Real or Evasion Responses to the Wealth Tax? Theory and Evidence from Sweden[*]

David Seim[†]

JOB MARKET PAPER

November 23, 2012

Abstract

This paper addresses the behavioral effects of an annual wealth tax. I use Swedish tax records over the period 1999-2006 to estimate bunching at kink points in the progressive tax schedule and find significant estimates of the implied elasticity of taxable net wealth at about 0.3. Exploiting features of the institutional setting, I decompose the effects into a reporting response and a saving response. The results suggest that an increase in the tax is likely to stimulate evasion rather than deter savings. I merge tax records with military enlistment records on cognitive ability and find that high-skilled individuals respond more to the tax. This suggests that the incidence of the tax falls disproportionally on the cognitively less able.

[†] Institute for International Economic Studies, Stockholm University, 106 91 Stockholm, Sweden. Email:david.seim@iies.su.se; phone: +46-8-163075; fax:+46-8-161443.

1 Introduction

Wealth taxes are surprisingly widespread given the lack of theoretical consensus on how wealth should be taxed, and the sparse knowledge about their effects. This paper addresses the latter problem. The effects that I document suggest a couple of reasons for why wealth taxes might be ineffective, inefficient, and unjust: the Swedish annual wealth tax was subject to evasion, in particular by households with higher cognitive ability.

Historically, many western economies have taxed net wealth and a number of countries, including France, Norway and Spain, still uphold these policies. Yet, wealth taxes remain controversial. Opponents of wealth taxation argue that the introduction of distortive taxes may hamper capital accumulation in the long run. In addition, a range of well-known problems are associated with the implementation of wealth taxes, including the difficulties to define a comprehensive tax base, appraise assets and prevent tax evasion (Adam et al., 2011; Brown, 1991). Proponents of wealth taxation, on the other hand, argue that it is a direct and effective way of reducing inequality. The question of whether or not wealth taxes are desirable is thus subject to a continuous debate that is far from settled.

In recent years, there has been renewed academic interest in the normative aspects of capital and wealth taxation. The proposition that capital taxes should be set to zero, as argued by Atkinson and Stiglitz (1976), Chamley (1986) and Judd (1985), has been scrutinized and challenged.[1] Within this growing strand of literature, Diamond and Saez (2011) question assumptions made in the optimal taxation literature and propose a rationale for positive capital income taxes. Piketty and Saez (2012) derive expressions for optimal policy as a function of empirically estimable parameters.[2]

The empirical contribution of this paper is to analyze how wealth taxation works in practice. To quantify welfare costs of taxes, Feldstein (1995, 1999) shows that the tax elasticity of the tax base is a so-called sufficient statistic for welfare analysis since it incorporates all responses. When approaching wealth taxation from an empirical perspective, our first objective should thus be to estimate the effect of the wealth tax on taxable net wealth, a sufficient statistic for the ensuing utility flow. However, even if such an elasticity can be credibly estimated, it only provides part of the picture. To understand the workings of a wealth tax, margins of adjustments might still matter, not just the composite, reduced form response. Taxation typically affects not only the incentives to save, but may also trigger undesirable behavior in terms of tax sheltering. Chetty (2009) challenges the results in Feldstein (1995, 1999) and shows that a distinction between real responses and evasion responses becomes crucial for welfare analysis if the marginal tax dollar lost to evasion is tantamount to a transfer across agents.[3] As noted

---

[1]Atkinson and Stiglitz (1976) argue that, in the presence of skill heterogeneity, the possibility of using non-linear income taxes implies that there is no role for capital taxes. Chamley (1986) and Judd (1985) show that optimal capital taxes should be zero in the long run. Here the result depends on the infinite-horizon setting and a tax distortion which grows exponentially over time and cannot be optimal from the social planner's point of view.

[2]Boadway, Chamberlain and Emmerson (2010), Banks and Diamond (2010) and Kopczuk (2013) provide extensive surveys of the literature on optimal wealth and capital income taxation.

[3]This decomposition hinges on evasion costs being transfers across agents. The effect on the tax base

by Kopczuk (2013), this is an area where evidence is particularly scant. The issue of tax sheltering may be particularly relevant in the context of wealth taxation. Many taxes in advanced economies are based on information solicited from third parties, which makes them largely impervious to abuse (Kreiner, Kleven, and Saez 2009). The wealth tax base, on the other hand, also involves an element of self-reporting, which makes the wealth tax susceptible to lower compliance rates and tax evasion.

Assessing the effects on the tax base and decomposing these into real and reporting responses may help us understand important aspects of policy. But even with perfect knowledge of these effects, the story would remain incomplete. If the ability to comprehend the tax system differs across individuals, they may be asymmetrically affected by tax reforms. Traditional models of optimal taxation generally assume that all agents are rational and respond optimally to policy. In practice, opaque reforms and salient policies may trigger widely different responses.[4] For instance, the tendency to confound marginal and average tax rates indicate that individuals may not necessarily respond as predicted to a progressive wealth tax.[5] If comprehension of the tax scheme is unevenly distributed, the tax burden may fall disproportionately on the less able.

Motivated by this general background, I investigate the behavioral effects of the Swedish annual wealth tax, using a unique dataset, comprising about 58 million observations in a panel of individual taxpayers. I quantify the effects of the tax by estimating the tax elasticities of taxable net wealth and decompose the response into a reporting effect and an intertemporal substitution effect. Using a measure of cognitive skills, I also test the hypothesis that high-skilled individuals understand the tax system better and respond differently to tax reforms than do low-skilled individuals.

Sweden introduced a progressive tax on net wealth in 1947. From 1991, when Sweden implemented an extensive tax reform, until 2007, when the tax was repealed, it had two brackets, separated by a threshold.[6] The marginal tax rate was zero below the threshold, and 1.5 percent above it and the threshold was changed a number of times. The design of the tax schedule gives rise to two sources of variation that can be exploited empirically. First, the threshold for taxable wealth creates a kink in the budget set which, under general assumptions about individuals' behavior, makes them bunch at the kink point. Second, the change in the threshold over time enables a difference-in-difference strategy.

Swedish taxpayers annually received a prepopulated tax return based on the net wealth reported to the tax authority by third parties, such as banks and financial institutions. However, they were required to report any omitted assets and liabilities themselves. Third-party reported net wealth corrected for the taxpayers' self-reported adjustments thus constituted taxable net wealth. With access to both third-party reported net wealth and self-reported net wealth, I can study the nature of household

_____

remains a sufficient statistic for welfare analysis when evasion is present and costs are not transfers but, for example, moral costs.

[4]See Chetty, Looney and Kroft (2009) for policy responses and salience.

[5]On the confusion of marginal and average taxes, see Liebman and Zeckhauser (2004) and Feldman and Katuščàk (2006).

[6]On the reform in 1991, see Agell et al. (1996).

responses.

By linking the wealth tax records to military enlistment data, which include a measure of cognitive ability, I study how responses differ across skill groups. Since military enlistment was mandatory for males in the sample cohorts, the measure is a comprehensive proxy for actual ability and is not plagued by selection bias.

The institutional setting thus makes Sweden a promising testing ground for identifying the behavioral effects of the wealth tax. I pursue a top-down approach, first assessing the data imposing as few assumptions as possible about preferences and behavior, and then gradually adding more structure to unravel the mechanisms behind the observed outcomes.

To my knowledge, this paper is the first to assess the tax elasticity of taxable net wealth. Using the variation in the marginal tax rate across tax brackets, I start by estimating bunching at the kink point, i.e., the excess mass in the distribution at the tax threshold. When applying the methods proposed by Saez (2010) and Chetty, Friedman, Olsen and Pistaferri (2011) to Swedish wealth data for 1999-2006, I find strong evidence of bunching at the kink point. The implied tax semi-elasticity of taxable net wealth lies in the absolute range $[0.12, 0.33]$, depending on the chosen bunching estimate. In other words, an increase in the tax by one percentage point leads to a reduction in taxable net wealth by 0.12-0.33 percent. I then exploit movements in the tax bracket over time in a difference-in-difference strategy. Using this approach, I find larger tax semi-elasticities of taxable wealth, ranging from 0.45 to 0.80 depending on the chosen time frame. The larger responses could be attributed to the larger salience of the tax on those once already close to the kink.

By comparing the distribution of third-party reported net wealth to taxable net wealth, I find that all bunching occurs along the self-reported margin. Kleven et al. (2011) argue that the high tax-compliance rates in modern tax systems are a result of third-party reported tax liabilities that are difficult to evade. Evasion is thus more likely to occur along the self-reported margin. I examine whether tax evasion is the source of bunching in two different ways. First, I empirically relate maximal pecuniary penalties for cheating to the difference between third-party reported net wealth and taxable net wealth and find that downward adjustments of taxable net wealth due to self-reporting are positively and significantly correlated with penalties. Second, I study dynamic responses to a change in the threshold. If a change with a new threshold were to reflect real intertemporal responses, the stock of wealth would adjust only gradually and bunching at the new threshold would therefore manifest itself over time. However, I find that bunching occurs immediately upon moving the threshold and there is no systematic increase over time. These findings suggest that the self-reported margin is a good proxy for wealth tax evasion.

Since the data not only support the occurrence of bunching but also suggest that changes to the tax scheme affect evasion rather than savings decisions, I impose some structural assumptions in order to address the underlying mechanisms at work. I start with a simple static model, where agents choose how much money to shelter from the

government, subject to a pecuniary convex cost of evasion consistent with the work of Allingham and Sandmo (1972) and Slemrod (2001).[7] I use bunching at kink points and data on penalties to estimate two parameters of the evasion cost function. Translating these findings into tax semi-elasticities of taxable net wealth, I find that they are in the range of $[0.18, 0.42]$. I then extend the model to a dynamic setting where agents are motivated to save in order to smooth consumption. Parameterizing the model, I estimate uncompensated tax semi-elasticities of wealth in the range $[0.20, 0.40]$.

I finally investigate whether some individuals are more prone to evade than others. If cognitively able individuals are indeed more able to understand the tax system, do they respond more? To allow for heterogenous effects in cognitive skills, I use the framework in Chetty, Looney and Kroft (2007) to specify a bounded-rationality model where agents compare the utility from taking the tax rate into account, to the utility from ignoring it. If the utility difference exceeds some exogenous cost, agents do internalize the tax rate. Assuming that the cost depends negatively on cognitive ability, the model predicts more bunching among high-skilled households. This prediction is confirmed by the data. Exploring various explanations for this finding, I argue that agents with high ability understand the tax system better than agents with low ability. I find no evidence for alternative explanations, such as access to different technologies for evasion or differences in risk and time preferences.[8]

To ease the exposition, the different theoretical approaches are collected in Section 2. Section 3 presents the institutional setup and data. Section 4 provides the estimates of bunching at kink points along with an argument for how to empirically gauge evasion. Section 5 uses the theory in Section 2 and the estimates of bunching in Section 4 to compute the elasticities of the tax base to the wealth tax. Key results are discussed in Section 6 and Section 7 concludes.

2   Theory

This section outlines a theoretical framework for studying intertemporal substitution and reporting responses to a wealth tax. This is done in four steps. The framework in the first step refrains from making functional-form assumptions about individual behavior. The second step features a sheltering decision in a parametrized, but static framework. The third step explores a dynamic model, allowing also savings to respond to the tax. The fourth step presents a simple bounded-rationality framework which incorporates heterogeneity in skills.

---

[7]Slemrod and Yitzhaki (2002) provide an extensive review of the literature on tax evasion. Theoretical predictions of the effect of a tax increase on sheltering behavior are shown to be highly model-dependent. Yitzhaki (1974) shows that if the penalty is proportional to the tax rate, as is the case for modest amounts of cheating in many countries, the amount of sheltering is independent of the tax rate.

[8]Although the data used for this investigation has unique coverage, the sample is small and the results should therefore be interpreted with caution. The literature seems to have settled on a positive relation between cognitive ability and patience and a negative relation between cognitive ability and risk aversion. See Frederick (2005) and Benjamin et al. (2006).

2.1 Small Kink Analysis

Individuals trade off consumption today, c, against savings, s (consumption in the future). They are heterogenous with respect to preferences and savings technologies, which are distributed according to some continuous and differentiable cumulative distribution function. With a linear tax on wealth (stock of savings), denoted by τ, and no uncertainty, individuals' wealth, Z, will be distributed according to a smooth density function $h(Z)$. This is true also when allowing for a tax-sheltering opportunity subject to a convex sheltering cost.

Introducing a kink in the budget set at wealth level z^*, by a higher marginal tax rate $\tau + d\tau$ to the right of the kink, will trigger a savings response and, potentially, a reporting response for agents with a taxable net wealth level above z^*. Since wealth is a stock variable, the initial effect should represent reporting responses rather than intertemporal-substitution effects.[9] Agents that under the linear tax scheme chose taxable net wealth levels in some interval $[z^*, z^* + dz]$, will bunch at the kink point. The number of households that bunch is thus $B = h(z^*) dz^*$.[10] Individuals who chose higher wealth levels in the absence of the higher tax, reduce their taxable wealth to the point where their indifference curves are tangent to the budget line under the higher tax (which has the slope $1 - \tau - d\tau$).

Under the assumption that the tax rate change at the kink is sufficiently small, so that the associated income effects are negligible, the response of taxable wealth can be interpreted by way of a compensated semi-elasticity. The semi-elasticity refers to the percentage change in wealth arising from a one-percentage point increase in the tax rate, $\varepsilon_{W,\tau} = -\frac{dz}{d\tau}\frac{1}{z}$. Combining this with the expression for bunching at the kink, I obtain

$$\frac{B}{h(z^*)z^*} = -\varepsilon_{W,\tau}^K d\tau, \tag{1}$$

where W denotes taxable net wealth. As the response can differ over time, estimating bunching dynamically can potentially identify a vector of elasticities until the steady-state wealth distribution is obtained along with a long-run elasticity. With heterogeneous elasticities, bunching identifies the average elasticity.

2.2 Static Model

In practice, agents may respond to the wealth tax by saving less or by reporting less wealth. In this subsection I focus on reporting behavior.

Assume that agents aim at maximizing their budget given some pecuniary cost of tax sheltering. In principle, the framework can be reinterpreted as a maximization problem where utility is linear in consumption and agents pay a convex moral cost of tax evasion. However, in order to estimate the parameters, I regard the cost as monetary. Although this distinction may appear peripheral, it has important implications

---

[9]This is true for unanticipated and immediately implemented tax changes.

[10]The expression is approximate, as it assumes a constant density over the net wealth interval of bunchers. The accuracy of the approximation increases as dz^* decreases.

for welfare. If the cost is a sunk resource cost, Feldstein (1999) shows in a labor income tax setting that whether agents respond through behavior or through reporting is not important from the viewpoint of the social planner. According to Chetty (2009), this is no longer true when the costs take the form of transfers across agents (via tax sheltering penalties for example).

Formally, an agent faces the following maximization problem:

$$\max_{e \leq s} (1 - \tau)(s - e) + e - C(e, s), \tag{2}$$

where s denotes exogenous savings, 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.[11] 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:

$$C(e, s) = \left(\frac{e}{s}\right)^{\frac{1}{\gamma}} \frac{pe}{1 + \frac{1}{\gamma}}, \tag{3}$$

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.[12] 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 (2) is given by

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

and the agent's taxable net wealth by

$$s - e^* = \left(1 - \left(\frac{\tau}{p}\right)^{\gamma}\right) s. \tag{5}$$

The cost function implies that evaded amounts are proportional to true wealth. This is a prediction that I empirically assess in Section 5.2. 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

---

[11] Allingham and Sandmo (1974) 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.

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

counterfactual density together with the fact that $\log(1+x) \approx x$ for small x, I obtain:

$$\frac{B}{h(z^*)z^*} \approx \log\left(\frac{1-\left(\frac{\tau_0}{p}\right)^\gamma}{1-\left(\frac{\tau_1}{p}\right)^\gamma}\right), \tag{6}$$

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 kink point in excess of the counterfactual density. If $\tau_0 = 0$, I approximate equation (6) as

$$\frac{B}{h(z^*)z^*} \approx \left(\frac{\tau_1}{p}\right)^\gamma. \tag{7}$$

Equation (7) allows me to identify the tax elasticity of evasion, γ, as a function of the observable parameters τ_1 and z^*, the estimable cost parameter p and bunching at the kink, $B/h(z^*)$. This formula has the intuitive feature that the constant fraction of wealth that is sheltered from the government (the right-hand side) and estimated excess mass are proportional to each other. As the tax rate increases, the fraction of evaded savings goes up and more households bunch at the kink.

The relevant semi-elasticities take the following form:

$$\varepsilon^S_{e,\tau} = \frac{\gamma}{\tau}$$

$$\varepsilon_{W,\tau} = -\frac{\gamma}{\tau}\frac{\left(\frac{\tau}{p}\right)^\gamma}{1-\left(\frac{\tau}{p}\right)^\gamma},$$

where W again represents taxable net wealth. The tax elasticity of tax evasion is decreasing in the tax rate and for $p > \tau$, the tax elasticity of taxable net wealth is increasing in the tax rate.

2.3 Dynamic Model

Section 2.2 assumes that all responses to the tax rate occur through tax sheltering. This section considers a dynamic model where agents also respond to tax changes by way of savings. The model abstracts from labor supply responses to the wealth tax.

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 in Section 2.2.

Since a change in the wealth tax may trigger a slow dynamic response of wealth, rather than a one-off adjustment, Appendix A lays out the infinite-horizon version of this dynamic model. This extension allows for a transition path of wealth to the new steady state.

The agent faces the following maximization problem:

$$\max_{s,e \leq s} U(c_1, c_2) = \max_{s,e \leq s} \frac{c_1^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} + \beta \frac{c_2^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} \tag{8}$$

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 $G(y)$.[13] The cost of evasion is assumed to take the same functional form as in Section 2.2, implying that the first-order condition that governs the tax sheltering response is given by equation (4), 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:

$$c_1^{-\frac{1}{\sigma}} = \beta \left(1 - \tau \left(1 - \left(\frac{\tau}{p}\right)^\gamma \frac{1}{1 + \gamma}\right)\right)^{1-\frac{1}{\sigma}} c_2^{-\frac{1}{\sigma}}. \tag{9}$$

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.[14] 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 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

$$s^* = f(\tau) y, \tag{10}$$

---

[13]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.

