Tax compliance and loss aversion*

Per Engström  Katarina Nordblom  Henry Ohlsson  Annika Persson

November 15, 2013
Incomplete, under revision

Abstract

We study if taxpayers are loss averse when filing returns. Preliminary deficits might be viewed as losses assuming zero preliminary balances as reference points. Swedish taxpayers can to try to escape such losses by claiming deductions after receiving information about the preliminary balance. Using a regression kink and discontinuity approach, we study data for 3.6 million Swedish taxpayers for 2006. There are strong causal effects of preliminary tax deficits on the probability of claiming deductions. Compliance will increase and auditing costs will be reduced if preliminary taxes are calibrated so that most tax payers receive refunds.

Keywords: tax compliance, loss aversion, prospect theory, quasi-experiment, regression kink, regression discontinuity

JEL Codes: C21, C26, D03, H24, H26

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 51 04, fax: +46 18 471 14 78

*We would like to thank Spencer Bastani, Fredrik Carlsson, Edoardo Di Porto, Martin Dufwenberg, Olof Johansson-Stenman, Martin Kocher, David Lee, Oskar Nordström Skans, Matthias Sutter, Måns Söderbom, Benno Torgler, Magnus Wikström, seminar participants at Deakin; Melbourne, Gothenburg, Växjö, UNSW; Sydney, UCFS, the UCFS Scientific Advisory Board, the European Commission, and the Swedish Tax Agency as well as conference participants at the Shadow 2011 Münster conference and the 2011 Swedish Economics Meeting in Uppsala 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. Nordblom also acknowledges financial support from Swedish Research Council, project no. 421-2010-1420. Some of the work was done when Ohlsson enjoyed the hospitality of School of Economics, UNSW, Sydney and Department of Economics, University of Melbourne during his sabbatical. Financial support for the sabbatical from the Wenner-Gren Foundations is gratefully acknowledged.
1 Introduction

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

We find, however, that preliminary tax payments are strongly related to tax compliance. This finding is not consistent with neoclassical theory. Reference dependence and loss aversion, as defined 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. In our particular application it is likely that a zero balance in preliminary tax payments is such a reference point.

Loss aversion implies that the individual values losses (in comparison to the reference point) more than gains by the same amount. An individual with a preliminary deficit will, therefore, perceive a higher marginal value of extra income than an individual with a preliminary surplus of the same amount. Those with a deficit would consequently be less inclined to comply. This line of reasoning is put forward in some theoretical studies in the area of prospect theory and tax compliance.\footnote{See, e.g., Yaniv (1999), Bernasconi and Zanardi (2004), and Dhami and al Nowaihi (2007).}

We study 3.6 million Swedish taxpayers in working ages and their behavior when filing their income tax returns. The high quality tax return data concern the income year 2006, the tax assessment year is 2007.

The Tax Agency reports a preliminary tax balance to the taxpayer before the income tax return has to be filed. Generally there is a good correspondence between preliminary and final taxes due, but they may differ. There are several reasons for why such differences arise. One is that actual local tax rates are set with two decimals of a percentage. Preliminary local tax rates used for tax collection at source, on the other hand, are the actual local tax rates rounded to the closest integer percentages tax rates. Preliminary balances, therefore, among other things vary depending on in which municipality the taxpayer lives.

Our hypothesis is that taxpayers view a zero tax balance as a reference point and that taxes due at the time of filing is regarded as a loss. A tax refund is regarded as a gain. We study if the probability that taxpayers claim deductions for “other expenses for earning employment income” depends on the preliminary balance. In many cases, these deductions express tax non-compliance: The Swedish Tax Agency reports that almost
Figure 1: The share claiming a deduction as a function of the preliminary deficit.

all audited claimed deductions for “other expenses for earning employment income” were rejected.²

Figure 1 shows that the probability of claiming a deduction is indeed higher for taxpayers with a preliminary deficit than for those with a preliminary surplus. The share claiming a deduction is, moreover, independent of the preliminary balance for those with a surplus. Slightly less than 4 percent of those with a preliminary surplus claim a deduction for “other expenses”. Each data point in the figure represents the share for taxpayers with a preliminary deficit in a SEK 300 interval.³ The share is, however, increasing in the preliminary deficit for those with a deficit.

