Loss evasion and tax aversion*

Per Engström  Katarina Nordblom  Henry Ohlsson  Annika Persson

10 October 2011

Abstract

The objective of this paper is to study if Swedish tax payers behave in a loss averse manner when not complying to taxation. This is important for tax design but also for understanding human behavior in general. The predictions of expected utility theory can be contrasted to those of prospect theory. We use data for 3.6 million tax payers for the income year 2006. Our research method is to use a regression kink approach and a regression discontinuity approach. We find strong evidence for loss aversion.

Keywords: loss aversion, prospect theory, tax evasion, tax avoidance, regression kink, regression discontinuity

JEL Codes: D03, D81, H24, H26

Preliminary, incomplete, not for circulation

Per Engström, Henry Ohlsson: Department of Economics, Uppsala University and Uppsala Center for Fiscal Studies (UCFS), Department of Economics, Uppsala University
Katarina Nordblom: Uppsala Center for Fiscal Studies (UCFS), Department of Economics, Uppsala University and Department of Economics, University of Gothenburg
Annika Persson: The Swedish Tax Agency

Corresponding author:
Henry Ohlsson, Department of Economics, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden, email: henry.ohlsson@nek.uu.se, phone: +46 18 471 11 04, fax: +46 18 471 14 78

*We would like to thank seminar participants at a UCFS workshop, the Swedish Tax Agency, and University of Gothenburg and conference participants at the 2011 Swedish Economics Meeting for their valuable comments and suggestions. Financial support to UCFS from Riksbankens Jubileumsfond (RJ), the Swedish Council for Working Life and Social Research (FAS), and the Swedish Tax Agency is gratefully acknowledged. Some of the work was done when Ohlsson enjoyed the hospitality of School of Economics, UNSW, Sydney during his sabbatical. Financial support from the Wenner-Gren Foundations for the sabbatical is gratefully acknowledged.
1 Introduction

Behavioral economics is a field that has received a rapidly increasing interest during the last decades. The theoretical advances are many. The concept of loss aversion has been recognized in the literature ever since Kahneman and Tversky (1979). Most empirical studies of loss aversion have used lab experiments and have found that people seem to be loss averse.\(^1\)

There is, however, much less real world evidence of loss aversion although the theory seems intuitive. Pope and Schweitzer (2011) is the only study using real world data that we are aware of. They study the putting behavior of professional golf players on the GPA tour. This is an important contribution as it uses actual data with 2.5 million observations, experienced players, and high stakes. There is clear evidence of loss aversion.

However, the sample is quite special. It is an open question if it is possible to generalize the results. In this paper, however, we study the whole population of Swedish tax payers. We study their behavior when filing their income tax returns. Very few become professional golfers but almost all of us pay taxes. This is the first study that we know of that documents loss aversion among a large general sample of the population when doing a very common real world duty.

Preliminary tax payments should have no effect on tax compliance according to standard neoclassical theory. This is still true if we incorporate psychological costs into the analysis, and not only the pure pecuniary aspects as in Allingham and Sandmo (1972) and Yitzhaki (1974). The final net-of-tax income is what matters for compliance irrespective of whether one has paid too much or too little in preliminary taxes.

There are, however, several indications that people who have paid too little in preliminary taxes are less likely to comply than those who have paid too much. Chang and Schultz Jr (1990) find for the US that those who owe additional taxes when they file their returns are less compliant than those who have refunds due. Persson (2003) is a previous descriptive study using Swedish data that reports similar results; she finds that people with a tax deficit are more prone to claim dubious deductions. Experimental studies also suggest that advance tax payments actually matter for compliance.\(^2\)

These findings are not consistent with expected utility theory. Reference dependence and loss aversion, as presented in *Prospect theory* by Kahneman and Tversky (1979); Tversky and Kahneman (1992), predict such outcomes on the other hand. Reference dependence means that the individual attaches a value to the deviation of an outcome from a reference point. It is likely that a zero preliminary tax balance is such a reference point in our particular application.

Some theoretical studies in the area of prospect theory focusing on tax compliance put forward the following line of reasoning:\(^3\) Loss aversion implies that the

\(^1\) (References to be added.)

\(^2\) See, e.g., Robben et al. (1990), Schepanski and Shearer (1995), and Copeland and Cuccia (2002).

\(^3\) See, e.g., Yaniv (1999), Bernasconi and Zanardi (2004), and Dhani and al Nowaihi (2007).
individual values losses in comparison to the reference point more than gains of the same amount. An individual with a preliminary tax deficit (more taxes due) will, therefore, perceive a higher marginal value of extra income than an individual with a preliminary tax surplus (some taxes will be refunded) of the same amount. Those with a preliminary tax deficit would consequently be less inclined to comply.

Most empirical studies testing these hypotheses are experimental. A reason for this is that it is difficult to get reliable real world data on tax compliance, tax evasion, and tax avoidance.

We present a theoretically founded empirical analysis using actual tax return data of very high quality. This is a major contribution of the paper. We study the probability that tax payers claim deductions for “other expenses for earning employment income”. These deductions are expressions of tax non-compliance in many cases.

RSV (2001) reports that almost all audited claimed deductions for “other expenses for earning employment income” were rejected. A random sample of claimed deductions was drawn in a more recent follow-up from 2006. There were errors in 93 percent of the cases. It is, therefore, likely that a large fraction of the claimed deductions actually is tax non-compliance. We simply label the behavior as tax aversion as it is not possible to determine if a particular rejected claimed deduction is tax evasion (the tax payer did not have any expenses but claimed a deduction anyway) or tax avoidance (the tax payer had expenses but these particular expenses where not deductable).

The data set we use consists of 3.6 million Swedish tax payers. The data concern the income year 2006, the tax assessment year is 2007. We use a regression discontinuity kink approach and a regression discontinuity approach. Our large data set allows us to eliminate potential problems of endogeneity and selection in ways that previous empirical studies have not been able to do.

We find behavior to be consistent with loss aversion in the following sense: Tax payers who have a preliminary deficit are more likely to claim deductions for “other expenses” than those who have a preliminary surplus. Loss aversion is the natural candidate for explaining the result theoretically. According to expected utility theory, on the other hand, the deduction probability should evolve smoothly around zero preliminary balance.

Can we rule out selection? The empirical analysis shows that none of the covariates shows a similar evolution around zero preliminary balance. So we are leaning towards a yes!

We also estimate the coefficient of loss aversion in our empirical analysis. Our estimate is lower than the estimate reported by Tversky and Kahneman (1992) but not extremely far off.

The remainder of the paper is organized as follows. In Section 2 we discuss prospect theory and its application to tax compliance. We present a simple model in Section 3, where the tax payers’ decisions are studied. The model provides predictions for the empirical analysis. Our data and the Swedish setting are presented in Section 4. Section 5 presents the empirical results. Concluding remarks are in


2 Prospect theory and tax compliance

Kahneman and Tversky (1979) were the first to define prospect theory. It was later refined by Tversky and Kahneman (1992) into the cumulative prospect theory. The elements of prospect theory we primarily use are reference dependence and loss aversion.