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

where

$$f(\tau) = \frac{\beta^\sigma \left(1 - \tau \left(1 - \left(\frac{\tau}{p}\right)^\gamma \frac{1}{1+\gamma}\right)\right)^{\sigma-1}}{1 + \beta^\sigma \left(1 - \tau \left(1 - \left(\frac{\tau}{p}\right)^\gamma \frac{1}{1+\gamma}\right)\right)^{\sigma-1}}, \tag{11}$$

and taxable net wealth becomes

$$s^* - e^* = f(\tau) \left(1 - \left(\frac{\tau}{p}\right)^\gamma\right) y. \tag{12}$$

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:

$$\frac{B}{k(z^*) z^*} \approx \left(\frac{f(\tau_0) \left(1 - \left(\frac{\tau_0}{p}\right)^\gamma\right)}{f(\tau_1) \left(1 - \left(\frac{\tau_1}{p}\right)^\gamma\right)} - 1\right). \tag{13}$$

In (13), $k(z^*)$ denotes the density of the distribution of taxable net wealth at the kink point with a linear tax rate. Equation (13) is a generalized version of equation (6). 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 $\tau_0 = 0$, the following approximation holds:

$$\frac{B}{z^* k(z^*)} \approx \log\left(\frac{\beta^\sigma}{1 + \beta^\sigma}\right) - \log\left(\frac{\beta^\sigma \left(1 - \tau_1 \left(1 - \left(\frac{\tau_1}{p}\right)^\gamma \frac{1}{1+\gamma}\right)\right)^{\sigma-1}}{1 + \beta^\sigma \left(1 - \tau_1 \left(1 - \left(\frac{\tau_1}{p}\right)^\gamma \frac{1}{1+\gamma}\right)\right)^{\sigma-1}}\right) + \left(\frac{\tau_1}{p}\right)^\gamma. \tag{14}$$

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 evaded savings, where higher evasion adds to bunching.[15] Equation (14) 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

---

[15]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.

bunching.

According to equation (14), 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:

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

$$\varepsilon_{e,\tau}^D = -(1-\lambda)(\sigma-1)\left(1-\tau+\tau\frac{1}{1+\gamma}\left(\frac{\tau}{p}\right)^\gamma\right)^{-1}(1-f(\tau))+\frac{\gamma}{\tau}$$ (16)

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. In Piketty and Saez (2012), agents' preferences are defined over wealth, bequests and consumption and therefore real savings responses to tax changes reflect many wealth accumulation motives.

2.4 Bounded Rationality

The framework considered so far may be appropriate for studying how the wealth tax affects savings, evasion, and aggregate wealth inequality. However, it does not allow for innate heterogeneity in the ability to fathom the tax system across groups.

To allow for heterogenous responses to the tax rate, based on the ability to perceive the tax, I follow the approach in Chetty, Looney and Kroft (2007). I formulate a model where agents rationally ignore the tax if the utility gain from including the tax in the optimization decision is lower than some exogenously given cost.[16] This is in line with the bounded-rationality literature, which assumes that it is costly to acquire or process information and that agents therefore overlook information that would lead to optimal choices in the absence of these costs.[17]

---

[16]This optimization problem has the perplex feature that the decision of whether to remain inattentive or not requires not only knowing the utility from full optimization but also the utility from inattentive optimization. Conlisk (1996) refers to this phenomenon as the *regression problem*. In practice, individuals will presumably know the losses from ignoring the tax, at least approximately.

[17]In this vein, Reis (2006) shows how agents optimally choose the interval length between updates of their financial portfolio. In a series of papers, Sims (2003 and 2006) develops the idea that agents are instead limited in the amount of information they can process.

Since an inattentive agent ignores the tax when optimizing, he evades no money in the static model. The gain – measured in monetary units – from taking the tax rate into account is thus given by

$$G\left(\tau\right) = s\left(1 - \tau + \tau\frac{1}{1+\gamma}\left(\frac{\tau}{p}\right)^{\gamma}\right) - s\left(1 - \tau\right) = \tau\frac{1}{1+\gamma}\left(\frac{\tau}{p}\right)^{\gamma}s. \qquad (17)$$

Agents optimize taking the tax rate into account if $G\left(\tau\right) > c$. In the following, I assume that c is heterogeneously distributed in the population and that c is negatively correlated with cognitive ability. If the gains from knowing the tax rate are small, aggregate inattention is higher. When the tax rate increases, the amount of inattention decreases.

The wealth tax considered in this paper is low which implies that the gains are small. For example, if the fraction evaded is 10 percent, $\gamma = 0.5$, $p = 0.86$ and $\tau = 0.015$ for wealth above SEK 1 million, an agent with SEK $1, 5$ million loses SEK 665, about USD 100, by ignoring the tax. However, if less able individuals are also myopic, ignoring the tax every year, small losses may grow to large differences over time. With an annual interest of 5 percent, the future value of the augmented losses, computed as an annuity, amounts to SEK $80, 211$ over a 40 year period. This is roughly equal to a third of median annual earnings in Sweden. Seemingly low gains from being attentive may thus produce amplified inequalities across skill groups. A natural policy intervention for mitigating the inequality effects of the tax would be to increase the penalty. Higher penalties decrease the redistributional distortions as evasion goes down and hence, the utility loss from ignoring the tax is lowered.

In the dynamic framework, an inattentive agent chooses how much to save while ignoring the tax rate. The perceived budget constraints are thus:

$$c_1 = y - s$$
$$c_2 = s.$$

However, the inattentive optimal-savings choice will entail an infeasible consumption plan since the agents are nonetheless tax liable. Consumption has to adjust in view of the optimization mistake. A natural assumption is that households choose first-period consumption according to their first-order condition and period-two consumption equals the residual after the tax has been paid.[18] Then actual consumption in period 2 is:

$$c_2 = s\left(1 - \tau\right). \qquad (18)$$

An inattentive agent's first-order condition reads $c_1^{-\frac{1}{\sigma}} = \beta c_2^{-\frac{1}{\sigma}}$ and hence, $\hat{c}_1 = \left(1 - \beta^{\sigma}/\left(1 + \beta^{\sigma}\right)\right)y$ and $\hat{c}_2 = \left(\beta^{\sigma}/\left(1 + \beta^{\sigma}\right)\right)\left(1 - \tau\right)y$. Agents rationally ignore the tax if the utility from this consumption plan is close to the utility obtained from

---

[18]The assumption of when the tax is payed does not influence the results.

optimizing with the tax rate. This is true if

$$G\left(\tau\right) = \frac{c_1^{*1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + \beta\frac{c_2^{*1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \frac{\hat{c}_1^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \beta\frac{\hat{c}_2^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} \tag{19}$$

is lower than some cognitive cost c. The above expression can be restated as

$$G\left(\tau\right) = \left(\left(1-f\left(\tau\right)\right)^{1-\frac{1}{\sigma}} - \left(1-f\left(0\right)\right)^{1-\frac{1}{\sigma}}\right)\frac{y^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} + \tag{20}$$

$$\beta\left(\left(\left(1-\tau+\tau\left(\frac{\tau}{p}\right)^{\gamma}\frac{1}{1+\gamma}\right)f\left(\tau\right)\right)^{1-\frac{1}{\sigma}} - \left(\left(1-\tau\right)f\left(0\right)\right)^{1-\frac{1}{\sigma}}\right)\frac{y^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}, \tag{21}$$

where I have used the fact that the savings rule of an inattentive agent is the same as that of a rational agent when the tax rate is zero.[19]

As before, a discontinuity in the marginal tax rate heterogeneously triggers bunching depending on the utility gain from bunching at the kink as opposed to ignoring the tax rate. Let us assume that the cost of internalizing the tax rate is heterogeneously distributed within the groups of low- and high-skilled, LS and HS respectively, and assume that $\mathrm{E}_{LS}\left[c\right] > \mathrm{E}_{HS}\left[c\right]$. Then, the fraction of agents who optimize under the tax rate close to the kink is larger among high-skilled than low-skilled. As the tax rate increases, the utility difference increases and more individuals rationally cease to ignore the tax. A testable prediction of this framework is thus heterogenous bunching across skill groups. Importantly, in this model, heterogeneity in bunching does not reflect preferential differences or discrepancy in the evasion technology. With high tax rates, the cost of ignoring the tax is large and aggregate behavior is similar to that of a rational agent.

3 Institutional Background and Data

In a comprehensive tax reform dated 1947, the Swedish parliament supplemented the existing inheritance and gift tax with a separate progressive annual wealth tax and an estate tax.[20] The adoption of the reform was preceded by intense debate. In the opening speech of a meeting with the Swedish Economic Association in 1947, Eli Heckscher criticized the higher taxes on wealth not only for their distortive effects on private savings but also for the risk of increased tax avoidance.[21] The difficulties associated with the legal implementation of an annual wealth tax are also recognized in the recent work by Adam et al. (2011). The proper valuation of assets is difficult and impractical when

---

[19]If agents with low ability, i.e. high costs, are also more myopic than agents with high ability, with $\beta_L < \beta_H$, low-skilled agents may require an even larger utility gain than high-skilled in order to optimize with the tax rate. This result is an effect of the assumed budget rule with period two consumption adjusting to the optimization mistake and it is obtained when the savings rate of an attentive agent is higher than that of an inattentive agent.

[20]Formally, wealth was taxed from 1910, but before 1947 a fraction of the wealth was added to income and taxed through the income tax system.

[21]See Nationalekonomiska föreningen (1948). Heckscher criticized the proposed wealth taxes for cultivating a tax-avoidance norm, rather than inducing individuals to behave as if the price on transfers or savings was raised. Ohlsson (2011) gives a detailed summary of the events surrounding the reform.

defining a broad tax base, which includes assets and liabilities that are not traded or priced on a regular basis. In addition, some taxable assets and liabilities are self-reported which opens up the possibility to evade. In line with these concerns, the results from a survey aimed at eliciting perceptions of tax cheating in Sweden in 2006 indicate that individuals perceiving tax evasion as common believe the wealth tax to be the tax most likely to be subject to evasion (Hammar et al., 2006).

3.1 Institutional Setup

While the Swedish estate tax was repealed in 1953, the annual wealth tax was in place until 2007.[22] Taxable assets consisted primarily of shares in publicly traded companies, bonds, bank-account holdings, real estate, cars, boats and capital insurance. The wealth tax base was defined as the total value of these taxable assets net of liabilities, like real-estate mortgages and consumption loans, i.e., net worth. In 1991, a system with three marginal tax brackets was converted into a two-bracket system, with a zero marginal tax rate for net wealth below SEK 900,000 and a marginal tax rate of 1.5 percent for net wealth above this threshold.[23] The tax was filed jointly for couples with children below 18 years of age. As of 2000, the threshold was different for singles and for couples who were required to file jointly. During the period 1999-2006, the threshold was increased several times, as displayed in Figure 1. Approximately 8 percent of the population paid the wealth tax in 1999.

The filing of the wealth tax occurred in the spring of year $t+1$ for wealth holdings as of December 31 in year t. The Swedish Tax Agency (Skatteverket) sent out prepopulated tax forms which were based on third-party reports from banks, investment funds, brokers and other financial institutions. Figure 2 displays a prepopulated tax return. As the tax base included non-third party reported assets such as cars, boats, securities and liabilities held abroad, and debt within families, tax payers were required to self-report such holdings. The taxable net wealth equalled the sum of third-party reported net wealth and self-reported net wealth. The form in Figure 3, which was appended to the return sheet, explains how to calculate the tax liability. Upon receiving the tax form, the taxpayer was allowed to make adjustments and submit a final return by May 1 in year $t + 1$. In the analysis, I investigate the effects of tax reforms on both third-party reported net wealth, and taxable net wealth.

The main purpose of the wealth tax was redistribution in recognition of the potentially distortive effects of the tax on investments. To avoid the displacement of firms abroad, various tax exemptions, which narrowed the tax base, were installed over time. Thus, stocks not listed or traded on organized exchanges were not subject to wealth taxes. Moreover, as the tax was considered to be a deterrent to stock enlistment of companies in publicly traded markets, company ownership above 25 percent was tax exempt. Although the taxable amount should, in principle, reflect market value, some stocks were taxed only at 80 percent of their market value, while other stocks remained

---

[22]The tax on inheritance and *inter vivos* gifts was abolished in 2005.

[23]$ 1 \approx 6.5$ SEK, implying a threshold equivalent to USD 130,000.

completely tax exempt. Incentives for the placement of wealth in third-party reported assets that legally avoided the wealth tax were amplified by retirement savings being exempt from the tax.

While individuals may have responded to tax changes through strategic portfolio choices, the wealth tax was not associated with deduction opportunities. However, individuals with low income and high net wealth were tax exempt to some degree. Both a general law stipulating that total tax liability should never exceed 60 percent of income and a specific law entailing a reduction in the wealth tax for households with low income and high real-estate value were in place during the latter years of the study (2005-2006).[24] In case of excess tax liability, the wealth tax was lessened, but not by more than 50 percent of the pre-limit liabilities. The wealth tax could not be exempted in its entirety.

In addition to the wealth taxes, real estate was taxed annually at 1 percent of the (assessed) taxable value. Movements in the wealth tax bracket between years t and $t+1$ were, in practice, indexed against changes in the taxable value of real estate.[25] When the tax value for real estate was revalued, the wealth tax was also reformed to avoid implausible increases in tax liability. By the end of the 1990s, the government renewed its procedure for computing the taxable value of real estate, implying substantial increases in tax liabilities. These were accommodated by movements in the wealth tax bracket several times at the beginning of the 2000s. The indexation of the wealth tax bracket to the real estate tax suggests that the increase in the threshold was not driven by a powerful lobby of wealthy households.

3.2   Data and Sample Restrictions

The data I use in this paper come from the following administrative registers provided by Statistics Sweden: (i) The Income and Tax Register (Inkomst- och Taxeringsregistret); (ii) The Integrated Database for Labour Market Research (LISA); and (iii) Military Enlistment Data from the National Service Administration (Pliktverket).

From The Income and Tax Register, I retrieve all third-party reported asset items, taxable assets, taxable liabilities, liable wealth taxes as well as cheating penalties for each Swedish taxpayer above 15 years of age over the years 1999-2006. Using the third-party reports and tax records from the Swedish Tax Agency, I calculate two distinct measures of net wealth: third-party reported net wealth and taxable net wealth. The latter measure coincides with the former if no self-reported adjustments were made by the taxpayer. In the tax records, taxable net wealth is often not reported for households that do not pay wealth taxes. For the tax agency, keeping track of taxable net wealth for households that were not tax liable was of minor interest. In case the third-party reported net wealth was below the kink and the household did not pay any wealth tax,

---

[24]The general tax reduction law was in place during the whole sample period.

[25]See the Swedish Government Official Reports (2004). Henrekson and Jakobsson (2001) provide a survey of the impacts of wealth taxes on business ownerships in Sweden and recapitulate the political agenda behind various reforms.

I assume that taxable net wealth was equal to third-party reported net wealth.[26]

For households with a prepopulated third-party reported net wealth above the kink that end up below the kink, taxable net wealth is also not reported in a number of cases. I deal with this in two ways. In the first approach, I do not include these households in the bunching estimation procedure at all. In the second approach, I assume that the distribution of reported taxable net wealth for those below the kink is representative for those where taxable net wealth is missing and extrapolate values to those with no reported taxable net wealth. I therefore refer to the second approach as the extrapolation method.[27] While both approaches imply a lower density to the right of the kink, the bunching estimate will be lower in the first approach.

Demographic information on individual characteristics such as age, education, occupation, wage earnings and family status, is collected from the LISA database, which includes both spouses' social-security numbers, enabling me to link couples filing the wealth tax jointly.[28] As in The Income and Tax Register, this database comprises individuals above 15 years of age. I am able to match 99.9 percent of the tax payers to the demographic database, yielding a matched dataset consisting of $58,015,897$ observations over the period 1999-2006.

Since the wealth tax was filed jointly by households, with the sum of all household members' net wealth constituting the taxable net wealth, lack of data on individuals below 16 years of age implies that I am not able to assess taxable net wealth figures for households with children. Fortunately, the demographic dataset contains information about household status, including information about children below 18 years of age. I thus confine the sample to consist of single households and couples without children. This results in $20,773,835$ observations of single households and $6,961,055$ entries of couples filing the tax jointly. In sum, this represents 60 percent of the total number of observations. Out of these, 6 percent, or $1,668,465$ observations include positive wealth tax payments. Self-employed individuals who used their assets in business activities were tax exempt. To avoid the results being driven by the self-employed, I restrict the sample by excluding individuals with assets in industrial property, agricultural property and rental property. Based on the aforementioned databases, couples in which one of the spouses possesses such assets are also removed. This implies dropping $2,011,649$ observations.[29]

---

[26]If the (imaginary) self-reported wealth for households below the kink behaved in the same way as for households above the kink, this procedure attenuates the bunching estimates. To the right of the kink, I find that households, in general, adjust their wealth downwards. If this were true to the left of the kink, such self-reports shift the distribution downwards. However, bunching at the kink reflects households coming from the right of the kink and does not shift. This implies that estimates are lower without knowing self-reports.

[27]This method may yield biased results if the sample with missing taxable net wealth values is endogenously selected. Running a regression on the sample of households with third-party reported net wealth above the kink who pay no wealth taxes, I find that an indicator for having a value is not significantly correlated with cognitive skills or wage earnings.

[28]In addition to married couples, the database includes information on common-law spouses. The final dataset comprises both types of relationships.

[29]To investigate whether this sample confinement spuriously generates bunching, I estimate bunching with the self-employed in the sample. The results are stronger when including self-employed with a bunching estimate $b = 1.01$, compared with 0.53 for the sample retained.

To obtain data on cognitive skills, I exploit psychological tests from the military draft. Before enlisting in the military, all men in Sweden were drafted and had to go through two days of various testing. The test procedure was in principle mandatory until 2010, but since the number of candidates fell in the 2000s, I confine the sample to include only cohorts of men born from 1951 until 1979. Approximately 90 percent of all men in my data who were born in the defined time period enlisted with the military.[30]

The enlistment usually took place the year in which the candidates turned 18. Apart from physical tests and a semi-structured interview with a psychologist that evaluated noncognitive skills, a cognitive skills test was taken.[31] This test consisted of four subtests with 40 questions each that evaluated logical, verbal and spatial capabilities and a test that evaluated the conscript's technical comprehension. Each subtest yielded a score between 4-36 points. As the test was subject to minor revisions in 1980 and 1994, I create a normalized measure of cognitive skills by ranking the sums of the subtest within enlistment years by percentile, and applying the inverse of the standard normal distribution to obtain a standard normal measure of cognitive skills. $4,103,044$ household-year observations of the remaining sample of tax payers were matched with enlistment data. No women appear in the resulting dataset and the unit of observation is either a single man or an enlisted man in a couple.