We design a quasi-experiment using a regression kink and discontinuity approach. This regression technique allows us to eliminate potential problems of endogeneity and selection in ways that most previous empirical studies have not been able to do. We find behavior to be consistent with loss aversion in the following sense: Taxpayers who have a preliminary deficit (taxes due) are more likely to claim deductions for “other expenses” than those who have a preliminary surplus (tax refund). None of the covariates exhibits a similar evolution around zero preliminary balance. Hence, we can rule out selection and claim that we

²See RSV (2001). A more recent follow-up in 2006 drew a random sample of claimed deductions. There were mistakes in 93 percent of the cases.

³The average SEK/USD exchange rate was SEK 7.38 during 2006.
have found evidence of loss aversion. The major contribution of the paper is thus that we find tax filing behavior to be consistent with loss aversion, which in turn has important policy implications.

We also use an alternative instrumental variable (IV) approach. The difference between the actual and the preliminary tax rates is used as an instrument when estimating probability models for claiming a deduction. The results are confirmed using this approach. Our conclusion is, therefore, that there is indeed evidence of loss aversion.

We also estimate the coefficient of loss aversion in our empirical analysis. The estimate, \( \hat{\lambda} = 2.17 \), for the full sample, is very close to the estimate reported by Tversky and Kahneman (1992), \( \hat{\lambda} = 2.25 \).

We indeed find a causal link from loss aversion to tax filing behavior. Hence, we contribute to the previous literature, which has suggested that people who have paid too little in preliminary taxes are less likely to comply than those who have paid too much, but without ability to establish causality.\(^4\)

One policy implication would be that tax authorities could increase tax revenue by withholding a little bit too much in order to keep people on the surplus side. This would increase compliance, which would reduce the need for audits and could enhance tax morale in society in the long run (Nordblom and Žamac, 2012). However, over-withholding should not be used in a too ambitious way, as that itself may cause credibility problems.

Not only do we come to the tax policy relevant conclusion that the degree of tax compliance could be affected by loss aversion. We also add to the growing literature, where loss aversion has been found in real-life situations outside experimental labs.\(^5\)

The remainder of the paper is organized as follows: We discuss prospect theory and its application to tax compliance in Section 2. Section 3 describes our data and the institutional setting. We then present a simple theoretical model, in which the taxpayers’ decisions are studied, in Section 4. The model provides predictions for the empirical analysis. Some descriptive results are presented in Section 5. Section 6 presents the empirical results using the regression kink and discontinuity approach. A robustness check using an IV-approach and our estimates of the coefficient of loss aversion is presented in Section 7.1. Section 8 concludes

\(^4\)See, e.g., Chang and Schultz Jr (1990), and Persson (2003). There are also experimental studies suggesting that advance payments actually matter for compliance, e.g., Robben et al. (1990), Schepanski and Shearer (1995), and Copeland and Cuccia (2002).

\(^5\)See Camerer (2000) and DellaVigna (2009) for overviews of a number of studies. In the field experiment by Fehr and Goette (2007) a majority of the studied bicycle messengers turn out to be loss averse, which has implications for their behavior. Pope and Schweitzer (2011) find loss aversion when studying the putting behavior of professional golf players on the PGA tour. Also among New York City cab drivers reference dependence has been found in a series of papers, see Crawford and Meng (2011) and references therein. Genesove and Mayer (2001), who study sellers of residential real estate find loss aversion where the purchase price seems to be the relevant reference point.
2 Prospect theory

2.1 Theory and evidence

Kahneman and Tversky (1979) defined prospect theory. The elements of prospect theory we primarily use are reference dependence and loss aversion. Reference dependence implies that people perceive outcomes as gains and losses compared to some reference point rather than caring for final states of total wealth. 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). This implies that the utility function is kinked at the reference point.

Although the concept of loss aversion was first introduced in risky settings (and has been proposed to explain parts of observed risk aversion), Tversky and Kahneman (1991) show that it can also explain behavior in the absence of risk.