People perceive outcomes as gains and losses rather than final states of wealth because of reference dependence. Loss aversion makes people consider losses as more salient than gains. “…[t]he function is steeper in the negative than in the positive domain; losses loom larger than corresponding gains.” (Tversky and Kahneman, 1991, p. 1039). In risky settings this implies that an individual is more willing to take risks when the individual faces a loss rather than a gain. Tversky and Kahneman (1991), however, show that loss aversion can also explain behavior in the absence of risk.

The existence of loss aversion has been studied and found in many experimental settings. Tversky and Kahneman (1992), Schmidt and Traub (2002), and Abdellaoui et al. (2008) are just a few examples. Loss aversion has been found to explain behavior also in real-life situations. Genesove and Mayer (2001) study the housing market. Pope and Schweitzer (2011) test for loss aversion using data on professional golfers’ performance on the PGA tour. The authors study data on more than 1.6 million putts. They find evidence of loss aversion.

There are several theoretical papers that incorporate prospect theory when studying tax evasion. Dhami and al Nowaihi (2007) set up a rigorous model and make a complete analysis to explain some “tax-evasion puzzles” and how these puzzles can be explained by various components in cumulative prospect theory. The most relevant result for our study is that the evaded amount is increasing in the degree of loss aversion.

Bernasconi and Zanardi (2004) also model tax evasion in the realm of cumulative prospect theory. Their simulations show that if people in the loss domain (who would receive an income lower than the reference income without evasion) evade. They will evade sufficiently to enter the gain domain if not caught. In the gain case, people do not evade as much that they risk an income lower than the reference income.

Some papers explicitly focus on the importance of advance tax payments to deter tax evasion in the light of prospect theory. Yaniv (1999) sets up a simple theoretical model to show how advance payments reduce tax evasion. The intuition is clear: Those who have paid too much in advance get a refund, which is considered a gain. Under prospect theory, the value function is concave for gains, implying risk aversion. On the other hand, if the advance tax payments are lower than actual tax liabilities, a tax payer is to pay the difference and, therefore, experience a loss. Under prospect theory the value function is convex for losses, implying risk seek-
Such a tax payer might, therefore, be more willing to take the risk of evading. It is, therefore, clear that advance tax payments higher than true tax liabilities deter tax evasion.

Elffers and Hessing (1997) also find that advance payments will promote compliance. They, however, also point to the fact that withholding too much may make people feel wrongly treated. This has an opposing effect. They also claim that a standard deduction would increase tax compliance.

There are also some empirical studies that focus the effects of loss aversion on tax evasion. Cox and Plumley (1988) find that the share of tax returns needing correction increases with balance due. Chang and Schultz Jr (1990) also study how compliance depends on over- and under-withholding at the time of filing using actual tax return data. They find that those who owe additional taxes when they file their returns are less compliant than those who have refunds due. This holds true in all income brackets.

Persson (2003) is a previous descriptive study using Swedish data. It reports results similar to the two above mentioned American studies. She finds that people with a tax deficit are more prone to claim dubious deductions. Moreover, all claimed deductions in one geographic region were audited. Those with tax deficits were more likely not to have their claimed deductions approved than those with tax surpluses.

Kirchler and Maciejovsky (2001) have data on self-reported tax evasion and use them to explain compliance behavior in an experimental setting. They find evidence of both reference dependence and loss aversion. Schepanski and Shearer (1995) especially focus on the effect of withholding in an experimental study where the subjects are undergraduate students. They find that hypothetical tax payers who are under-withheld are more likely to underreport income. Also Robben et al. (1990) use students in a hypothetical experiment where they find that having to pay extra tax after insufficient tax has been withheld leads to more evasion than receiving a refund after too much has been withheld.

All the above mentioned studies have focused on tax evasion as a risky action. If the tax payer evades and is caught, the tax payer needs to pay a fine in addition to the tax. The specific situation that we study is different, however. We study deductions for expenses that the tax payer may or may not have had. The deduction may not be justified even if the tax payer have had the expense as the particular expense may not be deductable. Hence, we cannot be entirely sure if it is tax evasion, tax avoidance, or something else going on. We can, however, be certain that those who claim the deduction are tax averse. They find it worthwhile to claim the deduction although that it comes at a cost (administrative and/or moral). Moreover, we restrict our analysis to small claimed deductions. The probability of having these claimed deductions audited was negligible at the time of the study. And if the deduction was audited and not accepted, there was no additional fine for sufficiently small deductions). Hence, the choice to claim the deduction was

4The tax law has changed since. The claimed amount deducted will have to be above a certain
completely without risk.

Some aspects of prospect theory, such as risk aversion in the gain domain and risk seeking in the loss domain are, therefore, not applicable in our case. Loss aversion and reference dependence may, however, be as important in a risk free situation as in a risky one (Tversky and Kahneman, 1991). We can identify the impact of loss aversion on tax compliance by studying how the preliminary tax balance affect deduction behavior.

3 An illustrative theoretical model

The purpose of this section is to provide predictions for our empirical exercise. We, therefore, keep the analysis as simple as possible. It is beyond our scope to provide a complete model of how tax compliance is related to advance tax payments. Hence, we do not use the complete prospect theory as presented by Kahneman and Tversky (1979) and Tversky and Kahneman (1992) to study tax compliance. Instead, we only use of the two components loss aversion and reference dependence.

Consider a tax payer \( i \) who is about to file his tax return. On his tax return, he receives information about his preliminary balance in taxes, \( B_{pi} \), to be refunded (\( B_{pi} > 0 \)) or taxes due (\( B_{pi} < 0 \)). He compares this preliminary balance to his reference point. What should be the valid reference point has been widely discussed.\(^5\) Dhami and al Nowaihi (2007) argue that legal after-tax income should be used as the reference point. This is related to the idea that the reference point should be based on rational expectations (Köszegi and Rabin, 2006), i.e., how much the tax payer expects to owe or to get refunded.

However, both these lines of reasoning require that the tax payers really know their true tax liabilities and can make correct calculations of what their preliminary tax balance should be. In Sweden, as in many other countries, preliminary tax advances are paid at source by the employers. It is likely that most people perceive their advance payments to be what they are supposed to pay unless there are special circumstances. We, therefore, assume that the reference point is a zero preliminary tax balance.\(^6\) Hence, if the preliminary balance is a deficit for tax payer \( i \), i.e., \( B_{pi} < 0 \) he is due to pay more taxes and he experiences himself to be in the loss domain. If, on the other hand, the preliminary balance is a surplus, he will get a tax refund, i.e., \( B_{pi} > 0 \) and he is in the gain domain.

In our case the presentation becomes more clear if we instead use the preliminary deficit defined as the negative of the preliminary balance. The preliminary deficit for tax payer \( i \) is \( D_{pi} \equiv -B_{pi} \). With reference dependence and loss aversion, we assume the following value function (c.f. Tversky and Kahneman, 1992; threshold amount. There are, therefore, no longer any small deductions with extremely low audit probabilities.