Table 1 summarizes the evolution of the number of taxpayers and government revenue over time. In 2001 and 2002, the threshold for paying wealth taxes was raised both for singles and couples filing jointly, thus reducing the revenue from the wealth tax, along with the number of taxpayers. Table 2 presents summary statistics for different subsamples. Consistent with a standard savings profile over the life-cycle, the table suggests that wealth tax payers are considerably older than the rest of the population. Since the oldest individuals in the sample with enlistment data were born in 1951, these households are, on average, younger and hold less assets than other households.

4   Empirical Analysis of Bunching

This section uses the data described in Section 3 to estimate the parameters needed for computing the tax elasticities implied by the framework in Section 2. These implied tax elasticities are presented in Section 5. I start by estimating bunching, i.e. the excess mass in the distribution of taxable net wealth at the kink point. In order to obtain an estimate, the counterfactual density, i.e. the mass at the kink point if the tax rate were zero, must be gauged. I do this in two ways. The first approach, described in Section 4.1, makes my results directly comparable to the literature. Here, I use the parametric

---

[30]There is some variation in the availability of enlistment data over the years. For example, the enlistment data only includes about 40 percent of the men enlisting in the year 1985. There is no official reason for the varying data availability according to officials at the National Service Administration. In general, compliance is large and, according to Grönqvist et al. (2010), only the physically and mentally handicapped were exempted from enlistment. Consequences of refusal included fines and, even, imprisonment.

[31]A comprehensive overview of the test procedure is found in Lindqvist and Vestman (2010). Carlstedt (2000) provides a review of the cognitive skills test, arguing that the test provides an accurate measure of general intelligence.

method for estimating bunching employed in Chetty, Friedman, Olsen and Pistaferri (2011). In Section 4.2, I instead take a non-parametric approach by estimating bunching from the two-dimensional data on third-party reported and taxable net wealth, thus capturing bunching occurring along the self-reported margin. In Section 4.3, I make the case that the self-reported margin may serve as a proxy for evasion. Finally, Section 4.4 calibrates the parameter p in the cost function (3), which is needed for identification of the parametrized models.

4.1 Parametric Estimation

Figure 4 plots the distribution of third-party reported net wealth for all Swedish households without children (below 18 years of age) for the years 1999-2006. In constructing the figure, I add up third-party reported assets, item by item, and subtract liabilities. The resulting number is cross-checked against the third-party reported measure of net wealth provided by Statistics Sweden. Figure 5 plots the corresponding distribution of taxable net wealth, obtained by adding the self-reported net wealth to the third-party reported net wealth. To obtain centralized measures of wealth while taking the changes in the threshold illustrated in Figure 1 into account, I calculate the difference between either wealth variable and the amount of wealth needed to reach the tax bracket with a positive marginal tax rate. Households are then grouped into SEK $5,000$ bins. Bin counts are plotted around the kink point, which is demarcated by the vertical line at zero.

Figure 4 reveals no bunching in the distribution of third-party reported net wealth around the kink point, but there is a marked spike in the otherwise smooth distribution of taxable net wealth at the kink point in Figure 5. To estimate bunching at the kink – a value needed to identify the parameters of interest – an estimate of the counterfactual density at the kink is required. However, in contrast to the theoretical predictions, the spike in the empirical distribution is diffused around the kink point. This presumably reflects households' inability to perfectly monitor savings to locate at the kink or, if households are bunching through evasion, their willingness to shroud tax evasion. When estimating bunching, I account for such noise. Specifically, to gauge the excess mass at the kink, I first estimate the distribution of taxable net wealth in the absence of taxation. Following Chetty, Friedman, Olsen and Pistaferri (2011), I fit a polynomial to the empirical distribution as follows:

$$N_j = \beta_0 + \beta_1 Z + \beta_2 Z^2 + \cdots + \beta_n Z^n + \sum_{i=-R}^{Q} \phi_i \mathbb{I}\left[Z_j = i\right] + \varepsilon_j, \qquad (22)$$

where N_j denotes the number of households in bin j, Z_j is wealth relative to the kink in SEK $5,000$ intervals and n is the order of the polynomial. The sum of indicator variables on the right-hand side reflects the exclusion of observations close to the kink point, as these reflect bunching. R and Q define the lower and upper bounds of this interval, respectively. The counterfactual density is given by the predicted values, $\hat{N}_j$, excluding

the contribution of the ϕ_i dummies around the kink. An estimate of the number of households bunching is thus $\hat{B} = \sum_{j=-R}^{0} N_j - \hat{N}_j$, i.e. the number of households in excess of the counterfactual density close to the kink point. The method overestimates the true value of $\hat{B}$, however, since the counterfactual density does not satisfy the integration constraint. To correct for this, I follow Chetty et al. (2011) and estimate a counterfactual density that shifts the counterfactual distribution to the right of the kink until the integration constraint is satisfied.[32] The relevant measure for the elasticity estimation is the estimated number of bunching households relative to the counterfactual density close to the kink point, measured as

$$\hat{b} = \frac{\hat{B}}{\frac{\sum_{j=-R}^{Q} \hat{N}_j}{Q+R-1}}. \tag{23}$$

Figure 6 plots the distribution of taxable net wealth together with the counterfactual distribution, estimated as a seven-degree polynomial. The window of bunching is defined as SEK $40,000$ below the kink, a value of -8 in the figure. Marginal changes in the window of bunching and the order of the polynomial do not affect the estimated excess mass to any considerable extent. I estimate bunching at $\hat{b} = 0.53$, meaning that there is 53 percent more mass relative to the counterfactual distribution within SEK $5,000$ of the kink.[33] Figure 7 presents the bunching estimates using the extrapolation method, i.e including households with missing taxable net wealth values, which produces an estimate of $\hat{b} = 1.25$.

The standard error for $\hat{b}$ is estimated by a parametric bootstrap procedure, as in Chetty et al. (2011). It reflects misspecification of the polynomial rather than sampling errors, as the estimate is constructed using the population distribution. The estimated standard error is 0.09, yielding a t-statistic on $\hat{b}$ of 5.89 and the null hypothesis of no bunching at the kink point is rejected with a p-value of 1.97×10^{-9}.

Figures 6 and 7 show a drop in the density to the right of the kink point, which is not explained by theory. Holes in the distribution arise in tax schemes with discontinuous jumps in average, not marginal, tax rates. One plausible explanation for this is that the time it takes to file the tax return and compute the value of non-third-party reported assets and liabilities constitutes a fixed cost. The existence of a fixed cost is analogous to a discontinuity in the average tax rate, i.e. a notch, predicting that there should be a hole in the distribution to the right of the kink.[34]

---

[32]The counterfactual density is in this case given by: $N_j \left(1 + \mathbb{I}[j > Q] \frac{\hat{B}}{\sum_{j=Q+1}^{\infty} N_j} \right) = \beta_0 + \beta_1 Z + \beta_2 Z^2 + \cdots + \beta_n Z^n + \sum_{i=-R}^{Q} \phi_i \mathbb{I}[Z_j = i] + \varepsilon_j$.

[33]One possibility that I am not able to address in the data, is that there is spurious bunching. Consider a household that is located to the right of the kink in third-party reported net wealth, but after truthful self-reporting would end up far to the left of the kink. If this household recognizes that the tax liabilities are equal just to the left of the kink and at the true point further to the left, it may choose to locate just below the threshold even though it is a tax complier. In ongoing work, I investigate this issue by comparing detailed administrative records on cars to the self-reported measure.

[34]Kleven and Waseem (2012) analyze bunching at thresholds in the income tax scheme when the average tax rate increases discontinually. This creates an even stronger incentive to bunch than do discontinuities in the marginal tax rate.

Figure 8 displays bunching in the distribution of third-party reported net wealth. Here, the estimated excess mass is only 0.10, but still statistically significant.

The identification of parameters in the theoretical framework relies on the assumption of the distribution of taxable net wealth being smooth in the absence of the wealth tax. This assumption can be relaxed by investigating movements in the tax bracket over time. The threshold defining the two tax brackets was different for singles and couples filing jointly. In addition, the threshold was changed several times and by separate amounts. Figures 9 and 10 display bunching around the kink in each year during the period 1999-2006 for singles and couples filing jointly, respectively. The estimated excess mass follows the kink closely over time. Figure 11 further investigates whether bunching follows the tax rate over time, or if alternative explanations account for the evolution of bunching. I compare the distribution of taxable net wealth for singles in 2001 to that of 2006, a time period marked by an increase in the threshold by SEK $500,000$. The excess mass in 2001 is located at the tax threshold and the figure presents three hypothetical scenarios for the location of the kink in 2006. The first placebo-kink denotes the threshold value that would be obtained had it followed inflation, the second location indicates the same value had it followed the riskfree interest rate and the final kink denotes the value had it followed the Stockholm Stock Exchange Index. The figure confirms that the kink does move to the 2006 tax threshold.

Turning to heterogeneity in bunching, Figures 9 and 10, in general, reveal no differences between singles and couples without children that file the tax jointly. The largest estimates of bunching for each subgroup are $\hat{b} = 1.4$ for singles and $\hat{b} = 1.54$ for couples.[35] I next turn to differences in excess mass across skill groups. To maintain statistical power, I employ the extrapolation method when performing this analysis. Figures 12 and 13 present taxable net wealth distributions by two skill groups defined by having obtained positive or negative z-scores on the cognitive military draft test.

The differences in bunching across cognitive skill groups are more pronounced over the years 2002-2006, a period during which the kink was held constant for single households, as indicated in Figures 14 and 15. The estimated bunching is 1.1 for the high-skilled and 0.42 for the low-skilled, i.e. almost three times as large for the high-skilled group. Comparing the top 25th percentile of the skill distribution to the bottom 25 percent in Figure 15, the coefficient of bunching is almost four times larger for high-skilled households than for low-skilled households. The difference in bunching across skill groups is a key finding, which translates into larger tax responses among high-skilled than low-skilled. Such differences will ultimately affect the wealth inequality across skill groups.

4.2 Nonparametric Estimation

The parametric estimates of bunching in Section 4.1 rely on the estimated counterfactual density accurately reflecting the distribution of taxable net wealth. Since a higher marginal tax rate does not only affect households close to the kink, but all households to the right of the kink point, it is not obvious that the estimated counterfactual density

---

[35]Figure 9 shows a deviating pattern for 1999. The reasons for this deviation are difficult to address.

closely matches the true one. A biased estimate of the counterfactual distribution would not only bias the counterfactual density close to the kink, but also the estimated number of households bunching.

My second approach instead estimates bunching nonparametrically, by exploiting the paired observations of third-party reported net wealth and taxable net wealth for each household. The idea is to gauge bunching by computing the number of households located above the kink in third-party reported net wealth but below the kink in taxable (self-reported) net wealth. Specifically, I use the following estimator of bunching:

$$\hat{B} = \sum_i^N \mathbb{I}\left[W^* - \delta < W_i < W^* \text{ and } T_i > W^*\right], \qquad (24)$$

where W_i denotes taxable net wealth, T_i is third-party reported net wealth, W^* defines the kink and δ is a lower bound on the amount of bunching per household. Thus, the counterfactual density is given by the number of households to the left of the kink in terms of third-party reported net wealth. I group the third-party reported net wealth distribution into bins of SEK $5,000$ and use the number of households in the bin closest to the kink as a counterfactual.[36]

Table 3 presents the results using this method. The standard errors are computed using a bootstrap method, in which new distributions of third-party reported net wealth and taxable net wealth are drawn with replacement from the true distribution. The standard error of the bunching coefficient is represented by the standard deviation of the distribution of estimated bunching coefficients. Estimated excess mass is always significant, with the magnitude increasing in the bandwidth of allowed bunching. The coefficient is also larger when using the extrapolation method. The point estimates are of the same order of magnitude as those obtained using the parametric method. The heterogeneity in bunching across skills is validated. In contrast to the results discussed in Section 4.1, estimated bunching is always significant also for the low skilled.

The nonparametric approach provides a measure of bunching at the household level, a novel feature in the bunching-literature. The graphs presented in Figure 16 show how bunching follows the threshold. Panel (a) displays bunching of couples within a window of SEK $25,000$ below the kink of SEK $900,000$. In the years 1999 and 2000, this constituted the threshold for the marginal tax rate. However, in 2001 all bunching disappears. The same pattern is supported by the other graphs.

To summarize, the two estimation procedures give statistically significant estimates of bunching. The parametric approach has the advantage that it allows for the response to taxes to occur both through real and reporting responses. The nonparametric procedure, in contrast, presumes that the response occurs exclusively along the self-reporting margin. In the case of savings responses, the second approach, described by equation (24), thus underestimates the amount of bunching. On the other hand, bunching estimates from the first approach rely on the estimated counterfactual distribution being

---

[36]Taking the average of the third-party reported net wealth in the interval of $[W^* - \delta, W^*]$ does not affect the results qualitatively, and the impact on estimated magnitudes is small.

correct. In practice, the estimated excess mass will depend on the imposed interval width of bunching and functional form assumptions about the order of the counterfactual polynomial in (22). When associating the excess mass estimates with the theoretical models, I explore a range of estimated values of bunching as a robustness exercise.

4.3   A Proxy for Tax Evasion

The bunching documented above begs the question of whether the observed behavior is due to evasion or real savings responses. To address these questions, an empirical proxy for evasion is needed. I argue that a natural candidate for such a proxy is self-reported net wealth.

First, a simple comparison of the distributions of third-party reported and taxable net wealth in Figures 4 and 5, respectively, suggests that responses occur within the category of self-reported assets and liabilities. However, the absence of bunching in third-party reported net wealth does not automatically imply that savings are unaffected by the tax. If agents' possibility to evade were to be eliminated, bunching through real responses might arise.

Another way of assessing bunching through evasion is to study dynamic responses over time. Since wealth is a stock and not a flow variable, any bunching through real intertemporal responses would only manifest gradually over time. The estimated bunching upon shifting the kink should thus be different immediately after shifting the kink, as compared to, say, five years after. However, Figures 9 and 10 lend no support to the hypothesis that the excess mass estimates change over time. The observed pattern could be consistent with bunching through real responses if self-reported assets were more liquid than assets included in third-party reported wealth. However, liquid assets, such as bank account holdings and funds, are typical third-party reported assets while self-reported assets mainly comprise more illiquid assets such as cars and boats.

Second, if self-reported wealth indeed constitutes a proxy for tax evasion, it should be mirrored in monetary penalties for tax sheltering. Regressing imposed penalties on self-reported wealth and the positive difference between third-party reported and taxable net wealth, I do find that larger differences are associated with larger fines. The estimates, presented in Table 4, suggest that a one SEK increase in the positive difference gives an increase in the imposed fine in the range of SEK 0.04-0.17, depending on the specification.

These findings suggest that the difference between third-party reported and taxable net wealth may serve as a proxy for tax evasion. In the analysis that follows, I therefore treat it as such and refer to this difference as my measure of tax evasion or tax sheltering.

4.4   Estimating p

In addition to bunching estimates, parametrizing the models of Section 2 and computing the implied tax elasticities requires knowledge of the unit cost of tax sheltering. The cost function, $C\left(e,s\right) = \left(\frac{e}{s}\right)^{\gamma}\frac{pe}{1+\gamma}$, features a unit cost of evasion, denoted p. To measure

p, I use data on the total value of sentenced fines per person. In practice, this cost should be viewed as a lower bound on the actual unit cost of evasion. First, it does not include transfers across non-government agents such as payments to accountants and tax planners. Second, according to officials at the Tax Authority, the detection of wealth tax sheltering increased the risk of being caught evading other taxes. Third, a history of tax fraud was a key parameter of the Tax Authority's auditing strategy, implying that detection was likely to lead to more frequent audits in the future. Finally, the measure does not include the (monetary) cost of going to court.[37] Nevertheless, sentenced fines should be able to serve as a proxy for the cost of tax sheltering.

Ignoring functional form assumptions regarding the relation between evasion and penalties, the mean amount of imposed fines, as a fraction of tax evasion, among the subset of individuals who do evade and are detected, is 0.86. In my derivation of the tax elasticities in the next section, $p = 0.86$ is thus chosen as the benchmark estimate.

5 Estimating Elasticities

In this section, I use the bunching and evasion cost estimates from Section 4 to compute tax elasticities according to the theories in Section 2. In Section 5.1, I use the framework for small-kink analysis, presented in Section 2.1. The sheltering parameter in the static model, layed out in Section 2.2, is identified in Section 5.2. In Section 5.3, I extend the analysis to include a savings response in accordance with the dynamic model of Section 2.3. Finally, in Section 5.4, I relate my results to the bounded-rationality model of Section 2.4 and analyze heterogenous bunching on the basis of cognitive ability.

5.1 Small Kink Analysis

Recall that equation (1), $B/h\left(z^*\right)z^* = -\varepsilon_{W,\tau}^K \mathrm{d}\tau$, identifies the compensated tax semi-elasticity of taxable net wealth, $\varepsilon_{W,\tau}^K$, without making any parametric assumptions about the preferences generating observed behavior. The approach is valid, as long as the kink is small and the associated income effects are negligible. The first column of Table 5 presents estimated elasticities based on different estimates of bunching obtained using the parametric and nonparametric methods of Section 4. The elasticities are in the range $[0.12, 0.32]$ with higher elasticities reflecting larger bunching estimates.

An important issue when analyzing the impact of wealth taxes is the time horizon considered. How much of the effect occurs already in the short run as compared to the long run? I use the small kink analysis to assess the evolution of the tax elasticities over time. Incorporating shifts in the threshold, the results are indexed by the time that the kink has remained constant. In 2002, the kink was shifted from SEK $1,000,000$ to $1,500,000$ for single households and was then held fixed at the latter level, until the tax was repealed in 2006. The year-by-year plots of taxable net wealth in Figure 9 reveal

---

[37]In an interview with representatives from the Tax Authority (Skatteverket), the evaded amount is only one parameter that determines whether the case is brought to court. Small amounts of tax sheltering may also lead to court if the tax authority can verify a history of cheating.

that the extent of bunching is fairly stable during the period 2002-2006, which yields steady elasticity estimates over time. Analyzing third-party reported net wealth each year, as in Figure 17, suggests no dynamic effects on bunching in accumulated savings over time.[38]

Next, I estimate taxable net wealth elasticities using variation over time, i.e., shifts in the threshold. Changes in the kink generate different changes in tax rates across households depending on their net wealth. This motivates a difference-in-difference research design.