The existence of loss aversion has been studied in many experimental settings, both risky and riskless, mostly using student subjects. Kahneman et al. (1990) conducted one of the first famous experiments in a riskless setting. Half of the subjects were given a mug, for which they were then asked to give a selling price. The rest were asked for their willingness to pay for a mug. The price mentioned by the former group widely exceeded the one mentioned by the latter. This “endowment effect” has been seen as a test of loss aversion and reference dependence.

Many experimental studies have followed. Tversky and Kahneman (1992), Schmidt and Traub (2002), Gächter et al. (2007), Abdellaoui et al. (2007), and Abdellaoui et al. (2008) are some studies finding evidence of loss aversion to various extent in within-subject comparisons. They let the subjects (in most cases rather few university students) make several choices and conclude that the same individual makes different choices depending on gains or losses. Camerer (2000) presents and discusses several field experiments that have found behavior to be consistent with both loss aversion and other components of prospect theory.

---

6 The original prospect theory by Kahneman and Tversky (1979) was later refined by Tversky and Kahneman (1992) into the cumulative prospect theory with the inclusion of weighting functions of probabilities. Loss averse behavior, therefore, does not stem solely from the utility function according to cumulative prospect theory. It is also determined by the weights put on the probabilities of different outcomes (Schmidt and Zank, 2005). We deal with riskless decision making in this paper, implicitly assuming original prospect theory.

7 See, e.g., Bateman et al. (1997).
(2003, 2004) finds in his field experiments that there is a pronounced endowment effect (like the one found by Kahneman et al., 1990) among inexperienced subjects, but that experienced subjects behave in line with neoclassical theory. However, Pope and Schweitzer (2011) find that even highly experienced professional golfers exhibit loss aversion. Crawford and Meng (2011) find reference dependence among New York cab drivers in a study where they control for selection in an elegant way. Fehr and Goette (2007) study bicycle messengers and find that the majority is loss-averse, which explains their reduced effort as response to a wage increase.

### 2.2 Tax compliance and prospect theory

Some theoretical studies in the area of prospect theory that focus on tax compliance suggest the following line of reasoning: Loss aversion implies that the individual values losses compared 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 more inclined to take the chance of non-compliance.

Dhami and al Nowaihi (2007) set up a rigorous model. They make a complete analysis to explain some “tax-evasion puzzles” and how these puzzles can be explained by various components in cumulative prospect theory.

Bernasconi and Zanardi (2004) also model tax evasion in the realm of cumulative prospect theory. Their simulations show that people in the loss domain are more prone to evade than people in the gain domain and that they evade larger amounts.

Some papers focus on how important advance tax payments are to deter tax evasion in the light of prospect theory. Yaniv (1999) sets up a simple theoretical model to show how advance payments reduce evasion. The intuition is clear: Those who have paid too much in advance get a refund, which is considered a gain. If the advance tax payments are lower than actual tax liabilities, on the other hand, one has to pay, which is considered a loss. Figure Y refers to a utility function under loss aversion. Since the utility function is concave for gains (implying risk aversion) and convex for losses (implying risk seeking), a taxpayer in the loss domain might be more willing to take the risk of evading. The theoretical reasoning is, therefore, that advance tax payments deter tax evasion.
Elffers and Hessing (1997) also find that advance payments promote compliance. They, however, also point to the fact that withholding too much may make people feel wrongly treated, which has an opposing effect. They also claim that a standard deduction would increase tax compliance.

Experimental studies also suggest that advance tax payments actually matter for compliance. Kirchler and Maciejovsky (2001) use data on self-reported tax evasion to explain compliance behavior in an experimental setting. They find evidence of both reference dependence and loss aversion. Similar results are found in the experiments by Schepanski and Shearer (1995), Robben et al. (1990), and Copeland and Cuccia (2002).

Previous studies using real-world data suggest that people who have paid too little in preliminary taxes actually are less likely to comply than those who have paid too much. Cox and Plumley (1988) find that the share of tax returns needing correction increases with taxes due. Chang and Schultz Jr (1990) find for the United States that those who owe additional taxes when they file their returns are less likely to comply than those who have refunds due. None of these studies can, however, establish a causal relationship.