\(^5\)See, e.g., the discussion in Kirchler and Maciejovsky (2001).

\(^6\)We present sensitivity analyses allowing other reference points in Section 5.
Köbberling and Wakker, 2005):

\[
V(D_i^p) = \begin{cases} 
  v(-D_i^p) & \text{if } D_i^p \leq 0, \\
  \lambda v(-D_i^p) & \text{if } D_i^p > 0,
\end{cases}
\]  

(1)

where \( \lambda > 1 \) represents the coefficient of loss aversion.

The tax payer may claim a deduction to reduce his tax liability. We restrict our analysis to small deductions, since a large majority of the small deductions have proven not to be justified. There may be other things going on for larger claimed deductions. This might give rise to selection problems in empirical analysis. Moreover, the perceived probability of audit is almost zero for the smallest deductions. This suggest that it is almost without any risks to claim the deduction.\(^7\)

3.1 The choice on the extensive margin

A first step is to predict the probability of actually claiming a deduction to reduce the tax liability. The tax payer’s only choice in our very simplified model is whether to claim a certain deduction. We limit the analysis to small deductions and assume that the deduction is a fixed amount, \( \delta \), the same for all tax payers. This is to make the analysis as illustrative as possible. The tax payer knows (or at least believes) that he is at no risk of being audited if he claims the deduction because the claimed deduction is so small.\(^8\)

However, claiming the deduction \( \delta \) comes at a certain cost, \( c_i \). This cost varies across tax payers, \( c_i \sim U[0, \bar{c}] \). It may reflect the administrative cost of claiming the deduction or the moral cost of doing so if the deduction is not rightful. The deduction is worth \( t\delta \), where \( t \) is the constant marginal tax rate, in monetary terms.

The tax payer compares the value of his preliminary tax balance, \( V(-D_i^p) \) to the value if he claims the deduction, \( V(-D_i^p + t\delta) - c_i \). He claims the deduction if the latter exceeds the former. Depending on the sign and amount of \( D_i^p \), there are three different domains where the tax payer may end up. The three conditions for claiming a deduction of size \( \delta \) are:

A: \( c_i < \lambda v't\delta \) if \( D_i^p > t\delta \),

B: \( c_i < v'[t\delta - D_i^p(1 - \lambda)] \) if \( D_i^p \in (0, t\delta] \),

C: \( c_i < v't\delta \) if \( D_i^p \leq 0 \).

(2)

For simplicity, we assume a linear value function, implying constant marginal values. We disregard declining sensitivity, another aspect of prospect theory, by this...\(^7\)The tax needs to be paid in case of audit. There are, however, no fines for incorrectly claimed deductions of low amounts.

\(^8\)Hence, we abstract from any risk of detection in our model contrary to previous studies of tax evasion. Our main results remain valid even if we include a risky choice of evasion.
assumption. A linear approximation should, however, work perfectly well to illustrate the reasoning since we limit our analysis to a narrow area around zero. The value of the deduction in the three cases above are illustrated in Figure 1.

![Figure 1: The value of claiming a deduction in three cases depending on the preliminary deficit.](image)

We assume a uniform distribution of the cost of claiming the deduction. This makes it straightforward to use the above conditions to predict the share of tax payers deducting at various preliminary deficits. The threshold cost for claiming the deduction in the loss domain is lower than in the gain domain as $\lambda > 1$. Hence, a larger share of the tax payers should claim $\delta$ in the loss domain than in the gain domain. Moreover, the share of taxpayers who deduct should be independent of the preliminary balance when the balance is positive as we assume a linear value function, i.e., that $v'( -D_p^i ) = v$. The same applies for those with a large preliminary deficit (where $D_p^i > t\delta$), who still will have a deficit even if they claim the deduction. It is only in the middle group, (where $D_p^i \in [ -t\delta, 0 ]$) where the share of tax payers deducting is expected to increase in the preliminary deficit:

$$
\Delta = v [ t\delta - D_p^i (1 - \lambda) ],
$$

$$
\frac{\partial \Delta}{\partial D_p^i} = v(\lambda - 1) > 0.
$$

(3)

Assuming loss aversion and reference dependence with a reference point of zero, we predict the pattern for the share of tax payers claiming the deduction to be as shown in Figure 2:

The pattern in Figure 2 can be summarized as follows:

**Prediction 1.** A larger share of tax payers with a preliminary deficit will claim the deduction $\delta$ than tax payers with a preliminary surplus.

**Prediction 2.** The probability of claiming the deduction $\delta$, as a function of the preliminary deficit, has the shape as in Figure 2 with kinks at 0 and at a preliminary deficit $D = t\delta$.  

7
The above connection can then be used to actually estimate the coefficient of loss aversion, $\lambda$. Tversky and Kahneman (1992) originally estimated $\hat{\lambda} = 2.25$ using experimental data where each individual was observed in both the gain and the loss domain. The studies following have used a similar technique to estimate $\lambda$ (see, e.g., Abdellaoui et al., 2007, and references therein).

Our approach is different. We plot the shares claiming the deductions for each level of preliminary tax deficits as sketched in Figure 2. The share claiming the deduction if they are in the gain domain is:

$$\int_{c=0}^{vt\delta} f(c)dc = X, \quad (4)$$

and in our data, we observe the actual share, $X$. If we assume the cost to be uniformly distributed, (4) is easily solved to yield

$$vt\delta = X. \quad (5)$$

Likewise, the share claiming the deduction in the loss domain, where $D_i > t\delta$, is

$$\int_{c=0}^{\lambda vt\delta} f(c)dc = Y \Rightarrow \lambda vt\delta = Y. \quad (6)$$

Equations (5) and (6) then gives $\lambda = \frac{Y}{X}$. Hence, with a uniformly distributed cost of detecting, the coefficient of loss aversion is simply the ratio between the shares claiming deductions.

---

*Köbberling and Wakker (2005) and Abdellaoui et al. (2008) are two recent papers that elaborate on how to define and measure the coefficient of loss aversion appropriately in situations involving risk.*
Prediction 3. If \( c_i \sim U[0, \bar{c}] \), where \( c_i \) is the cost of claiming the deduction, the coefficient of loss aversion, \( \lambda \), is:

\[
\lambda = \frac{Y}{X},
\]

where \( Y \) is the share deducting in the loss domain, where \( D_i > t\delta \), and \( X \) is the share deducting in the gain domain.

3.2 The choice on the intensive margin

We will now drop the assumption of a constant deducted amount and instead allow the deducted amount to vary. What will determine the tax payer’s choice on the intensive margin? In other words, we will study what determines the size of the deducted amount conditional on claiming a deduction.