Let $\Delta \log W_{i,t}^x = \log W_{i,t} - \log W_{i,t-x}$ represent the log change in taxable net wealth between year t and $t - x$, (for $x = 1, 2, 3$) and let $\Delta \log NTR_{i,t}^x$ denote the log change in net-of-tax rates over the same time period. Following Gruber and Saez (2002), I then estimate:

$$\Delta \log W_{i,t}^x = \alpha_i + \theta \Delta \log NTR_{i,t}^x + f(W_{i,t-x}) + \mu X_{i,t-x} + \varepsilon_{i,t} \qquad (25)$$

using two-stage-least-squares (2SLS). I instrument for the log change in the net-of-tax rate, $\Delta \log NTR_{i,t}^x$, by $\Delta \log NTR_{i,t}^{x,sim}$, the simulated change in the net-of-tax rate holding the household's wealth equal to its level in the base year. $f(W_{i,t-x})$ denotes a 10-piece linear spline in taxable net wealth and $X_{i,t-x}$, a vector of household characteristics. Both are measured at the base year. The regression includes a household fixed effect α_i, capturing household-specific time trends.

When estimating (25) by 2SLS, the first-stage regressions of $\Delta \log NTR_{i,t}^x$ on $\Delta \log NTR_{i,t}^{x,sim}$ yield coefficients around 0.75 with implied t-statistics above 50. To avoid mean reversion in wealth driving the results, I exclude households with base-year taxable net wealth values below SEK $800,000$. Columns (5) in Table 5 present estimated tax elasticities of taxable net wealth. All specifications include year fixed effects and a spline in base-year taxable net wealth. The estimated elasticities are approximately semi-elasticities, given that $\Delta \log(1 - \tau) \approx \Delta \tau$ for low taxes. Exploiting shifts in the threshold yields larger estimates of the tax elasticity, in the range $[0.45, 0.80]$, depending on the chosen time interval.

5.2   Static Model

The static model in Section 2.2 assumes that the response to the wealth tax occurs along the evasion margin. In the parametrized model, bunching at the kink reflects the fraction of sheltered wealth. Larger estimates of excess mass thus imply larger fractions of evasion. Using the benchmark estimated value of the unit cost of evasion, $\hat{p} = 0.86$, column (2) of Table 5 presents tax semi-elasticities of taxable net wealth for a range of bunching estimates. The elasticities are slightly larger compared to the small-kink analysis and take values in the range $[0.18, 0.42]$.

---

[38]Figures 10 and 18 display bunching of taxable net wealth and third-party reported net wealth, respectively, over time, for couples. These graphs corroborate the finding that there are no dynamic responses of savings.

The semi-elasticities of evasion, presented in column (1) of Table 6, are substantial, ranging between $[87.8, 105.7]$. The large magnitudes arise as the estimated fraction of evasion is small. At low levels of evasion, the percentage change in evasion arising from a one percentage point increase in the tax rate is large. However, the tax elasticities of evasion are smaller and the estimates are slightly larger than one, indicated by column (2) of Table 6. At higher tax rates, for example, $\tau = 0.35$, the response is lower partly due to higher fractions evaded, with semi-elasticities around 2.5. Such a hypothetical tax rate is high in the context of a wealth tax, but when applying the framework to tax sheltering of, for instance, labor income taxes, such an example is realistic. On the other hand, the evasion cost technology is different for higher tax rates and other tax rates, in particular labor-income taxes, so out-of-sample predictions should be interpreted with caution.[39] The implied semi-elasticities of around 2.5 are in line with those in Fisman and Wei (2004), that estimate a semi-elasticity of evasion of 3 using tax rates in the range of 30 to 50 percent, and within the same range as the estimates between 0.5 and 3 reported by Clotfelter (1983).

Using data on self-reported downward-adjustments of taxable net wealth, I next study how much money is saved by bunchers. The estimates vary slightly depending on whether extrapolation is used. A conservative estimate not using the extrapolation method reveals that the median shelters SEK $220,000$ from the government and therefore gains SEK $3,300$ from not paying the tax. The average values correspond to SEK $400,000$ and SEK $6,000$.[40]

The parameterized model predicts a linear relation between sheltered resources and wealth. Figure 19 tests this by plotting mean evasion against third-party reported net wealth. When estimating this relationship by means of regression, higher order terms are statistically significant but the magnitude of the estimated coefficient is practically zero.

5.3 Dynamic Model

When I allow bunching to arise from adjustment on two margins, equation (14) states that bunching comprises the sum of the savings responses and reporting responses. In the parameterization of the model, I use standard values from the macro literature: $\beta = 0.96$ and $\sigma = 0.25$.[41] The results are not sensitive to how these parameters are chosen and I also present results for $\sigma = 1.1$. With $\sigma < 1$, the fraction of income saved is larger on the right-side of the kink than on the left-hand side. To make estimated bunching consistent with this fact, the estimated evaded fraction has to be larger here compared to the case when $\sigma > 1$. Moreover, such a parameter value implies that a tax increase triggers evasion both directly and indirectly due to increased savings. In contrast, if

---

[39]Kleven et al. (2011) and Slemrod and Kopczuk (2002) argue that the responsiveness of evasion to tax changes depends on various parameters including the enforcement regime and the definition of the tax base.

[40]These computations allow for bunching in an interval of SEK $100,000$ below the kink.

[41]$\beta = 0.96$ reflects that the period length corresponds to one year.

$\sigma > 1$, agents respond to a tax increase by both reducing savings and increasing evasion. The response of evasion is mitigated by a decrease in savings.

Columns (3) and (4) of Table 5 present uncompensated tax semi-elasticities of taxable net wealth using different excess mass estimates.[42] The uncompensated elasticities – now reflecting both intertemporal substitution and tax sheltering responses – are close to the estimates obtained in the static framework. A lower σ implies more evasion and a larger tax elasticity of taxable net wealth. For $\sigma = 0.25$, the elasticities range in $[-0.20, -0.40]$.

The bunching estimates and the parameter choices imply that the real responses to tax changes are almost zero and the responses to tax changes occur mainly along the tax-sheltering margin. This is manifested in columns (3) - (6) in Table 6, displaying uncompensated and income-compensated semi-elasticities and regular elasticities of evasion. Evasion responses are slightly larger in the dynamic case. Even though the savings response is an outcome of the parameterization, varying the parameter values persistently yields this result.

5.4   Bounded Rationality

In the theory presented in Section 2.4, the cognitively less able find it too costly to take the tax rate into account when optimizing. Figure 12 presents the distribution of taxable net wealth for high- and low-skilled households over the full sample period, while Figures 13 and 14 present the corresponding distributions and an associated bunching estimate for the years 2002-2006. To maintain statistical power, I retain the extrapolation method described in Section 3 for dealing with missing taxable wealth values throughout the analysis of cognitive skills. The analysis in this section is primarily conducted on the stable years 2002-2006, i.e., the period during which the kink was constant for single households and only shifted once for couples.

The underlying cognitive ability variable is normally distributed with zero mean and unit variance. The samples in the figures are based on dividing the sample into two groups with positive and negative z-scores, respectively. Figure 15 compares the top 25th percentile of the skill distribution to the bottom 25th. The observed elasticities, estimated using the small kink analysis summarized by equation (1), $\frac{B}{h(z^*)z^*} = -\varepsilon_{W,\tau}^K d\tau$ and the parametric estimates in Figures 14 and 15, are presented in Table 5, in the rows denoted $Parametric$ and $High$ and Low.

The nonparametric estimates of bunching based on cognitive skills are displayed in Table 3, denoted by b_{HIGH} and b_{LOW}. The associated elasticities are found in the rows labeled $Nonparametric$ and $High$ and Low, respectively, of Table 5. For the high-skilled, the elasticities range between $[0.20, 0.50]$, whereas the corresponding interval for low-skilled is $[0.05, 0.32]$. These elasticities are not structural parameters, as they also reflect the optimization errors incurred by bounded rationality.

---

[42]The compensated tax elasticities of taxable net wealth are positive. Neither savings nor taxable net wealth are arguments in the agents' utility function. A tax increase implies lowered consumption in the second period, also if the tax increase is compensated for, but the impact on taxable net wealth is positive.

Exploiting my empirical proxy for tax sheltering, I test the theoretical prediction that high-skilled households evade more. This is done in two steps. First, I investigate self-reports that involve bunching, i.e., households locating above the kink in third-party reports and below the kink in taxable net wealth. Second, I study the intensive margin, namely the amount of self-reported adjustment in SEK.

To investigate extensive-margin responses, I regress a binary indicator of bunching on a 10-piece linear spline in third-party reported net wealth, wage earnings and education, on the sample of men with enlistment data located between SEK 1.5 and 2 million in terms of third-party reported net wealth. Figure 20 plots the residuals from this regression against cognitive skills. This captures the partial relation between skills and bunching while controlling for wealth, income and education. The results suggest that, conditional on important factors, cognitive skills and bunching are positively related.

The intensive-margin response is investigated in a similar way. I regress the evaded amount on a 10-piece linear spline in third-party reported net wealth, wage earnings and education. Figure 21 plots the residuals from this regression against cognitive skills. Although less obvious, there is still a positive relation between self-reported wealth and cognitive skills also along this margin.

I quantify this positive relation in the regressions presented in Table 8. Controlling for education, third-party reported net wealth and wage earnings, a one-standard-deviation increase in cognitive ability is associated with an increase in the probability of bunching by around 1 percent. The same increase in skills increases the percentage evaded by about 0.14 percent.

These results provide consistent support for the prediction that low-ability households respond less to tax changes than households with higher ability. I discuss alternative mechanisms behind this pattern in the next section.

6 Heterogeneous Responses to Wealth Taxes

The results in Section 5.4 provide support for the prediction that households with a low cognitive ability respond less to tax changes than households with a higher ability. The bounded-rationality model presumes that these findings are driven by different abilities to understand the tax system. However, alternative channels could also explain the correlation between cognitive ability and evasion. I next explore the two alternative mechanisms that are arguably the strongest contenders: heterogenous costs of tax sheltering and heterogenous preferences.

To address the possibility that agents have access to different technologies for tax evasion, I regress sentenced fines on cognitive skills, controlling for evasion, wealth and income levels. The results, presented in columns (1) and (2) in Table 9, confirm that once evasion, wealth and income are taken into account, cognitive skills are of no importance for the amount of sentenced penalties. In rare instances, it is marginally statistically significant but with the wrong sign, i.e., suggesting a positive correlation between skills and

sentenced fines. These findings are consistent with the view that the same technologies are available to all.

To investigate whether different preferences are driving the results, I use a unique dataset on risk and time preferences obtained from a survey in Mollerstrom and Seim (2012). The survey elicits these preferences by survey questions developed in lab experiments and proven to correlate closely with results from actual experiments.[43] To compute a risk-aversion index, the subjects were asked to choose between a riskfree alternative and a gamble with two outcomes, holding the expected value of the gamble fixed but ranging the riskfree amount over a number of questions. The number of times the subject chose the riskfree alternative comprises a measure of risk aversion.

Similarly, a time-preference index is based on a choice between receiving money today and money in twelve months, where today's value was held fixed and the value in the future was varied. The impatience index measures the number of times the subject opted for receiving money today.

Matching these survey data at the individual level to enlistment data on cognitive ability yields a sample of 153 individuals. The results from regressing risk and time preferences on cognitive skills, displayed in columns (3) and (4) in Table 9, suggest no statistically significant correlation between them. In contrast to the previous literature, the point estimates would suggest that high-skilled individuals are more risk averse while, in line with earlier studies, the point estimates on impatience are negative.[44]

To summarize, I find support neither for heterogeneity in evasion technology across skill groups nor for heterogeneity in risk- and time preferences across skill groups.

7 Conclusion

In this paper, I address the behavioral effects of a wealth tax from different perspectives. I merge detailed administrative data with Swedish tax records over the period 1999-2006 and exploit features of the institutional setup to identify tax elasticities of taxable net wealth. In addition, I address potential mechanisms for the observed behavior and study whether effects are heterogeneous across skill groups.

The Swedish wealth tax comprised two tax brackets: one with a zero marginal tax rate and one with a marginal tax rate of 1.5 percent. I find strong evidence of bunching at the threshold separating the two brackets and report tax semi-elasticities of taxable net wealth in the absolute range $[0.12, 0.33]$. When I instead exploit the variation induced by movements of the threshold over time, I find larger semi-elasticities of taxable net wealth: in the range of 0.45 to 0.80.

Importantly, the tax base comprised both net wealth reported by third-parties, as well as self-reported assets and liabilities. This institutional feature allows me to disentangle savings responses from reporting responses. My results suggest that the

---

[43]The questions build directly on the work by Falk and Becker (2012), which aims at finding survey methods for eliciting risk and time preferences that are akin to lab experiments.

[44]Benjamin et al. (2006) and Frederick (2005) find that individuals with a high cognitive ability are less risk averse and less impatient.

observed bunching occurs exclusively along the self-reported margin. I conjecture that the tax triggers evasion and corroborate this by two pieces of evidence. First, I find that self-reported downward adjustments of taxable net wealth are positively correlated with monetary penalties for evasion. Second, following movements of the threshold over time, I find no evidence of dynamic adjustments in bunching along third-party reported net wealth and therefore no signs of intertemporal adjustments in terms of savings.

To learn more about the underlying mechanisms behind the observed effects, I explore different theoretical models and take them to the data using the bunching estimates. When translating my findings into tax semi-elasticities of taxable net wealth, my results suggest that they are in the range $[0.18, 0.42]$ when I employ a simple static setting, and in the range $[0.20, 0.40]$ when I explore a dynamic framework.

Since the tax system often is perceived as difficult to comprehend, I conjecture that individuals with high cognitive ability are more likely to understand the workings of the tax scheme. I test this by exploiting results from cognitive-skills tests from military enlistments. My findings suggest that the cognitively able respond more. The mass of households who locate just below the threshold for paying the positive marginal tax rate, in relation to an estimated counterfactual distribution, is greater for the subsample of high-skilled households than for low-skilled households. Also, on the intensive margin, I find that self-reported downward adjustments are larger for high-skilled households compared to low-skilled, even when controlling for income, wealth levels, type of wealth held and education. To address whether cognitive ability indeed lies behind the differences across groups, I explore key alternative explanations: having access to different technologies for evasion or different preferences with respect to risk and time. None of these alternatives are supported by the data.

Due to the scarce empirical evidence on taxes on wealth and intergenerational transfers, future research should aim at identifying the long-term effects of such taxes. The data used in the paper at hand only comprise nine years of observations and the horizon may be too short for long-run effects on accumulated assets to be observable. Additionally, the wealth-accumulation responses in the present paper arise from parameter values chosen from the macro literature. Future research should estimate, simultaneously, parameters that determine real and reporting responses. Obtaining accurate measures of the long-term impact of wealth and capital taxes on wealth accumulation may provide guidelines in the design of public policy.

Another interesting avenue of future research would be to explore taxes on bequests. Understanding whether households circumvent estate and inheritance taxes, or if these taxes induce real responses, is crucial for the design of future tax schemes related to inter-generational transfers.

On a final note, my findings on heterogeneous effects across skill groups demonstrate the merits of salient and simple tax policies. If governments cannot reduce the costs of internalizing tax policies, for instance by providing information about the workings of different tax schemes, they should make transparent and easily understandable tax policies a priority. Exploring how household heterogeneity affects optimal public policy

is a challenging but truly important task.

References

[1] Adam, Stuart, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles and James Poterba [2011], "Taxes on Wealth Transfers", in *Tax by Design. The Mirrlees Review*, Oxford University Press, chapter 14, 347–67.

[2] Agell, Jonas, Peter Englund and Jan Sodersten [1996], "Tax Reform of the Century – The Swedish Experiment", *National Tax Journal*, 49(4),643-664.

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

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

[5] Banks, James and Peter Diamond [2010], "The Base for Direct Taxation", in James A. Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Steve Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth D. Myles, and James Poterba, eds., *Dimensions of Tax Design: the Mirrlees Review*, Oxford University Press, 2010, chapter 6, 548-674.

[6] Benjamin, Daniel J., Sebastian A. Brown and Jesse Shapiro [2006], "Who is "Behavioral"? Cognitive Ability and Anomalous Preferences", Mimeo.

[7] Boadway, Robin, Emma Chamberlain and Carl Emmerson [2010], "Taxation of Wealth and Wealth Transfers", in James A. Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Steve Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth D. Myles, and James Poterba, eds., *Dimensions of Tax Design: the Mirrlees Review*, Oxford University Press, 2010, chapter 8, 737-836.

[8] Brown, Robert D. [1991], "A Primer on the Implentation of Wealth Taxes", *Canadian Public Policy / Analyse de Politiques*, 17(3), 335-350.

[9] Carlstedt, Berit [2000], "Cognitive Abilities: Aspects of Structural, Process and Measurement", Mimeo, Gothenburg University.

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

[11] Chetty, Raj [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.

[12] Chetty, Raj, Adam Looney and Kory Kroft [2007], "Salience and Taxation: Theory and Evidence", NBER Working Paper 13330.

[13] Chetty, Raj, Adam Looney and Kory Kroft [2009], "Salience and Taxation: Theory and Evidence", *The American Economic Review*, 99(4), 1145-1177.

[14] 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", *The Quarterly Journal of Economics*, 126(2), 749-804.

[15] Clotfelter, Charles T. [1983], "Tax Evasion and Tax Rates: An Analysis of Individual Returns", *The Review of Economics and Statistics*, 65(3), 363-373.

[16] Conlisk, John [1996], "Why Bounded Rationality", *Journal of Economic Literature*, 34, 669-700.

[17] Diamond, Peter and Emmanuel Saez [2011], "The Case for a Progressive Tax: From Basic Research to Policy Recommendations", *Journal of Economic Perspectives*, 25(4), 165-190.

[18] Falk, Armin and Anke Becker [2010], "Economic Preference Module", Mimeo, University of Bonn.

[19] Feldman, Naomi E. and Peter Katuščàk [2006], "Should the Average Tax Rate be Marginalized?", Center for Economic Research and Graduate Education Economics Institute Working Paper 304.

[20] Feldstein, Martin [1995], "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act", *Journal of Political Economy*, 103(3), 551-572.

[21] Feldstein, Martin [1999], "Tax Avoidance and the Deadweight Loss of the Income Tax", *Review of Economics and Statistics*, 81(4), 674-680.

[22] Fisman, Raymond and Shang-Jin Wei [2004], "Tax Rates and Evasion: Evidence from "Missing Imports" in China", *Journal of Political Economy*, 112(2), 471-496.

[23] Frederick, Shane [2005], "Cognitive Reflection and Decision Makin", *Journal of Economic Perspectives*, 19, 25-42.

[24] Gruber, Jon and Emmanuel Saez [2002], "The elasticity of taxable income: evidence and implications", *Journal of Public Economics*, 84, 1-32.

[25] Grönqvist, Erik, Björn Öckert and Jonas Vlachos [2010], "The Intergenerational Transmission of Cognitive and Non-Cognitive Abilities", IFAU Working Paper 2010:12.

