estimated at particular points in the wealth distribution and may not be representative of responses to thresholds at different locations, even though elasticity estimates are very similar at the various positions of the wealth tax threshold.

Bearing in mind that the results may not be generalizable to large tax reforms, the results in this section suggest that, at least locally, a wealth tax is likely to generate a higher social surplus and therefore outperform income taxes as a redistributive tool. This key finding is robust to various measures of the elasticity of taxable income.

FIGURE H.1: INEQUALITY DEFLATOR.

(A)
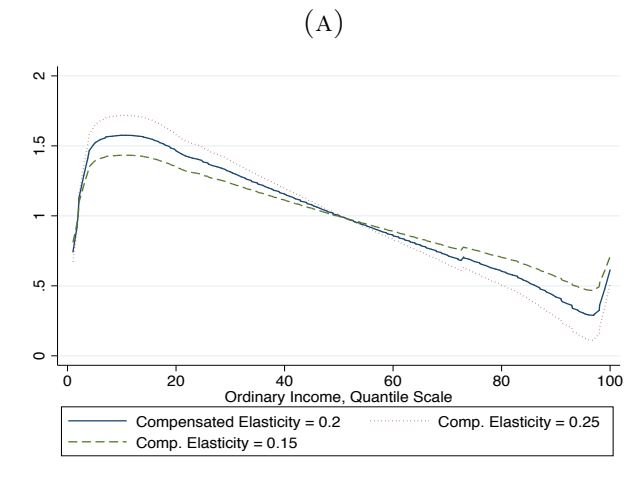
Notes: This graph displays the inequality deflator, defined in Appendix Section H and estimated on Swedish data for different compensated taxable income elasticities.

References

[1] Allingham, M. G. and A. Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics*, 1:323-338.

[2] Atkinson, A. and J. E. Stiglitz. 1976. "The Design of Tax Structure: Direct Versus Indirect Taxation." *Journal of Public Economics*, 6(1-2):55-75.

[3] Bastani, S. and H. Selin. 2014. "Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule." *Journal of Public Economics*, 109:36-49, Elsevier.

[4] Boadway, R. 1974. "The Welfare Foundations of Cost-Benefit Analysis." *The Economic Journal*, 84(336):926-939.

[5] Chamley, C. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica*, 54(3):607-622.

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

[7] Chetty, R., J. N. Friedman, T. Olsen and L. Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics*, 126(2):749-804.

[8] Feldstein, M. 1978. "The Welfare Cost of Capital Income Taxation." *Journal of Political Economy*, 86(2):29-51.

[9] Fleurbaey, M. 2009. "Beyond GDP: The Quest for a Measure of Social Welfare." *Journal of Economic Literature*, 47(4):1029-1074.

[10] Gelber, A. M., D. Jones and D. W. Sacks. 2015. "Earnings Adjustment Frictions: Evidence from the Social Security Earnings Test." GSPP Working Paper.

[11] Hendren, N. 2014. "The Inequality Deflator: Interpersonal Comparisons without a Social Welfare Function." Mimeo, Harvard University.

[12] Hicks, J. R. 1940. "The Valuation of Social Income." *Economica*, 7(26):105-124.

[13] Hylland, A. and R. Zeckhauser. 1979. "The efficient allocation of individuals to positions." *The Journal of Political Economy*, 87(2):293-314.

[14] Judd, K. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, 28(1):59-83.

[15] Kaldor, N. 1939. "Welfare Propositions of Economics and Interpersonal Comparisons of Utility." *The Economic Journal*, 49(195):549-552.

[16] Kaplow, L. 1996. "The optimal supply of public goods and the distortionary cost of taxation." *National Tax Journal*, 49(4):513-534.

[17] Kaplow, L. 2004. "On the (ir)relevance of the distribution and labor supply distortion to government policy."

[18] Mayshar, J. 1991. "Taxation With Costly Administration." *Scandinavian Journal of Economics*, 93(1):75-88.

[19] Modigliani, F. and R. Brumberg. 1954. "Utility Analysis and the Consumption Function: an Interpretation of Cross-Section Data." In K. K. Kurihara, ed., *Post Keynesian Economics*, 15:388-436. New Brunswick, NJ: Rutgers University Press.

[20] Saez, E., J. Slemrod and S. H. Giertz. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature*, 50(1):3-50.

[21] Slemrod, J. 2001 "A General Model of the Behavioral Response to Taxation" *International Tax and Finance*, 8:119-128.

[22] Yitzhaki, S. 1974. "A Note on Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics*, 3(2):475-480.