Previously, we assumed a certain cost \( c_i \) associated with claiming the given deducted amount. We will now assume a convex cost function, \( c_i(\delta) \), where \( c_i' > 0 \) and \( c_i'' > 0 \). We continue to assume a linear value function. This should be a good approximation in the neighborhood where \( D_p^i \) is close to zero.

Hence, the value function in the gain domain is:

\[
V_i = v(t\delta - D_p^i) - c_i(\delta),
\]

and the interior solution is characterized by:

\[
vt = c_i' \Rightarrow \delta_{ig} : vt = c_i'.
\]

However, the individual will only claim the deduction if the benefit exceeds the cost, i.e., if \( vt\delta_i \geq c_i(\delta_i) \) for the optimal \( \delta_{ig} \).

The solution from (8) is valid, not only for \( D_p^i < 0 \), but for all interior solutions where \( D_p^i \leq t\delta_{ig} \), i.e., as long as the optimal deduction, \( \delta_{ig} \) is sufficiently large to bring the tax payer into the gain domain.\(^{10}\)

For who start in the loss domain and remain there:

\[
V_i = \lambda v(t\delta - D_p^i) - c_i(\delta),
\]

the interior solution is characterized by:

\[
\lambda vt = c_i' \Rightarrow \delta_{il} : \lambda vt = c_i'.
\]

and is valid for large initial deficits \( D_p^i > t\delta_{il} \). Moreover, the condition for an interior solution is \( \lambda vt\delta_{il} \geq c_i(\delta_{il}) \).

For those who end up in either of these solutions, the absolute value of \( D_p^i \) is not important for the amount deducted. This is an artefact of the linear value function. The individual heterogeneity instead comes from variations in the cost

\(^{10}\)If originally in the loss domain, the condition for an interior solution according to (8) is that \( \lambda vtD_p^i < v(t\delta_i - D_p^i) - c_i(\delta_i) \).
function. However, for a given cost function, the individual makes a larger deduction if ending up in the loss than in the gain domain. Hence, $\delta_{il} > \delta_{ig}$.\footnote{When claiming a deduction in the gain domain (8) states that $vt = c_{ig}'$, and in the loss domain (10) that $\lambda vt = c_{il}'$. Since $\lambda > 1$, this means that $c_{il}' > c_{ig}'$, which in turn implies that $\delta_{il} > \delta_{ig}$ for a given cost function.}

For an individual where $D^p_i \in (t\delta_{ig}, t\delta_{il})$, the deduction is $\delta_i = \frac{D^p_i}{t}$, i.e., $\delta_i \in [\delta_{il}, \delta_{ig})$ implying that he reaches exactly the reference point by claiming the deduction.\footnote{Assume $\delta_i < \frac{D^p_i}{t}$. Then he will remain in the loss domain and since $\lambda vt > c'(\delta_l)$ this would yield a lower value than without the deduction. Instead assume $\delta_i > \frac{D^p_i}{t}$. Then he will enter the gain domain and since $vt < c'(\delta_l)$ this deduction would be too expensive on the margin. Hence, $\delta_i = \frac{D^p_i}{t}$.}

Hence, in which domain one ends up given a certain deficit, $D^p_i$, depends on the individual’s cost function. For a certain cost function, the deducted amounts in interior solutions are depicted in Figure 3:

![Figure 3: The size of the deduction depending on the preliminary deficit.](image-url)

### 3.3 Simulating deductions

We select a specific cost function and simulate the effects to illustrate what we could expect in terms of predictions for the intensive margin:

$$c_i = \bar{c}_i + \gamma_i \delta_i^\sigma, \quad \sigma > 1, \quad \gamma_i > 0,$$

where

$$\bar{c}_i = c_0 + (1 - \kappa) e_c + \kappa e$$
$$\gamma_i = \psi(\gamma_0 + (1 - \kappa) e_\gamma + \kappa e)$$

determine the fixed and variable parts of the cost, respectively. $\psi > 0$ is a simple scaling parameter; $e_c$, $e_\gamma$ and $e$ are uniformly distributed between $e_{j\min}$ and $e_{j\max}$.\footnote{Assume $\delta_i < \frac{D^p_i}{t}$. Then he will remain in the loss domain and since $\lambda vt > c'(\delta_l)$ this would yield a lower value than without the deduction. Instead assume $\delta_i > \frac{D^p_i}{t}$. Then he will enter the gain domain and since $vt < c'(\delta_l)$ this deduction would be too expensive on the margin. Hence, $\delta_i = \frac{D^p_i}{t}$.}
The parameter \( \kappa \) determines the covariance structure between \( \gamma_i \) and \( \bar{c}_i \). If \( \kappa > 0 \) those who have a high fixed cost also have (on average) higher marginal cost of making deductions. Baseline is \( \kappa = 1 \) (corr=1).

We simulate the deduction choices over the range \( D^p_i \in [-15,000, 15,000] \) in two steps. The first, namely the probability of claiming the deduction, is shown in Figure 4. Note that it is preliminary deficit, i.e., the negative of preliminary balance on the x-axis.

![Figure 4: The share of tax payers claiming a deduction depending on the preliminary deficit.](image)

The pattern is the same as the one shown in Figure 2. The average simulated amounts deducted for those who claim a deduction are shown in Figure 5 as a function of the preliminary deficit.

The deducted amounts are larger in the loss domain than in the gain domain according to the simulations. We also see that the average deducted amount actually increases in the preliminary deficit in the “middle domain”. We can, therefore, make further predictions for the empirical analysis:

**Prediction 4.** Conditional on claiming a deduction, tax payers in the loss domain on average make larger deductions than those in the gain domain.

### 4 Descriptives

Our entire data set covers data on 4.6 million Swedes, 16–67 years old, filing their tax returns in 2007 for the income year 2006. We have access to a limited number of variables: employment income, marginal tax rate, gender, age, claimed deduction for other expenses for earning employment income, and the preliminary balance in taxes due.

Assessed employment income includes salaries, social insurance system benefits (such as sickness benefits and parental benefits), and unemployment benefits.
Approved costs for commuting to work are subtracted. Assessed employment income also includes public pensions and occupational pensions. Retirees cannot be expected to have salary income and, therefore, claim deductions other expenses for earning employment income. This explains the upper age limit in our sample.

The sequence of events for a tax payer is as follows:\textsuperscript{13} Employees pay employment income taxes at source during income year. Employers transfer the tax payments to the Tax Agency.

The Tax Agency sends a preliminary income tax return to the tax payer in April the following year (the assessment year). This tax return is based on the statements of income the Tax Agency has received from employers, banks, etc. The Tax Agency also calculates a preliminary balance in taxes due. The preliminary balance can show a deficit (more taxes due) or a surplus (there will be a tax refund). The tax payer can then add missing information to the tax return, for instance claiming deductions.

We would like the preliminary balance to be a correct signal from the Tax Authority to the tax payer. The preliminary balance should also be exogenous to the tax payer. We, therefore, exclude tax payers who have earned business income from the sample. The preliminary balance can be expected to be endogenous for self-employed.