[26] Hammar, Henrik, Sverker C. Jaegers and Katarina Nordblom [2006], "Skatter och Skattefusk", Mimeo, University of Gothenburg.

[27] Henrekson, Magnus and Ulf Jakobsson [2001], "Ägarpolitik och ägarstruktur i efterkrigstidens Sverige", In L. Jonung (Ed.), *Vem ska äga Sverige?*, Stockholm: SNS Förlag.

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

[29] Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Soren Pedersen and Emmanuel Saez [2011], "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark", *Econometrica*, 79(3), 651-692.

[30] Kleven, Henrik Jacobsen and Mazhar Waseem [2012], "Tax Notches in Pakistan: Tax Evasion, Real Responses, and Income Shifting", Mimeo, London School of Economics.

[31] Kopczuk, Wojciech [2013], "Taxation on Intergenerational Transfers and Wealth", prepared for *Handbook of Public Economics*, 5.

[32] Liebman, Jeffrey B. and Richard J. Zeckhauser [2004], "Schmeduling", Mimeo, Harvard University.

[33] Lindqvist, Erik and Roine Vestman [2011], "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment", *American Economic Journal: Applied Economics*, 3(1), 101-128.

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

[35] Mollerstrom, Johanna and David Seim [2012], "Does the Demand for Redistribution Rise or Fall with Cognitive Ability?", Mimeo, Harvard University.

[36] Nationalekonomiska Föreningen [1948], "De Nya Skatteförslagen", *Förhandlingar 1947*, Stockholm: Hæggströms, 1-31.

[37] Ohlsson, Henry [2011], "The Legacy of the Swedish Gift and Inheritance Tax, 1884-2004", *European Review of Economic History*, 15(3), 539-569.

[38] Piketty, Thomas and Emmanuel Saez [2012], "A Theory of Optimal Capital Taxation", NBER Working Paper 17989.

[39] Reis, Ricardo [2006], "Inattentive Consumers", *Journal of Monetary Economics*, 53, 1761-1800.

[40] Saez, Emmanuel [2010], "Do Taxpayers Bunch at Kink Points?", *American Economic Journal: Economic Policy*, 2, 180-212.

[41] Sims, Christopher A. [2003], "Implications of Rational Inattention", *Journal of Monetary Economics*, 50, 665-690.

[42] Sims, Christopher A. [2006], "Rational Inattention: Beyond the Linear-Quadratic Case", *The American Economic Review*, 96(2), 158-163.

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

[44] Slemrod, Joel and Wojciech Kopczuk [2002], "The Optimal Elasticity of Taxable Income", *Journal of Public Economics*, 84(1), 91-112.

[45] Slemrod, Joel and Shlomo Yitzhaki [2002], "Tax Avoidance, Evasion, and Administration", *Handbook of Public Economics*, 3, 1423-1470.

[46] Swedish Government Official Reports [2004], "Reformerade Egendomsskatter", SOU 2004:36.

[47] Wolff, Edward N. [2010], "Recent Trends in Household Wealth in the United States: Rising Debt and the Middle-Class Squeeze – an Update to 2007", Working Paper no. 589, Levy Economics Institute of Bard College.

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

A   Infinite Horizon Model

This section presents the infinite-horizon version of the two-period model laid out in section 2.3. Utility is additively separable in consumption and has the common *CES* form. Formally, the optimization problem is:

$$U = \max_{A_{t+1}, e_t} \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} \tag{26}$$

subject to

$$c_t + s_t = y_t$$

$$A_{t+1} = s_t + (1 - \tau)(A_t - e_t) + e_t - \left(\frac{e_t}{A_t}\right)^{\frac{1}{\gamma}} \frac{pe_t}{1 + \frac{1}{\gamma}}$$

$$c_t \geq 0$$

$$A_0 \text{ given.}$$

In an interior solution, the problem has first-order conditions

$$e_t = \left(\frac{\tau}{p}\right)^{\gamma} A_t$$

$$c_t^{-\frac{1}{\sigma}} = \beta \left(1 - \tau + \tau \left(\frac{\tau}{p}\right)^{\gamma} \frac{1}{1 + \gamma}\right) c_{t+1}^{-\frac{1}{\sigma}}.$$

These equations describe the dynamics of savings, wealth, evasion and consumption. Starting out with any tax rate and wealth distribution, it is possible to describe the evolution of the response to a tax reform. On a balanced growth rate, all variables grow at a constant rate given by:

$$g = \left(\beta \left(1 - \tau + \tau \left(\frac{\tau}{p}\right)^{\gamma} \frac{1}{1 + \gamma}\right)\right)^{\sigma} \tag{27}$$

An increase in the tax rate has a distortionary effect on the growth rate. However, this effect is attenuated by agents' tendency to shelter money from the government. From a growth-enhancing point of view, sheltering behavior is not too bad.

B Tables

Table 1: Wealth taxpayers and government revenue.

Year	Number of Wealth Tax Payers	Share of Population Paying the Tax	Government Revenue (Million SEK)	Share of Total Government Revenue
1999	707997	0.080	8380	0.012
2000	685929	0.077	7990	0.010
2001	467534	0.053	6270	0.008
2002	258724	0.023	3790	0.005
2003	312705	0.035	4720	0.007
2004	328162	0.036	5140	0.007
2005	223139	0.025	4760	0.006
2006	276373	0.030	5870	0.007

Sources: Wealth tax data from Income- and Tax Register, population data from Statistics Sweden and government revenue data form The Swedish National Financial Management Authority.

Table 2: Summary statistics for the Swedish population and different subsamples, 1999-2006.

	Population	Singles and couples without children			
		All	Excluding self-employed		
			All	Wealth Tax Payers	With Enlistment Data
	(1)	(2)	(3)	(4)	(5)
Demographics					
Age	47.704	53.068	52.699	65.970	37.919
Children	0.570	0	0	0	0
Primary School (%)	0.302	0.332	0.328	0.283	0.152
Secondary School (%)	0.436	0.425	0.426	0.361	0.534
Higher Education (%)	0.261	0.244	0.246	0.356	0.314
Female (%)	0.509	0.500	0.506	0.507	0
Married (%)	0.429	0.362	0.341	0.642	0.135
Income and Taxes					
Wage Earnings	130,747	113,056	113,396	116,792	220,515
Wealth Tax Payed	794	951	798	13,079	351
Third-Party Reported:					
Assets	404,722	418,724	343,164	1,878,774	256,197
Debt	199,058	151,973	135,651	100,727	242,403
Portfolio					
Real Estate (%)	0.312	0.254	0.252	0.374	0.252
Bank Account (%)	0.183	0.219	0.211	0.207	0.183
Funds (%)	0.188	0.203	0.201	0.233	0.182
Stocks (%)	0.056	0.063	0.059	0.143	0.087
Bonds (%)	0.018	0.023	0.022	0.040	0.011
Skills					
Cognitive Skills					-0.069
Observations	58,015,897	37,220,782	33,752,654	2,059,074	4,122,141

Notes: Table entries are means unless otherwise stated. Monetary values are in SEK. The population in column 1 comprises tax payers aged 16 and older. Columns (2)-(5) describe the subsample of single households and couples without children below 18 years of age. Children refers to the number of children below 18 years of age, living with the individual. Entrepreneurs are individuals who possess industrial property. The variable Cognitive Skills denotes the standardized value of the sum of subscores from the cognitive skills test taken at enlistment.

Table 3: Nonparametric estimates of bunching, $\hat{b}$, 1999-2006.

	(1)	(2)	(3)	(4)
All	0.489***	0.950***	1.447***	2.837***
	(0.009)	(0.014)	(0.019)	(0.033)
Couples	0.533***	1.036***	1.443***	2.864***
(0.013)	(0.021)	(0.028)	(0.046)	
Singles	0.453***	0.880***	1.447***	2.805***
	(0.012)	(0.017)	(0.025)	(0.044)
Men with military	0.740***	1.455***	1.949***	3.794***
enlistment data	(0.046)	(0.120)	(0.097)	(0.176)
High-skilled	0.894***	1.710***	2.257***	4.359***
	(0.073)	(0.120)	(0.150)	(0.248)
Low-skilled	0.542***	1.127***	1.444***	2.996***
	(0.056)	(0.096)	(0.115)	(0.214)
δ	40,000	100,000	40,000	100,000
Extrapolation	No	No	Yes	Yes

Notes: The number of households that bunch are computed in accordance with equation (24) and the counterfactual density at the kink is obtained using third-party reported net wealth. δ denotes the width of the bunching interval. Extrapolation refers to the method described in Section 3.2. High-skilled and low-skilled households are defined as having, respectively, positive and negative z-scores in cognitive ability. Standard errors are computed using a nonparametric bootstrap procedure, in which new distributions of paired third-party reported net wealth and taxable net wealth are drawn with replacement from the true distribution. The standard error of each estimate is the standard deviation of the distribution of the $\hat{b}$':s. Significance codes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Estimation results from regressing sentenced fines on downward self-adjustments using, OLS, 1999-2006. Dependent variable: penalties.

	(1)	(2)	(3)	(4)
Self-reported wealth	-0.010 (0.007)	-0.009 (0.007)	-0.002 (0.018)	0.011 (0.016)
I (Downward-adjustment) × Self-reported wealth	0.041** (0.019)	0.039** (0.019)	0.168*** (0.055)	0.139*** (0.051)
Wage Earnings		0.070 (1.790)		2.98 (4.700)
Third-party Net Wealth		0.001 (0.001)		0.007 (0.005)
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Lower bound on penalties	0	0	10,000	10,000
Observations	7,556	7,556	2,247	2,247

Notes: The variable self-reported wealth denotes self-reported wealth adjustment of taxable net wealth. I (Downward-adjustment) is an indicator of self-reported downward adjustment. All regressions include the following controls: I (Downward-adjustment), self-reported wealth squared and the squared interaction between self-reported adjustment and I (Downward-adjustment). Standard errors clustered at the household level. Significance codes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Tax semi-elasticities of taxable net wealth, $\varepsilon_{W,\tau}$, 1999-2006.

| Method | Sample | $\hat{b}$ | Small Kink Analysis | Static Model | Dynamic Model $\sigma = 0.25$ | Dynamic Model $\sigma = 1.1$ | IV |
			(1)	(2)	(3)	(4)	(5)
Parametric	All	0.53	-0.118	-0.18	-0.203	-0.134	
		1.25	-0.278	-0.377	−0.354	-0.339	
	High	1.1	-0.244				
		0.97	-0.216				
	Low	0.42	-0.093				
		0.21	-0.047				
Nonparametric	All	0.49	−0.109	-0.172	-0.195	-0.121	
		0.95	-0.211	-0.300	-0.293	-0.259	
		1.45	−0.320	-0.420	-0.395	-0.389	
	High	0.89	-0.199				
		1.71	-0.380				
		2.26	-0.502				
	Low	0.54	-0.120				
		1.13	-0.250				
		1.44	-0.320				
IV							
$x = 1$							-0.799***
							(0.070)
$x = 2$							-0.539***
							(0.079)
$x = 3$							-0.452***
							(0.101)

Notes: The bunching estimates, $\hat{b}$, are obtained using either the parametric method described in Section 4.1 or the non-parametric method described in Section 4.2. The estimates are estimated on the entire sample (All) or the subsamples comprising individuals with high cognitive ability (High) and low cognitive ability (Low). The two entries for each subsample using the parametric approach denote the benchmark estimate and the estimate obtained using the extrapolation method, respectively. The three entries for each subsample using the non-parametric approach denote the benchmark estimate, the results for a broader bunching interval and the estimate obtained using the extrapolation method, respectively. Columns (1)-(4) display the estimate implied by the indicated method/model. Where applicable, $\beta = 0.95$ and $p = 0.86$. The IV-estimates in column (5) display the results from estimating equation (25) by 2SLS and x denotes lag length. In these estimations, the standard errors, displayed within parenthesis, are clustered at the household level. Significance codes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Tax semi-elasticities and elasticities of evasion, $\varepsilon_{e,\tau}$, 1999-2006.

		Static Model		Dynamic Model			
				Uncompensated		Compensated	
		Semi-elasticity	Elasticity	Semi-elasticity	Elasticity	Semi-elasticity	Elasticity
Method	$\hat{b}$	(1)	(2)	(3)	(4)	(5)	(6)
Parametric	0.53	104.370	1.566	81.269	1.219	81.765	1.227
	1.25	90.241	1.354	76.659	1.150	77.155	1.157
Non-parametric	0.49	105.66	1.585	81.567	1.224	82.064	1.231
	0.95	94.760	1.421	78.424	1.176	78.920	1.184
	1.45	87.797	1.317	75.582	1.134	76.077	1.141

Notes: The bunching estimates, $\hat{b}$, are obtained using either the parametric method described in Section 4.1 or the non-parametric method described in Section 4.2. The two entries using the parametric approach denote the benchmark estimate and the estimate obtained using the extrapolation method, respectively. The three entries using the non-parametric approach denote the estimate obtained when using a window of SEK $40,000$ on the 1999-2006 sample, the estimate obtained when using a window of SEK $100,000$ on the 1999-2006 sample and the estimate obtained when using the extrapolation method, respectively. Columns (1)-(6) display the estimate implied by the indicated model. Where applicable, $\beta = 0.96$, $\sigma = 0.25$ and $p = 0.86$.

Table 7: Estimation results from regressing downward adjustments of taxable net wealth on third-party reported net wealth using OLS, 1999-2006. Dependent variable: Downward adjustment.

	(1)	(2)
Third-party Net Wealth	0.837***	0.522***
	(0.051)	(0.109)
Third-party Net Wealth Squared	3.47×10^{-12}**	2.60×10^{-11}***
	(1.49×10^{-12})	(6.84×10^{-11})
Third-party Net Wealth Cubed		-3.65×10^{-22}***
		(1.01×10^{-22})
Year FE	Yes	Yes
Observations	4,835,932	4,835,932

Notes: Downward adjustment is defined as downward-adjusted self-reported taxable net wealth. The standard errors, displayed within parenthesis, are clustered at the household level. Significance codes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Estimation results from regressing bunching and downward adjustments of taxable net wealth on cognitive skills, 1999-2006.

Dependent variable:	Bunching				Log Downward Adjustment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Skills	0.007***	0.006***	0.014***	0.011***	0.149***	0.123***	0.225***	0.167***
	(0.002)	(0.003)	(0.003)	(0.004)	(0.032)	(0.037)	(0.039)	(0.045)
Education controls	No	Yes	No	Yes	No	Yes	No	Yes
Top percentile excluded	No	No	No	No	Yes	Yes	Yes	Yes
Sample period	99-06	99-06	02-06	02-06	99-06	99-06	02-06	02-06
Observations	36,072	36,072	18,376	18,376	70,347	70,347	38,996	38,996

Notes: The bunching variable assumes the value 1 if the household is bunching, zero otherwise. Columns (1)-(4) thus represents the results from estimating a linear probability model on the subsample of households within SEK $500,000$ above the kink in third-party reported net wealth. Downward adjustment is defined as downward-adjusted self-reported wealth. Both dependent variables are computed using the extrapolation method. All regressions include controls for wage earnings, a 10-piece linear spline in third-party reported net wealth and type of wealth held in the portfolio. The education controls comprise four dummies representing the highest education obtained: primary school, secondary school, tertiary education and PhD degree. Type of wealth refers to the share of real estate and financial property in the portfolio. Top percentile excluded refers to the exclusion of the 99th percentile of the wealth distribution. Significance codes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Estimation results from regressing sentenced fines, risk aversion and impatience on cognitive skills.

Dependent var:	Sentenced Fines		Risk Aversion			Impatience		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cogn. Skills	0.112	0.110	0.120	0.124	0.149	-0.295	-0.291	-0.255
	(0.079)	(0.080)	(0.184)	(0.178)	(0.184)	(0.208)	(0.202)	(0.204)
Log Income	-0.1647***	-0.1653***	-0.122	-0.109	-0.081	-0.207	-0.192	-0.151
	(0.0243)	(0.0244)	(0.078)	(0.079)	(0.086)	(0.211)	(0.201)	(0.222)
Age	0.004	0.004		0.036			0.048**	
	(0.009)	(0.009)		(0.024)			(0.024)	
Divorced					0.653			0.915*
					(0.487)			(0.553)
Year FE	yes	yes	no	no	no	no	no	no
Spline in net wealth	yes	yes	no	no	no	no	no	no
Observations	576	576	153	153	153	153	153	153

Notes: The models in columns (1) and (2) are based on administrative data on sentenced fines on tax evasion and include controls for wealth. The variables risk aversion, impatience and the dummy variable divorced are based on self-reported survey data from Mollerstrom and Seim (2012). Significance codes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C   Figures

Figure 1: Marginal tax rates over time.
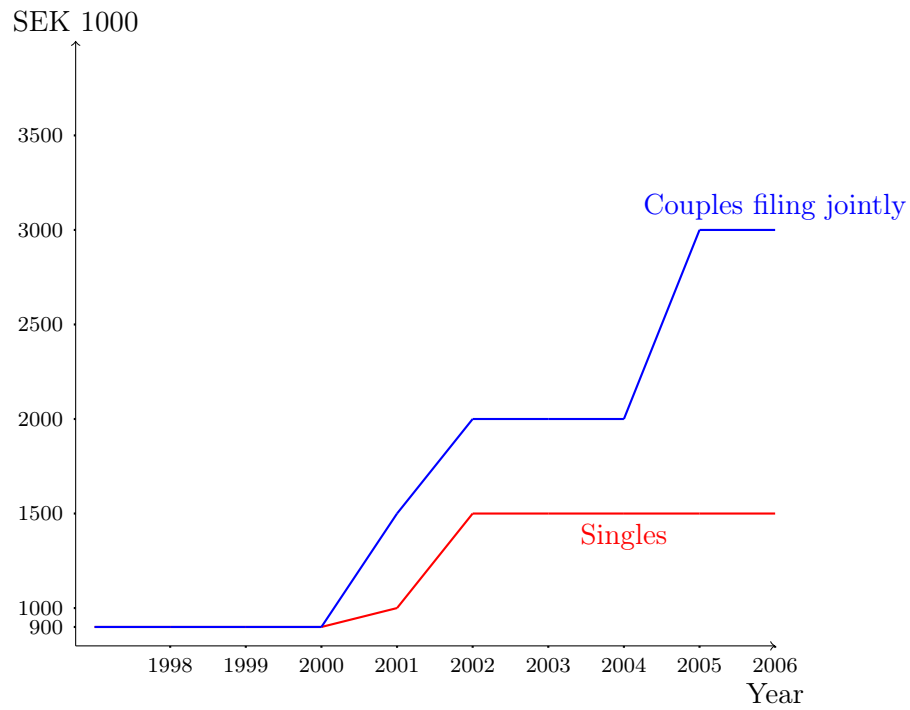
Notes: The figure shows the evolution of the marginal tax rate bracket over time, for single households and couples without children who file the tax jointly.