Tax payers can ask taxes at source during the income year to be based on their individual economic situation instead of the default rules. The preliminary balance can be expected to be endogenous for these tax payers too. We exclude these tax payers from the sample. Finally, we exclude taxpayers who actively have made

\textsuperscript{13}See also the information from the Tax Agency reproduced in Appendix A.
advance tax payments directly to the Tax Agency at their own initiative in the beginning of the assessment year. We can expect the preliminary balance to be endogenous also for these tax payers. These selection criteria leaves us with a sample of 3.6 million tax payers.\footnote{Descriptive statistics are presented in Appendix B.}

Why not use the defaults rules? And why make advance payments? One reason is that taxation at source and the Tax Authority’s preliminary calculations of taxes due cannot correctly take some taxes into account. This is because the Tax Authority does not have all the necessary information from the statements of income received from employers, banks, etc. This concerns the wealth tax (repealed in 2007), capital taxation of realized capital gains (financial instruments, real estate), and taxation of business income. Our selection criteria should exclude most tax payers who pay these taxes.

In addition, only tax payers with normal annual employment income are included in the sample. We interpret normal annual employment income to be in the interval SEK 100,000–1,000,000.

At the time for our study, the income year 2006, assessed employment income was reduced by approved other expenses for earning employment income exceeding SEK 1,000 (USD 150 in December, 2006). So which types of expenses were approved? Some examples are expenses for safety equipment, safety clothes, tools, and instruments related to work not paid for by the employer, expenses for an office if the employer does not provide one (an office at home is not approved in general), expenses for books and journals related to work for some occupations if not paid by the employer, and expenses for phone calls related to work if not paid by the employer (not the phone and not the subscription).

In previous studies where claimed deductions for “other expenses” have been randomly audited, 90–95 percent have not been approved. Hence, claiming such deductions is a clear sign of tax aversion. In some cases, the tax payer may simply not understand the rules and believe that he is entitled to the deduction. In other cases, the tax payer may take a chance. But some claimed deductions are clearly not attempts not to comply. This is likely to be true especially for large claimed deductions. It is difficult to “make up” very large expenses. Suppose that there is a tax payer with a large preliminary deficit who claims a large deduction. One might then suspect that the preliminary balance is endogenous, the tax payer has in advance planned to claim the deduction to reach a zero tax balance in the end. We, therefore, exclude those with substantial claimed deductions. And we only study those with a preliminary balance of ± SEK 5,000 in this section. This further reduced to only those with a preliminary balance of ± SEK 3,000 in the empirical analysis reported in Section 5.

Figure 6 shows that the probability of claiming a deduction is independent of the preliminary balance for those with a surplus. It is, however, increasing in the preliminary deficit for those with a deficit. Most importantly, the probability of claiming a deduction is higher for tax payers with a preliminary deficit than for
those with a preliminary surplus. But is this a causal effect of the preliminary balance or simply selection? A simple graphical test of selection consists of plotting the distribution of individuals over preliminary deficits. If the distribution changes around zero we might have problems with selection. As we can see from Figure 7 below, the distribution does not seem to kink or jump at zero.

We could, however, still have problems with selection that do not show up in the frequency distribution plot. The individuals slightly below zero could be very different from the individuals slightly above even if the distribution is smooth around the reference point. We, therefore, need to take a closer look at how the covariates evolve around the reference point. If there is selection based on any of the covariates, or selection based on any unobservable factor that is correlated with the covariates, it would show up as a kink or discontinuity around zero. The pattern in Figure 6 could then be due to selection. If, on the other hand, all covariates evolve smoothly around zero it suggests that there is a causal effect of the preliminary balance.

We will here present descriptive checks if the predetermined covariates show kinks or discontinuities at zero preliminary balance. The formal econometric tests are presented in Section 5. Employment income is, of course, a crucial variable. We check this first.

The blue dots in Figure 8 show the relationship between the preliminary deficit and employment income. The kink slightly below zero preliminary deficit is very natural. The higher employment income, the more difficult it is to calibrate the

Figure 6: The probability of claiming a deduction as a function of the preliminary deficit.
tax payments at source correctly – the errors will be scaled up in proportion to the income. We should, therefore, end up further away from zero preliminary balance, the higher the employment income is. This is exactly the pattern we see in the absolute version of the relationship.

The employment income tax is progressive. The tax payer may move into a higher tax bracket limit if he earns unexpectedly much during a year. This might contribute to a higher preliminary deficit. The relationship is steeper on the deficit side than on the surplus side. This may be because of the design of the income tax.

The absolute relationship in Figure 8 suggests that we do not have a correct specification. There is, however, a possible solution to the problem. We might ask: Is the impact of having SEK 1,000 in preliminary deficit the same for a high income tax payer and a low income tax payer? It can be argued that the marginal utility of claiming a deduction is higher for a low income tax payer. There are, therefore, reasons to instead measure the preliminary deficit in relation to the income level.

Define the relative preliminary deficit of tax payer \(i\) as:

\[ d_i^p = (D_i^p / E_i) * \bar{E}, \]

where \(d_i^p\) is the relative preliminary deficit, \(D_i^p\) is the absolute preliminary deficit, \(E_i\) is employment income, and \(\bar{E}\) is average employment income. The kink at zero preliminary deficit disappears when we measure the preliminary deficit relative to employment income, see the red dots in Figure 8.

The other conditioning variables do not seem to have kinks or discontinuities at zero preliminary deficit. Figure 9 and Figure 10 show this. And it does not matter...
According to Prediction 3, the coefficient of loss aversion could be expressed as the share of deducting taxpayers in the loss domain (with sufficiently large losses) over the share of deducting taxpayers in the gain domain for a certain deduction.

The share of tax payers with preliminary deficits that claim deductions is 6.2 percent in our full sample of 3.6 million tax payers. The corresponding share in the full sample for those with preliminary surpluses is 4.4 percent. A rough estimate of loss aversion, given these assumptions, therefore, is:

$$\hat{\lambda} = \frac{\hat{Y}}{\hat{X}} = \frac{6.2}{4.4} \approx 1.41.$$  \hspace{1cm} (12)

Our estimate is lower than the estimate reported by Tversky and Kahneman (1992), $\hat{\lambda} = 2.25$, but not extremely far off.
Figure 9: Gender as a function of the preliminary deficit.

Figure 10: Age as a function of the preliminary deficit.
5 Estimations

5.1 Baseline models

A main objective in this section is to formally test whether the observed impact of the preliminary deficit on the probability of claiming deductions is causal and not just due to selection. We follow some of the empirical strategies suggested in the manuals for regression discontinuity design (Lee and Lemieux, 2010) and regression kink design (Card et al., 2009). The empirical tests essentially consist of answering two questions:

- Does the impact of the preliminary balance on the probability of claiming a deduction have a statistically significant kink or discontinuity around zero preliminary balance?

- Can we rule out corresponding statistically significant kinks or discontinuities for the predetermined covariates?

We, therefore, estimate spline models of the following type (additional covariates are suppressed):

\[ \Delta_i = \alpha_0 + \alpha_1 I_i + \gamma_1 D^p_i + \beta_1 I_i D^p_i + \epsilon_i, \]  

where \( \Delta_i \) is an indicator of claiming a deduction, \( I_i \) is an indicator for a positive preliminary deficit, \( D^p_i \) is the preliminary deficit, and \( \epsilon_i \) is an error term. The coefficient \( \alpha_0 \) measures the intercept, \( \alpha_1 \) measures a potential discontinuity at the reference point (zero preliminary deficit), and \( \beta_1 \) measures the change in linear relationship at the reference point.

We iterate the estimation over a large number of different bandwidths. Bandwidths are defined by a maximum and minimum preliminary deficit. We use symmetric bandwidths around the reference point (zero preliminary deficit), ranging from SEK 3,000 to SEK 500. A bandwidth of, e.g., SEK 3,000 means that we only include individuals with preliminary balance in the range SEK -3,000 to SEK 3,000. A large bandwidth gives a larger sample size and, therefore, more precise estimates, while a smaller bandwidth gives sharper identification but less precise estimates.

It is reasonable not to limit the analysis to a linear model for larger bandwidths. We, therefore, also estimate models that include higher order polynomials. We set the maximum polynomial order to 3 and the minimum to 0. The estimated coefficients of interest are \( \beta_1 \) for a possible kink and \( \alpha_1 \) for a possible discontinuity.

The second and third (polynomial) order specifications are:

\[ \Delta_i = \alpha_0 + \alpha_1 I_i + \gamma_1 D^p_i + \gamma_2 (D^p_i)^2 + \beta_1 I_i D^p_i + \beta_2 I_i (D^p_i)^2 + \epsilon_i, \]  

where \( \Delta_i \) is an indicator of claiming a deduction, \( I_i \) is an indicator for a positive preliminary deficit, \( D^p_i \) is the preliminary deficit, and \( \epsilon_i \) is an error term. The coefficient \( \alpha_0 \) measures the intercept, \( \alpha_1 \) measures a potential discontinuity at the reference point (zero preliminary deficit), and \( \beta_1 \) measures the change in linear relationship at the reference point.

We iterate the estimation over a large number of different bandwidths. Bandwidths are defined by a maximum and minimum preliminary deficit. We use symmetric bandwidths around the reference point (zero preliminary deficit), ranging from SEK 3,000 to SEK 500. A bandwidth of, e.g., SEK 3,000 means that we only include individuals with preliminary balance in the range SEK -3,000 to SEK 3,000. A large bandwidth gives a larger sample size and, therefore, more precise estimates, while a smaller bandwidth gives sharper identification but less precise estimates.

It is reasonable not to limit the analysis to a linear model for larger bandwidths. We, therefore, also estimate models that include higher order polynomials. We set the maximum polynomial order to 3 and the minimum to 0. The estimated coefficients of interest are \( \beta_1 \) for a possible kink and \( \alpha_1 \) for a possible discontinuity.

The second and third (polynomial) order specifications are:

\[ \Delta_i = \alpha_0 + \alpha_1 I_i + \gamma_1 D^p_i + \gamma_2 (D^p_i)^2 + \beta_1 I_i D^p_i + \beta_2 I_i (D^p_i)^2 + \epsilon_i, \]  

where \( \Delta_i \) is an indicator of claiming a deduction, \( I_i \) is an indicator for a positive preliminary deficit, \( D^p_i \) is the preliminary deficit, and \( \epsilon_i \) is an error term. The coefficient \( \alpha_0 \) measures the intercept, \( \alpha_1 \) measures a potential discontinuity at the reference point (zero preliminary deficit), and \( \beta_1 \) measures the change in linear relationship at the reference point.

We iterate the estimation over a large number of different bandwidths. Bandwidths are defined by a maximum and minimum preliminary deficit. We use symmetric bandwidths around the reference point (zero preliminary deficit), ranging from SEK 3,000 to SEK 500. A bandwidth of, e.g., SEK 3,000 means that we only include individuals with preliminary balance in the range SEK -3,000 to SEK 3,000. A large bandwidth gives a larger sample size and, therefore, more precise estimates, while a smaller bandwidth gives sharper identification but less precise estimates.

It is reasonable not to limit the analysis to a linear model for larger bandwidths. We, therefore, also estimate models that include higher order polynomials. We set the maximum polynomial order to 3 and the minimum to 0. The estimated coefficients of interest are \( \beta_1 \) for a possible kink and \( \alpha_1 \) for a possible discontinuity.
The zero order specification, exclusively focusing on a possible discontinuity at zero preliminary balance, is simply:

\[ \Delta_i = \alpha_0 + \alpha_1 I_i + \epsilon_i. \]  

(16)

For each bandwidth we, therefore, estimate the four models (13)–(16). The optimal model for each bandwidth is then determined based on the minimization of a Schwarz Bayesian criterion (SBC) which is an information criterion that essentially weighs a better goodness of fit against lost degrees of freedom. SBC is a modification of AIC that punishes lost degrees of freedom harder. Kink and discontinuity estimates from the optimal model are then plotted over a range of symmetric bandwidths.

Our primary interest (which is based on the simple theoretical predictions made in section 3) is in the \( \beta_1 \) estimate. This parameter measures the change in the derivative at the reference point or, in other words, the kink. But we are also interested in whether there is a discontinuity in the relationship at the reference point. This is measured by \( \alpha_1 \).

Figure 11 reports our estimates of a possible kink at zero preliminary balance, for the deduction indicator variable. The optimal models include polynomials of the second order for bandwidths above SEK 2,500, see the (blue) thick bracketed line. First order models are optimal for bandwidths down to SEK 800. For smaller bandwidths zero order models are optimal with some exceptions.

The estimated kink is statistically significant for bandwidths wider than SEK 800 and for bandwidths around SEK 600. The (red) thick solid line reports the estimated coefficients while the (red) thin solid lines provide the limits of 95 percent confidence intervals. The figure also shows that the estimated coefficients change level downwards when the polynomial order decreases. The estimated coefficient remains stable in a wide range of bandwidths, from SEK 2,500 to SEK 800.

It is also possible that there is a discontinuity in the relationship between the preliminary deficit and the probability of claiming a deduction at zero preliminary balance. Figure 12 reports our estimates of this. The estimated coefficients of a discontinuity at zero are statistically significant for all bandwidths of SEK 800 except for a few bandwidths around SEK 600. It is also clear from the figure that the estimated coefficients change level upwards when the polynomial order decreases. There is an underlying downward sloping trend in the estimated coefficient for the whole range of bandwidths.