Figure 2: Prepopulated tax return.

Skattemyndigheten

Post-adress
Besöks-adress
Skatteupp-lysningen

SKATTEKONTORET I WISTAD
BOX 123
123 45 WISTAD
STORGATAN 35

0771-567 567

Inkomstdeklaration 1   2003
Inkomståret 2002

Person-/Organisationsnummer
580301-2805   PROVEXEMPLAR

Senast måndagen den 5 maj 2003
ska deklarationen finnas hos skattemyndigheten
för att du ska undgå förseningsavgift.

Läs först i "Dags att deklarera" hur du ska fylla i blanketten.

Andersson, Margareta
BADHUSPLANEN 2
123 48 WISTAD

Kontrollera belopp som skattemyndigheten fyllt i. Om en uppgift är fel, stryk över den och fyll i rätt totalbelopp i vit ruta. **Ändra** eller **lägg till** uppgifter endast i vit ruta.

Behåll den ena som kopia!

① Inkomster - Tjänst

Lön, förmåner, sjukpenning m.m.	230 780	03
Kostnadsersättningar		05
Allmän pension och tjänstepension		14
Privat pension och livränta		15
Andra inkomster som inte är pensionsgrundande		16
Inkomster, t.ex. hobby, som du själv ska betala egenavgifter för		19
Inkomst enligt blankett K10, K4C eller N3A		22
K10 har lämnats via ADB. Detta gäller delägare m.fl. i fåmansföretag.		79

② Avdrag - Tjänst

Resor till och från arbetet
Du får endast avdrag för den del som **överstiger 7 000 kr. Fyll i totalbeloppet.** (Skattemyndigheten drar automatiskt av 7 000 kr.) | 01

Tjänsteresor | 08

Tillfälligt arbete, dubbel bosättning och hemresor | 09

Övriga utgifter
Du får endast avdrag för den del som **överstiger 1 000 kr. Fyll i totalbeloppet.** (Skattemyndigheten drar automatiskt av 1 000 kr.) | 06

③ Allmänna avdrag

Pensionssparande m.m. | 6 000 | 43

④ Kapital

Ränteinkomster, utdelningar m.m.	7 515	50
Överskott vid uthyrning av privatbostad enligt blankett K3		51
Avdrag för ränteutgifter m.m.	7 232	53
Avdrag för förvaltningsutgifter Du får endast avdrag för den del som **överstiger 1 000 kr. Fyll i totalbeloppet.** (Skattemyndigheten drar automatiskt av 1 000 kr.)		61
Kapitalvinst	1 160	56
Kapitalförlust		57

⑤ Förmögenhet

Här fyller du i din förmögenhet | Sambeskattads personnummer
• om du **ska sambeskattas** och om din, din makes och dina hemmavarande barns sammanlagda förmögenhet överstiger **2 000 000 kr.**
• om du **inte ska sambeskattas** och om din och dina hemmavarande barns sammanlagda förmögenhet överstiger **1 500 000 kr.**
Läs i "Dags att deklarera" på sidan 10 om hur barns förmögenhet redovisas.

Tillgångar	1 798 296	66
Skulder	115 200	67

⑥ Utländsk försäkring - Avkastningsskatt

Skatteunderlag för kapitalförsäkring		62
Skatteunderlag för pensionsförsäkring		63

⑦ Underlag för fastighetsskatt

Småhus/	1,0 %	0*	80
Tomt för småhus	0,5 %		82
Privatbostad i utlandet	1,0 %		84

* Lägg till underlaget för fastighet som du ägt under endast en del av året.

Observera! Innan du fyller i ruta 56 eller 57, läs i Dags att deklarera på sidan 9. Annars är det risk att du får fel skatt.

Skriv under på baksidan!

580301-28051L030121

NRV	FÅM		Region	Kontor	Sektion	Grupp	Person-/Organisationsnummer	Fåmansföretag		Tax.år
			01	21	P01	99	580301-2805			2003

Webbplats www.rsv.se

Allmän självdeklaration enligt lagen SFS 2001:1227
RSV 2000 K utgåva 12 02-10   Nærljarydenics AB

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 3: Formula for computation of tax liability.

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

Figure 4: Third-party reported net wealth distribution around the threshold.
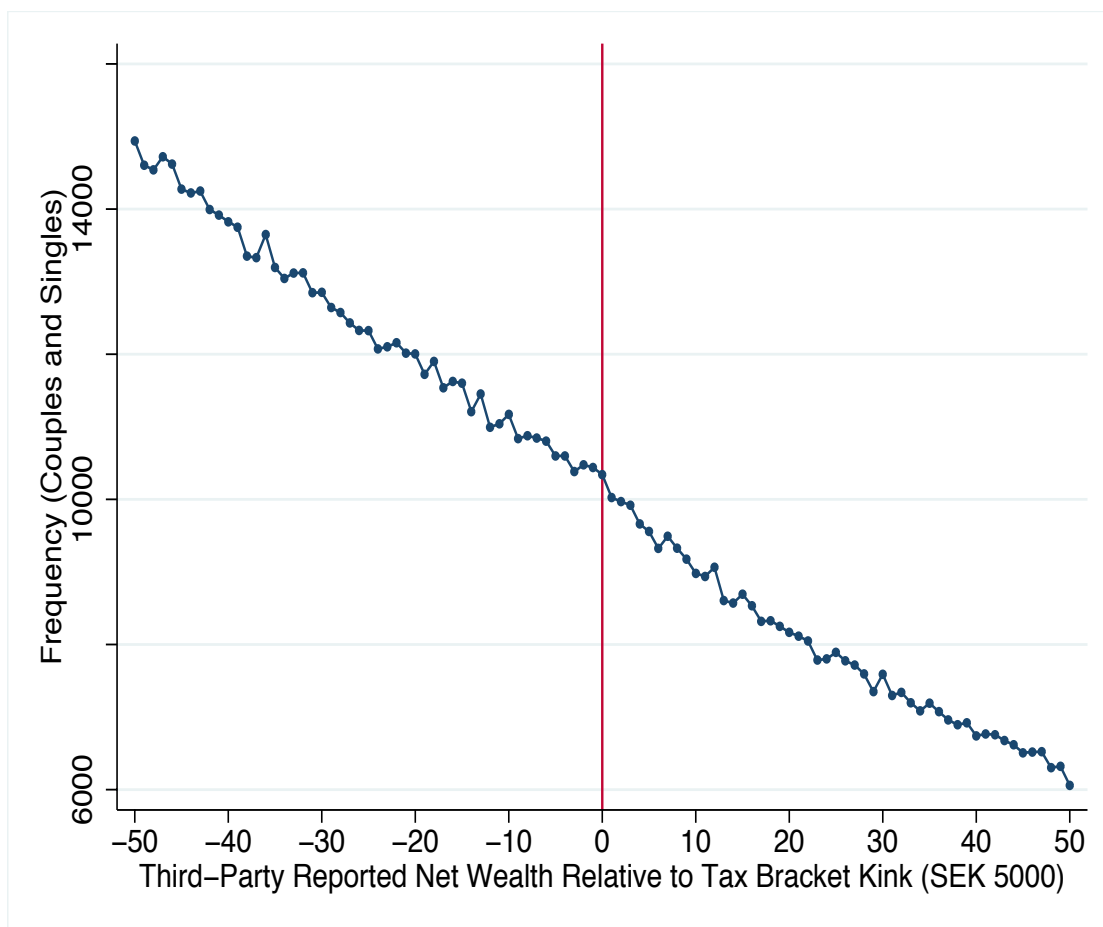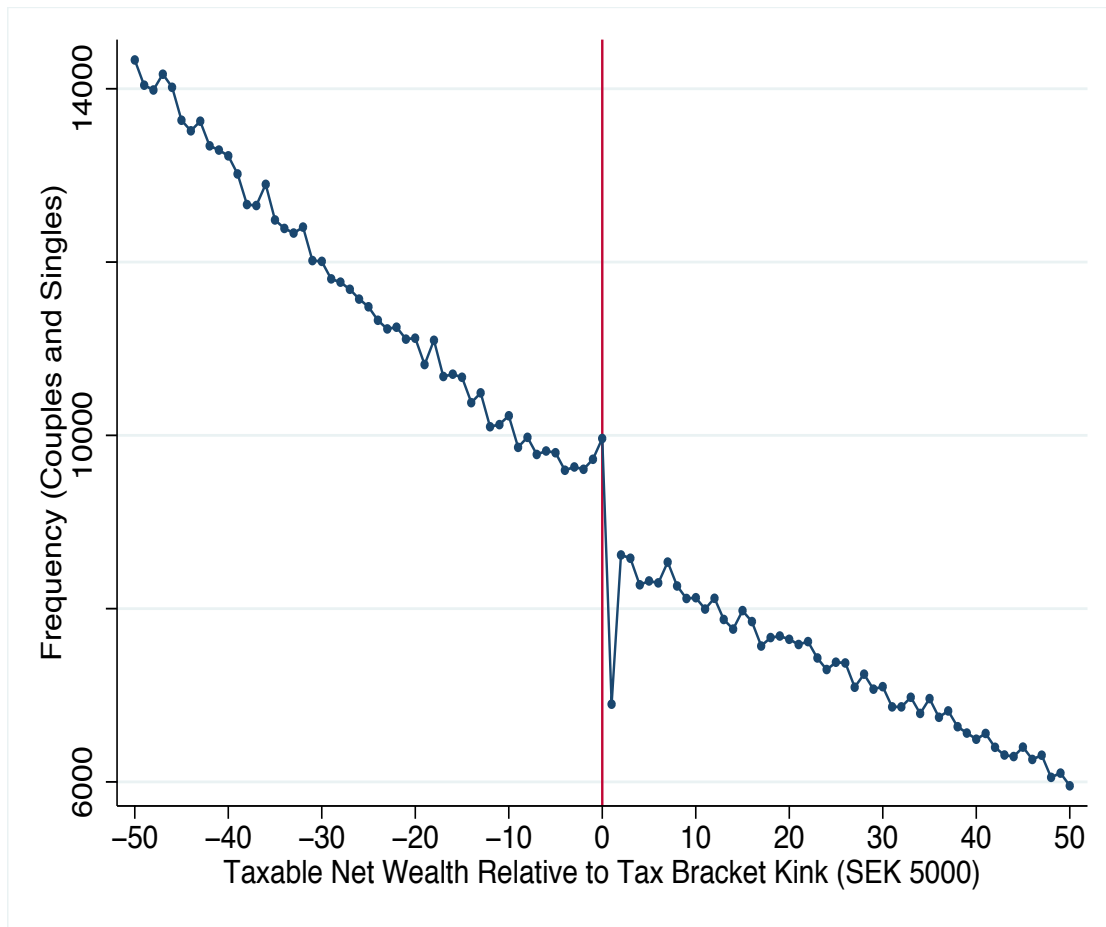
(a)

Notes: The figure shows the distribution of third-party reported net wealth around the shift in the tax brackets, demarcated by the vertical at 0, for the years 1999-2006. 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.

Figure 5: Taxable net wealth distribution around the threshold.

(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 the years 1999-2006. 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.

Figure 6: Estimated bunching of taxable net wealth at the threshold.

(a)
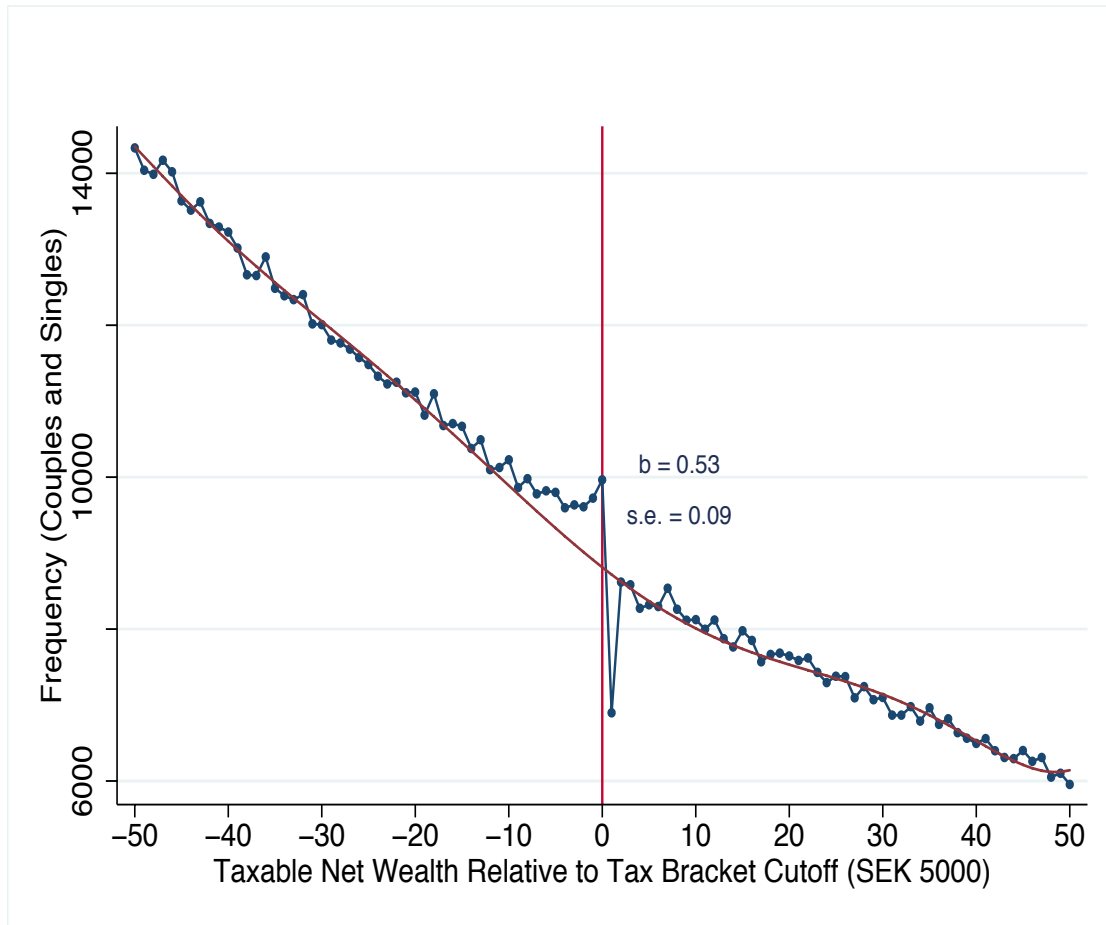
Notes: The figure replicates Figure 5 and adds the estimated counterfactual density, displayed by the solid line in red. The counterfactual density was obtained by fitting a seven-degree polynomial to the density, excluding points within SEK 40,000 below the kink. *b* denotes the estimated excess mass.

Figure 7: Estimated bunching of taxable net wealth at the threshold, extrapolation method.

(a)
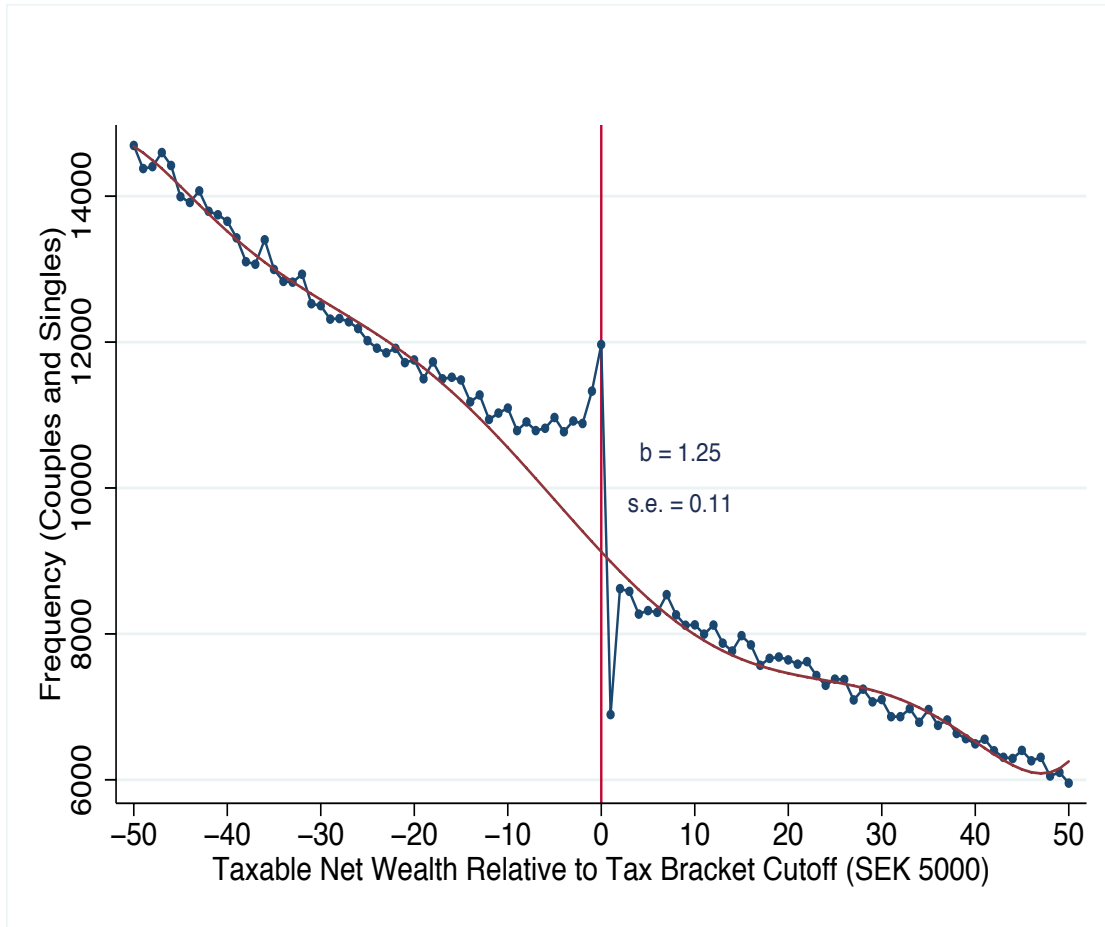
Notes: This figure replicates Figure 6 but uses the extrapolation method – which is described in the text – to deal with missing taxable net wealth values.

Figure 8: Estimated bunching of third-party reported net wealth at the threshold.