We now turn to the corresponding estimates for the three covariates. Figures corresponding to Figure 11 and Figure 12 are presented in Appendix C. Starting with the gender indicator, we estimate a significant negative kink for large bandwidths while the optimal polynomial order is one. For bandwidths smaller than SEK 1,300, however, the optimal polynomial order decreases to zero. For smaller bandwidths than this we estimate a significant negative discontinuity instead of a kink. The discontinuity estimate slopes upwards with respect to the bandwidth and remains significant up until a bandwidth of SEK 750.
Figure 11: Kink - estimates of $\beta_1$.

Figure 12: Discontinuity - estimates of $\alpha_1$. 
The corresponding results for age show decreasing kink estimates for bandwidths lower than SEK 2,500. The estimates becomes insignificant below a bandwidth of about SEK 1,400. The optimal polynomial order is one except for very large bandwidths and very low bandwidths. The discontinuity estimate is insignificant and stable for bandwidths below SEK 1,900 as long as the optimal polynomial order is one. For very small bandwidths, when the kink estimate is kicked out, the discontinuity estimates jumps upwards and becomes significant.

Finally, the results for employment income shows that the kink estimate is negative and significant in the range of bandwidths between SEK 2,700 and SEK 1,400. The optimal polynomial order is zero for smaller bandwidths. The discontinuity estimate is only significant for a small range of bandwidths between SEK 1,400 and SEK 1,000.

Let us compare the results for the deduction probability with the corresponding results for the covariates. The estimated kinks and discontinuities for the covariates jump around much more. The results for deduction probability shows a stable and significant kink for a wider range of bandwidths than the covariates. The discontinuity estimates for the deduction variable slopes downward with respect to the bandwidth, but it also remains significant in a much wider range than the estimates for the covariates.

5.2 Sensitivity analysis – placebo kinks and discontinuities

There is a risk that even the smallest and economically insignificant estimates becomes statistically significant when the sample size is very large. It is, therefore, relevant to ask whether kinks and discontinuities at the zero reference point are more pronounced than kinks and discontinuities at other reference points. We, therefore, present estimates of kinks and discontinuities based on a range of different “placebo” reference points in this subsection. We let the reference points vary between SEK -3,000 to SEK 3,000. The bandwidth is fixed at SEK 1,000 in all regressions.

The analysis serves two closely related purposes. First, the causal story gets stronger if we find that the kinks and discontinuities in deduction probability are more pronounced for reference points predicted by our theoretical model. Second, it also speaks against selection driving the increase in deduction probability around zero if the kinks and discontinuities for the covariates are not more pronounced around zero than elsewhere.

Figure 13 below shows the estimated kinks in deduction probability for different placebo reference points. We clearly see that the kink estimates peak slightly below zero (at a reference point of about SEK -200). Furthermore, the concave increase in deduction probability at the deficit side (see Figure 6) generates moderate negative kinks as long as the optimal polynomial order is one. Further to the right, for reference points above SEK 1,700, the optimal polynomial order is zero.

The corresponding discontinuity estimates displayed in Figure 14 below mirror the kink estimates. There is a local peak in the discontinuity estimates for reference
points close to zero. Further to the right, when the kink estimate is kicked out of the model, the moderate increases in deduction probability (see Figure 6) generate large and significant discontinuity estimates.

We conclude that the estimated kinks for different reference points follow a pattern that is highly consistent with loss aversion based on a reference point close to zero. Loss aversion does not, in its simplest form, generate a discontinuity in the deduction probability around zero. The empirical evidence for a discontinuity is also weaker as the largest, and most significant, discontinuity estimates are found at reference points between SEK 1,700 and SEK 3,000.

We now turn to the corresponding analysis of the covariates. See Appendix D for the figures. We find that neither kink nor discontinuity estimates are more pronounced around the zero reference point than elsewhere for any of the covariates. We are, therefore, confident to conclude that selection is a very unlikely source of the increase in deduction probability around zero. It would require very strong selection on some unobservable factor to produce the pattern we see in Figure 6. It is hard come up with a candidate for such an unobservable factor in itself. It is even harder provided that it also needs to be virtually uncorrelated with gender, age, and employment income.

5.3 Interpretation and discussion

Figure 15 reports the combined effects of estimated kinks and discontinuities for the different bandwidths. Suppose that we compare the probability of claiming a deduction for a tax payer with a SEK 2,000 in preliminary deficit with and with-
out the estimated effects of kinks and discontinuities. The figure shows that the differences in the probabilities of claiming a deduction is 2.5 percentage points for bandwidths larger SEK 2,500 (when the optimal models are of the second order). The differences in probabilities are about 2 percentage points for smaller bandwidths down to SEK 800 (when the optimal models are of the first order). For the smallest bandwidths the probability differences are about 1 percentage point (when the optimal models are of the zero order).

Let us briefly summarize the econometric results. We find clear evidence of a kink at zero preliminary balance for the deduction probability for a wide range of bandwidths. Employment income and age does not kink or jump at zero preliminary balance for smaller bandwidths. The fraction of men shows a very small but statistically significant jump at zero. However, a complementary analysis with variable reference points shows that the fraction of men jumps even more at other points than zero.

6 Concluding remarks

The objective of this paper is to study if Swedish tax payers behave in a loss averse manner when not complying to taxation. This is important for tax design but also for understanding human behavior in general. The predictions of expected utility theory can be contrasted to those of prospect theory. We have access to data for 3.6 million tax payers for the income year 2006. Our research method is to use a regression kink approach and a regression discontinuity approach.

We find behavior to be consistent with loss aversion. Tax payers who have a
preliminary tax deficit are more likely to claim deductions for “other expenses for earning employment income” than those who have a preliminary surplus. According to expected utility theory the deduction probability should evolve smoothly around zero. Loss aversion is the natural candidate for explaining the result theoretically. Can we rule out selection? The empirical analysis shows that none of the covariates shows a similar evolution around zero preliminary deficit.

We also estimate the coefficient of loss aversion in our empirical analysis. Our estimate, \( \hat{\lambda} = 1.41 \), is lower than the estimate reported by Tversky and Kahneman (1992), \( \hat{\lambda} = 2.25 \), but not extremely far off.

There are two additional observations that we have made when studying these data. Both of these are starting points for possible future research.

- Conditional on actually claiming a deduction, the deductions of those with preliminary deficits have larger variance than the deductions of those with preliminary surpluses.
- We also have information on corrections and rejections of claimed deductions. The Tax Agency makes larger corrections of the claimed deductions for tax payers with preliminary deficits than for those preliminary surpluses.
References


Appendix A  The Tax Agency’s information to tax payers
Filing an income tax return

Everyone receiving an income is required to file a tax return (self assessment) the year after the income year. The income year is the year in which the income (e.g., wages or pensions) is paid out and your employer – or whoever pays out your pension – makes a tax deduction for it.

At the beginning of April, the Tax Agency will send you:

- **A tax return form**
  Everyone required to declare income will receive a tax return form, ‘Inkomstdeklaration’. Many particulars on the form have already been filled in by the Tax Agency (Skatteverket).

- **An income specification**
  This is a list itemising the income statements (kontroll-uppgifter) sent to both you and the Tax Agency.

- **A preliminary tax estimate**
  You will also get a preliminary estimate of your tax. It will be based on the information the Tax Agency has filled in on your tax return.

- **Payment slips**
  In addition, you will be sent tax payment slips that you can use if you need to pay more tax.

---

### Declare your income on the Internet or by phone

Most taxpayers can file their tax returns on the Internet. Visit [www.skatteverket.se](http://www.skatteverket.se) to find out who can use this service and how to proceed it.

To file your tax return on the Internet you go into [www.skatteverket.se (E-jänster → Inkomstdeklaration)](http://www.skatteverket.se (E-jänster → Inkomstdeklaration)).

If you don't need to make any changes, you can approve your income tax return by telephone (020-567 100) or via sms (711 44).

---

If 2 May is a Saturday or Sunday, tax returns must be submitted by Monday at the latest.

---

<table>
<thead>
<tr>
<th>Income year</th>
<th>Assessment year (the year after the income year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>January-December</td>
<td>January-May</td>
</tr>
<tr>
<td>All the income you have received during this year must be declared in the following year’s income tax return.</td>
<td>In January, you will be sent income statements for your earnings the previous year. Your employer sends the income statement both to you and to the Tax Agency. You will be sent your tax return form in April.</td>
</tr>
</tbody>
</table>

By 2 May at the latest the Tax Agency must have received your income tax return.
**Kontoutdrag**

Datum: S

Kontonummer

Vid inbetalning: Ditt referensnummer (OCR):

Ditt konto vid utbetalning:

- Plusgiro
- Bankgiro

580303-2805

2009-08-03

Skattekontoret Wistad

Box 122

123 45 Wistad

Andersson, Margareta

Badhusplanen 2

123 48 Wistad

Skatteupplysningen 0771-567 567

489 01 03-7

19580303280541

9999-99 999 99

5050-1055 1958030328054

Kontoutdrag för perioden 3 augusti 2008 - 1 augusti 2009

Specifikation Ränta fr.o.m.

<table>
<thead>
<tr>
<th>Belopp (kr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ingaende saldo 2006-08-02</td>
</tr>
<tr>
<td>Avdragen skatt enligt kontrolluppgifter</td>
</tr>
<tr>
<td>Intäktsränta</td>
</tr>
<tr>
<td>Överskott före eventuell utbetalning</td>
</tr>
<tr>
<td>Utbetalning</td>
</tr>
<tr>
<td>Utgående saldo 2007-07-31</td>
</tr>
<tr>
<td>070215</td>
</tr>
<tr>
<td>070215 0</td>
</tr>
<tr>
<td>-73400</td>
</tr>
<tr>
<td>76800</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>3420</td>
</tr>
<tr>
<td>-3420</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

Utbetalningsorsak:

- Slutf’int skatt enligt 2007 års taxering

Utbetalningsbart belopp:

- Överskott före eventuell utbetalning

Uppgifter om utbetalning från skattekonto:

**Besked**

Slutf’int skatt enligt 2009 års taxering

Kommunal, Län, Kom, Förs region kontor, Sek, Grupp, Person-/Org.-/Reg.nr *

Uppges utöver namn och adress vid skriftväxling

**Everyone has a tax account**

The tax account provided by the Tax Agency shows your preliminary tax figure based on income statements as well as your own tax payments, your final tax figure and other details.

**Tax refund**

If you are salaried or a pensioner you can receive your tax refund by June. But to do this you must

- file your tax return via the Internet, phone or sms and

- have specified a bank account for refund of tax.

**How to fill out the form:**

- Check that all income statements are included in the specification sent to you and that the amounts are correct. Most of the information in the specification is also filled in on your tax return form.

- Are the amounts filled in on the tax return form correct and is all the information included?

- File your tax return via the Internet with security code, by phone, by sms – or sign the form and send it in!

- Correct what is not correct! Add any missing information! You can do this on the Internet with electronic legitimation or on the form.

- Sign (electronically or on the form) and send it in!

**June**

In June a final tax statement (slutskattebesked) and a statement of account (kontoutdrag) are sent out to most employees and pensioners who have filed a tax return, via the Internet, by phone or by sms and who are due to get a refund. This applies only to those who have specified a bank account.

**August - September**

By September at the latest, final tax statements and statements of account are sent out to most people who have sent in an income tax return and who have not received a final tax statement in June. Those who have paid too much tax will now get their money back.

**December**

By mid-December at the latest, final statements and statements of account are sent out to those who did not get them before. Those who have paid too much will now get their money back.

<table>
<thead>
<tr>
<th>June</th>
<th>August - September</th>
<th>December</th>
</tr>
</thead>
<tbody>
<tr>
<td>In June a final tax statement (slutskattebesked) and a statement of account (kontoutdrag) are sent out to most employees and pensioners who have filed a tax return, via the Internet, by phone or by sms and who are due to get a refund. This applies only to those who have specified a bank account.</td>
<td>By September at the latest, final tax statements and statements of account are sent out to most people who have sent in an income tax return and who have not received a final tax statement in June. Those who have paid too much tax will now get their money back.</td>
<td>By mid-December at the latest, final statements and statements of account are sent out to those who did not get them before. Those who have paid too much will now get their money back.</td>
</tr>
</tbody>
</table>
## Appendix B  Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>preliminary deficit</th>
<th>preliminary surplus</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of observations</td>
<td>810,690</td>
<td>2,800,282</td>
<td>3,610,972</td>
</tr>
<tr>
<td>deducting, fraction</td>
<td>0.062</td>
<td>0.044</td>
<td>0.048</td>
</tr>
<tr>
<td>deduction, SEK</td>
<td>291 (1,563)</td>
<td>156 (1,010)</td>
<td>186</td>
</tr>
<tr>
<td>preliminary balance, SEK</td>
<td>-10,757 (51,232)</td>
<td>6,553 (6,762)</td>
<td>2,667</td>
</tr>
<tr>
<td>employment income, SEK</td>
<td>279,770 (135,523)</td>
<td>269,296 (115,198)</td>
<td>271,648</td>
</tr>
<tr>
<td>men, fraction</td>
<td>0.48</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>age, years</td>
<td>46.8 (12.5)</td>
<td>42.0 (12.1)</td>
<td>43.1</td>
</tr>
</tbody>
</table>

*Note.* mean (standard deviation)
Appendix C  Estimates for covariates

Figure 16: Kink - estimates for employment income.

Figure 17: Kink - estimates for gender.
Figure 18: Kink - estimates for age.

Figure 19: Jump - estimates for employment income.
Figure 20: Jump - estimates for gender.

Figure 21: Jump - estimates for age.
Appendix D  Placebo estimates for covariates

Figure 22: Kink - placebo estimates for employment income.

Figure 23: Kink - placebo estimates for gender.
Figure 24: Kink - placebo estimates for age.

Figure 25: Jump - placebo estimates for employment income.
Figure 26: Jump - placebo estimates for gender.

Figure 27: Jump - placebo estimates for age.