(a)
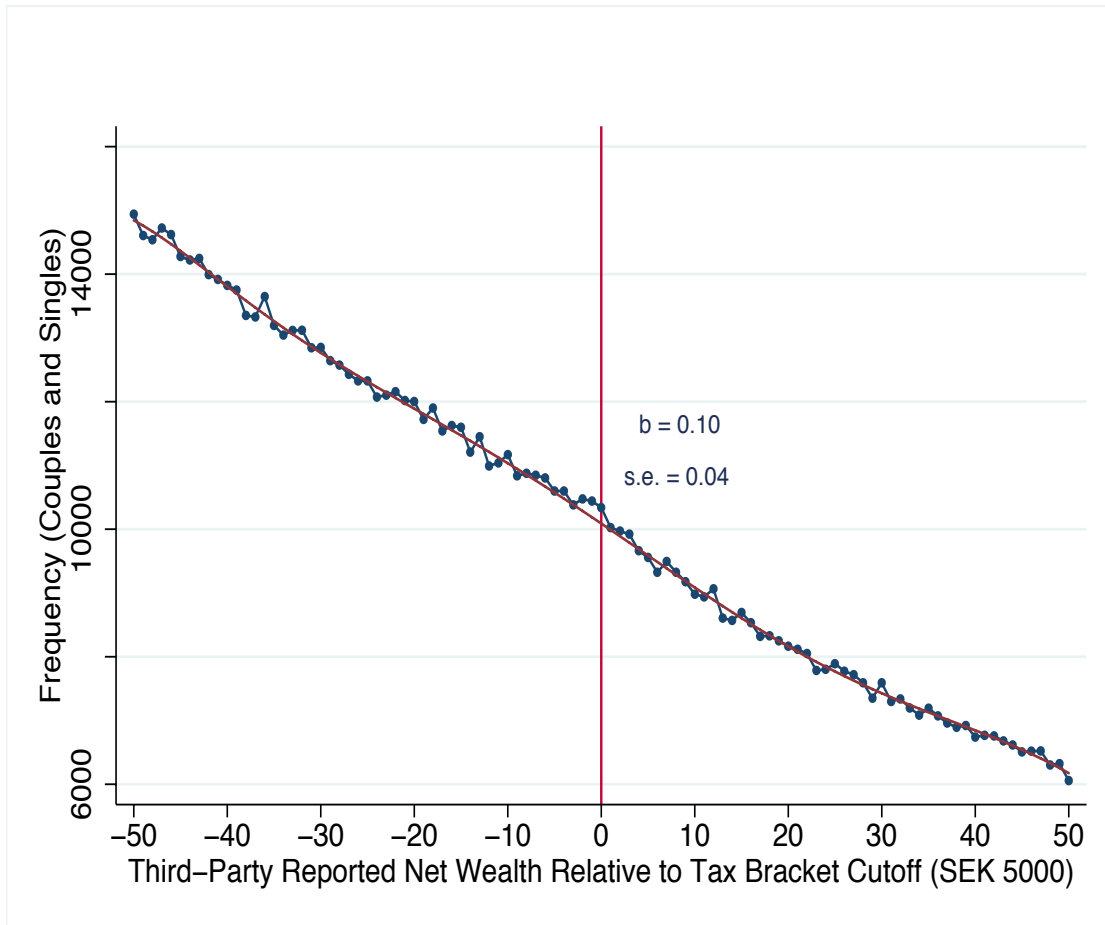
Notes: The figure replicates Figure 4 and adds the estimated counterfactual density, displayed by the solid line in red. The counterfactual density was obtained by fitting a seven-degree polynomial to the density, excluding points within SEK 40,000 below the kink. *b* denotes the estimated excess mass.

Figure 9: Taxable net wealth around the threshold 1999-2006, singles.
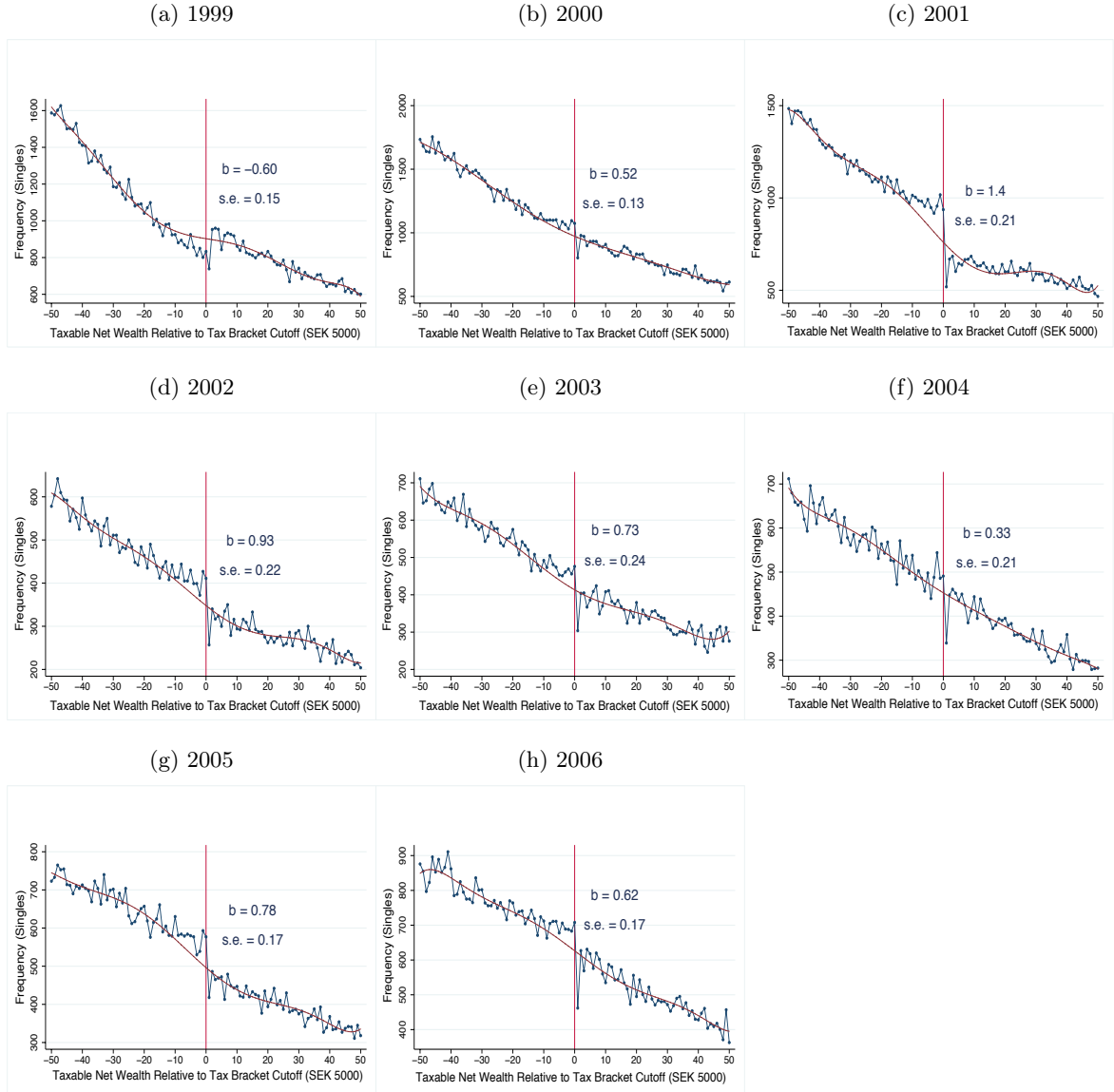
(a) 1999

(b) 2000

(c) 2001

(d) 2002

(e) 2003

(f) 2004

(g) 2005

(h) 2006

Notes: These figures plot the empirical distribution of taxable net wealth for single households around the kink point in each year from 1999-2006. The vertical line denotes the location of the threshold. The counterfactual density, graphed by the solid curve in red, was obtained as in Figure 6.

51

Figure 10: Taxable net wealth around the threshold 1999-2006, couples.

(a) 1999

(b) 2000

(c) 2001

(d) 2002

(e) 2003

(f) 2004

(g) 2005

(h) 2006
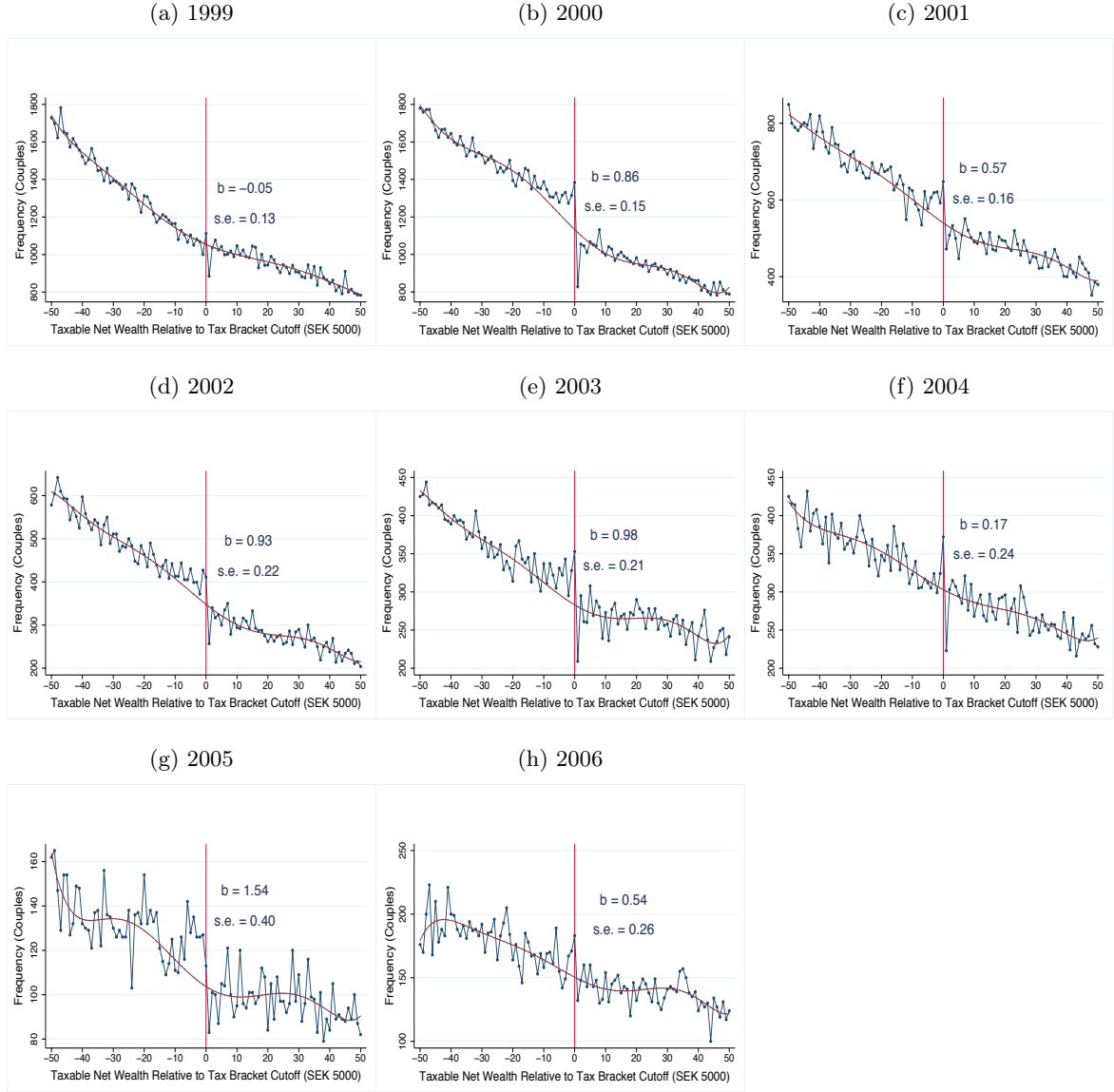
Notes: These figures plot the empirical distribution of taxable net wealth for couples without children, who file jointly, around the kink point in each year from 1999-2006. The vertical line denotes the location of the threshold. The counterfactual density, graphed by the solid curve in red, was obtained as in Figure 6.

Figure 11: Does bunching track the tax? Bunching in 2001 and 2006.

(a) 2001

(b) 2006
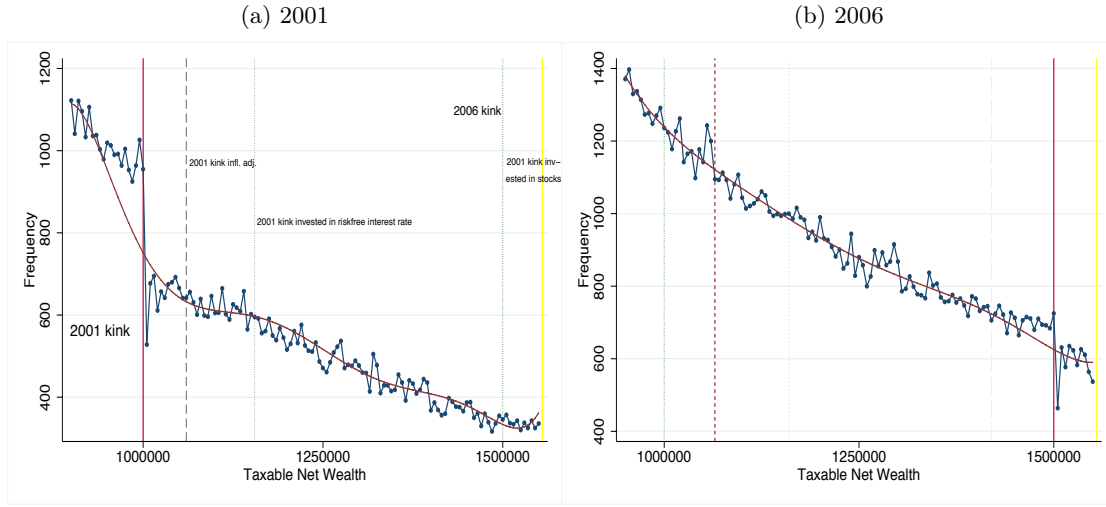
Notes: These figures present the taxable net wealth distribution for singles in 2001 and 2006. The figure shows the kinks in 2001 and 2006, located at SEK 1 million and SEK 1.5 million, respectively. The additional vertical lines represent the position of the 2001-kink if it followed - from the left 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 12: Taxable net wealth distributions, by cognitive skills.

(a) High Skilled

(b) Low Skilled
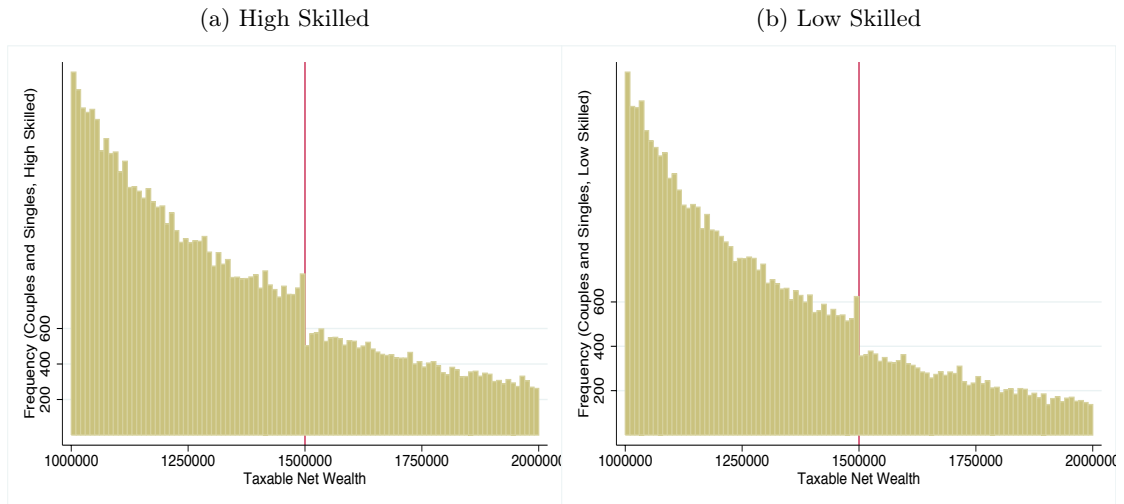
Notes: These figures show the distribution of taxable net wealth around the kink point, demarcated by the vertical at SEK 1,500,000, for the years 1999-2006. Wealth distributions are recentered each year by family status to obtain the recentered kink. Each bin corresponds to the number of households within SEK 10,000. High skilled is defined as having a positive z-score. For couples, the skill variable measures the z-score of the male.

Figure 13: Taxable net wealth distributions, by cognitive skills, 2002-2006.

(a) High Skilled

(b) Low Skilled
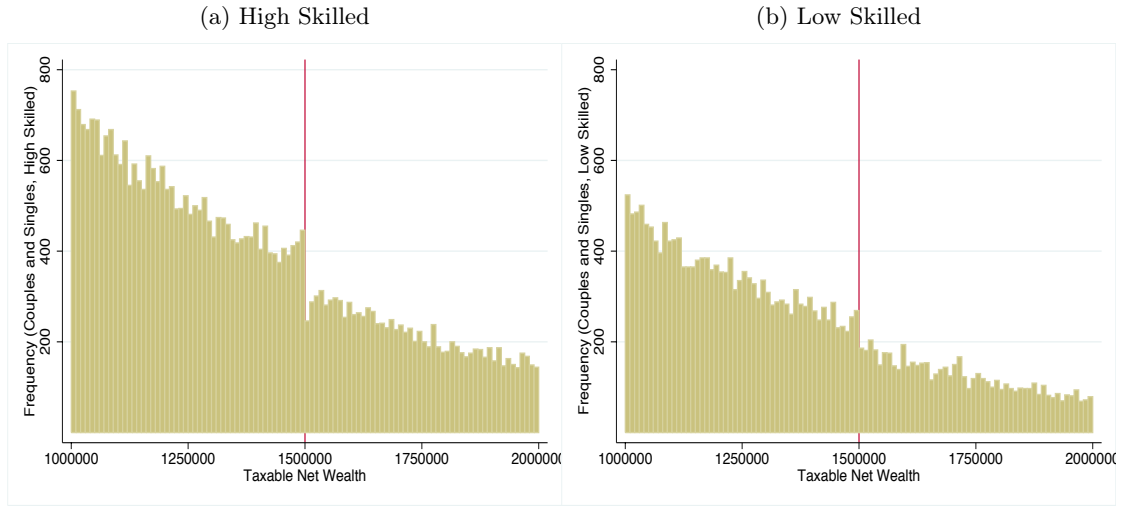
Notes: These figures replicate Figure 12 for the years 2002-2006.

Figure 14: Estimated bunching of taxable net wealth at the threshold, by skill groups.

(a) High Skilled

(b) Low Skilled

Notes: These figures show the distribution of taxable net wealth close to the kink point in the years 2002-2006, by skill groups. They add the estimated counterfactual density, displayed by the solid lines in red. The counterfactual densities were obtained by fitting a seven-degree polynomial to the densities, excluding points within SEK 40,000 below the kink. b denotes the estimated excess mass. High skilled are defined as having a positive z-score. For couples, the skill variable measures the z-score of the male.

Figure 15: Estimated bunching of taxable net wealth at the threshold, by skill groups, alternative measure.

(a) High Skilled   (b) Low Skilled

Notes: These figures replicate the plots in Figure 14 but change the skill group definition. High skilled is defined as having a score above the 75th percentile in the skill distribution. Low skilled below 25th. For couples, the skill variable measures the z-score of the male.

Figure 16: Bunching over time - couples.

(a) 1999, 2000 kink

(b) 2001 kink

(c) 2002,2003,2004 kink

(d) 2005,2006 kink
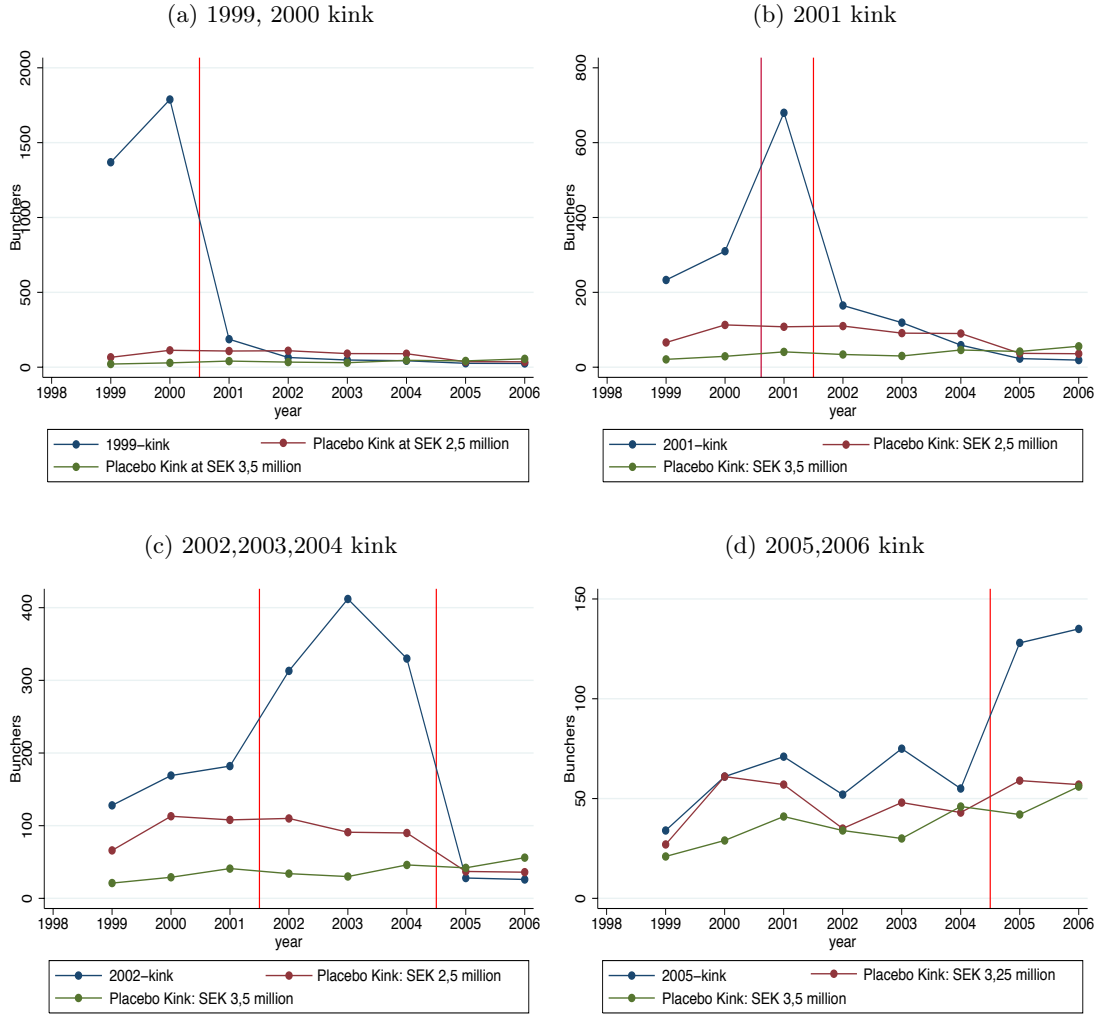
Notes: These graphs show bunching – computed by the nonparametric method – over time at different kinks for couples. In (a), bunching at the threshold of SEK $900,000$, which was the threshold in 1999 and 2000, with a window of SEK $25,000$ is displayed. (b), (c) and (d) present similar graphs for bunching at thresholds corresponding to the years 2001, 2002-2004 and 2005-2006, respectively. All graphs additionally plot bunching at placebo kinks of SEK 2.5 million and SEK 3.5 million. All bunching estimates are computed using the extrapolation method.

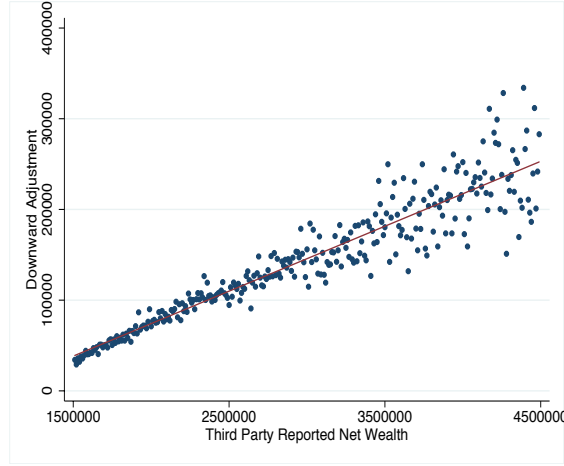Figure 17: Third-party reported net wealth around threshold 1999-2006, singles.
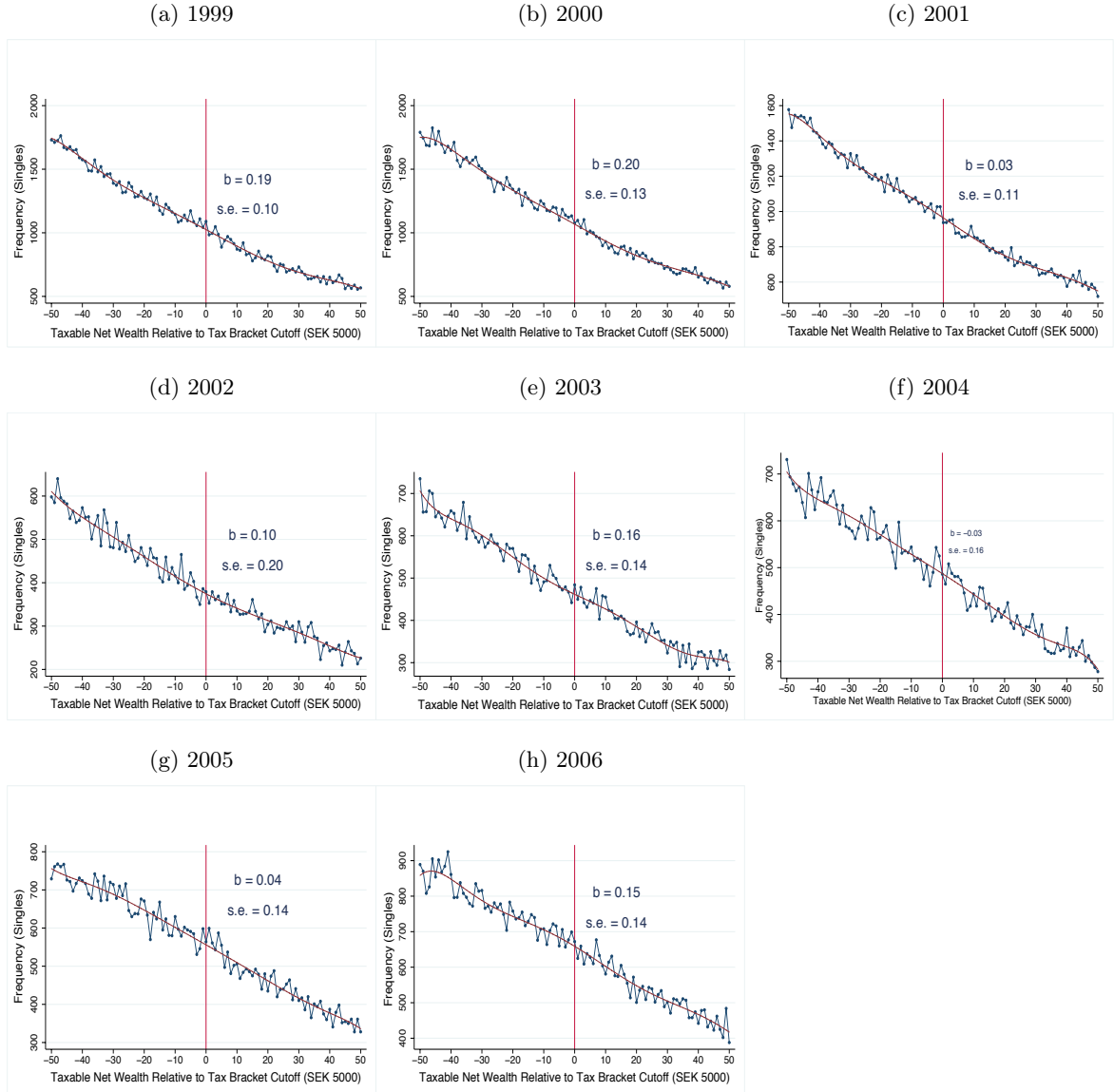
(a) 1999

(b) 2000

(c) 2001

(d) 2002

(e) 2003

(f) 2004

(g) 2005

(h) 2006

Notes: These figures plot the empirical distribution of third-party reported net wealth for single households around the kink point in each year from 1999-2006. The vertical line denotes the location of the threshold. The counterfactual density, graphed by the solid curve in red, was obtained as in Figure 6.

Figure 18: Third-party reported net wealth around the threshold 1999-2006, couples.
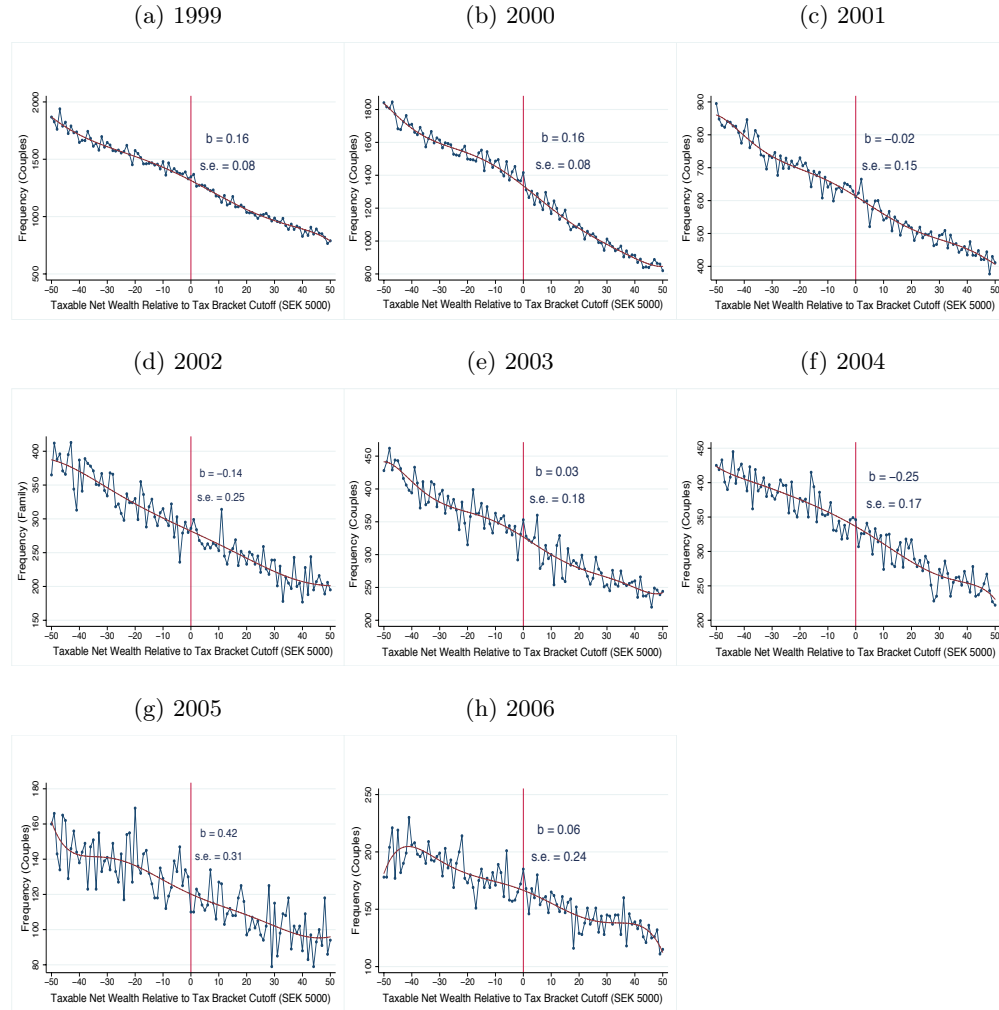
(a) 1999　　　　　　　　(b) 2000　　　　　　　　(c) 2001

(d) 2002　　　　　　　　(e) 2003　　　　　　　　(f) 2004

(g) 2005　　　　　　　　(h) 2006

Notes: These figures plot the empirical distribution of third-party reported net wealth for couples without children who file the tax jointly, around the kink point in each year from 1999-2006. The vertical line denotes the location of the threshold. The counterfactual density, graphed by the solid curve in red, was obtained as in Figure 6.

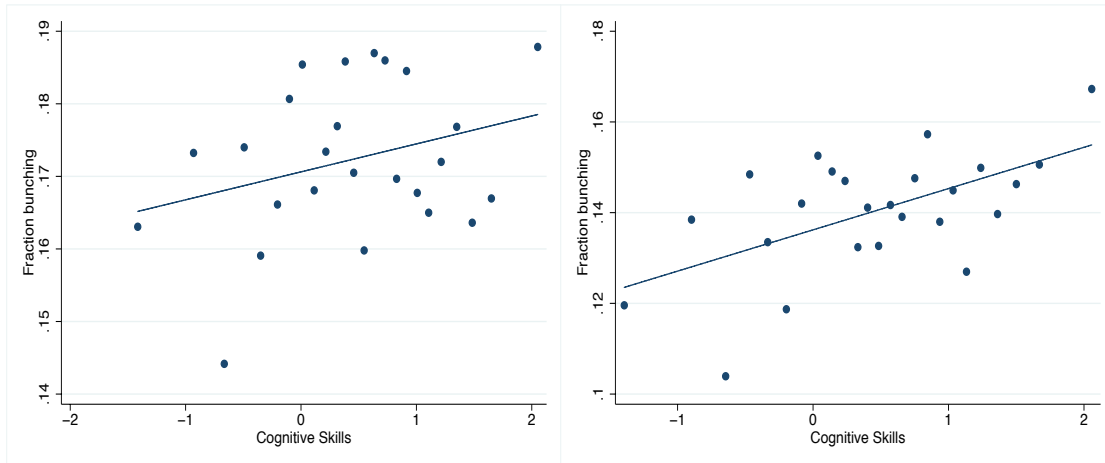Figure 19: Downward adjustment and third-party reported net wealth.

Notes: The figure shows mean of self-reported downward adjustments of taxable net wealth against third-party reported wealth. Third-party reported wealth is binned in SEK 10, 000 intervals.

Figure 20: Bunching and cognitive skills by quantile.

(a) 1999-2006

(b) 2002-2006

Notes: The scatter plots present residuals, after regressing an indicator variable of bunching on a 10-piece spline in net wealth, wage earnings, education, share real estate and share financial property of the portfolio, on the y-axis, against cognitive skills on the x-axis. The sample consists of households above the kink but within SEK 500, 000 above the kink. Each quantile comprises 4 percent of the sample.

Figure 21: Log downward adjustment and cognitive skills by quantile.

(a) 1999-2006   (b) 2002-2006
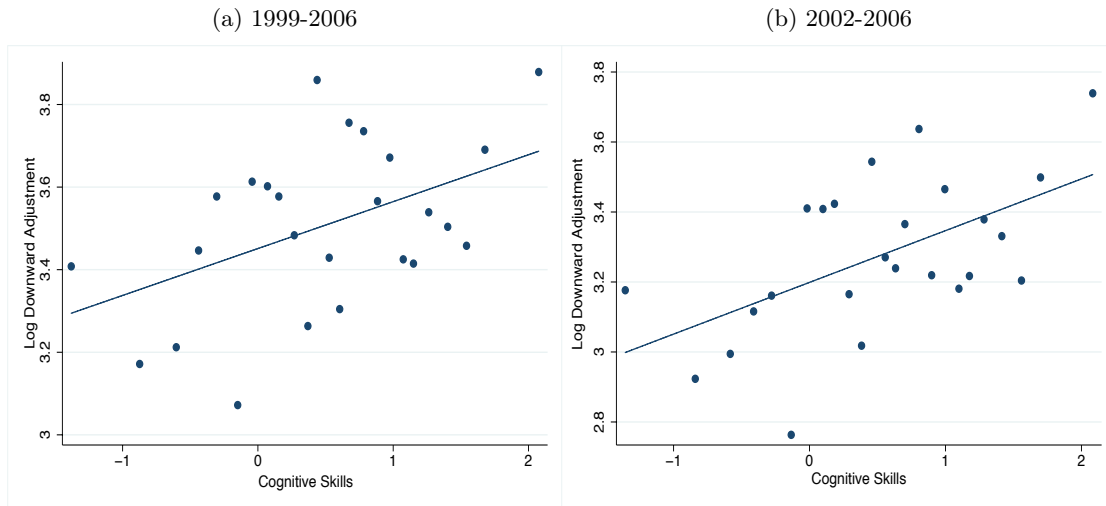
Notes: The scatter plots present residuals, after regressing self-reported downward adjustments of taxable net wealth on a 10-piece spline in net wealth, wage earnings, education, share real estate and share financial property of the portfolio, on the y-axis, against cognitive skills on the x-axis. The sample consists of households above the kink but within SEK 500, 000 above the kink. Each quantile comprises 4 percent of the sample.