Persson (2003) is a descriptive study using Swedish data on actual taxpayers. She reports that people with tax deficits are more prone to claim dubious deductions but she cannot establish causality. In one region, all claimed deductions were audited. Those with tax deficits were less likely to have their claimed deductions approved than those with tax surpluses, which suggests non-compliance.

3 Data and institutional setting

Our entire dataset covers all 4.7 million Swedes, 16–67 years old, who had employment income, but were not self-employed, 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.

The decision we study in this paper is whether or not the taxpayer claims a deduction for “other expenses for earning employment income” while filing their tax return. Filing the tax return is one event in the annual sequence of taxpaying.

The sequence of events for a taxpayer is as follows: Employers withdraw preliminary taxes before the

---

8 See also the information from the Tax Agency reproduced in Appendix A.
salary is paid to the employees. The Tax Agency decides on tax tables that employers use when withdrawing tax payments that are transferred to the Tax Agency during the income year. Hence, typically, taxpaying is perceived as automatic for Swedish taxpayers, without the need to take any action of their own.

In April the following year, the Tax Agency sends a preliminary income tax return to the taxpayer. This tax return is based on the statements of income that the Tax Agency has received from employers, banks, etc. Virtually all incomes are specified on the tax return form, due to the legislated reporting system. Moreover, the Tax Agency calculates a preliminary balance in taxes due (actual tax liability – preliminary payments), which is also written on the tax return. Since preliminary taxes are withdrawn according to tax tables on all income, the preliminary balance is close to zero for most taxpayers. The preliminary balance may, however, show a deficit (more taxes due) or a surplus (there will be a tax refund). When filing their tax return, the taxpayer may add missing information. Because of the extensive reporting from employers etc., there is usually not much information to add, but it could for instance be claiming deductions for “other expenses for earning employment income”.

The filing must be done by 2 May and after that the Tax Agency calculates the final tax for the taxpayer and depending on the preliminary tax payments, it may be a surplus or a deficit. If there is a surplus after the filing, refunds are paid out in June or August. If there is a deficit, the taxpayer has to pay the taxes due in November or December. For the taxpayer there is almost no administrative cost in either case. If there is a refund, one should specify a bank account and the refund is automatically paid out. With a deficit, the taxpayer receives a payment slip from the Tax Agency.

It is a major event in Sweden when the Tax Agency starts sending out the final tax statements. Media often fuels the sense of a giant mandatory lottery taking place. An illustrative example is given by an article in the tabloid Aftonbladet that gave notice to the event. The preamble exclaims: “Praise mammon, here comes the tax refunds” and continues: “If you are lucky the money could be on your account before Midsummer”. One could in some sense describe the event as Christmas for grownups–but there is always a risk that it is not Santa but the tax collector that is coming down your chimney. Others would say that this a

---

9 Analogous principles apply for capital income taxation (interest received and paid, dividends, etc). As Sweden has a dual income tax system, all capital incomes are taxed at source with 30 percent.

10 The interest paid on surpluses and the interest levied on deficits are independent of the filing date. However, for deficits smaller than SEK 20,000, no interest had to be paid if the taxes due were paid by 4 May at the latest. Appendix A presents the details that applied during the assessment year 2007. The issues that Slemrod et al. (1997) and Jones (2012) discuss for the United States do therefore not arise in the same way in Sweden.

11 The article is from 22 March 2011 and was published in the smart-phone version of the tabloid under the name “Tax refund? You can now log in and check!”, our translation.
clear example of mental accounting.

The Swedish income tax system is a dual one. Capital income is taxed at a flat national rate of 30 percent, while employment income is taxed progressively and by both municipalities (and counties) and the central government. Most Swedes only pay the local income tax, which during 2006 ranged between 28.8 and 34.24 percent, with an average of 31.6 percent. The local income tax consist of two parts – the municipality tax and the county tax. There was also a personal exemption of between 11 and 30 thousand SEK, depending on income, but the same in all localities. The reason for the exemption to change with income was to reduce the marginal effects from reduced means tested benefits when income increased. This makes it hard for a single individual with varying income to fully predict the value of this personal exemption. High-income earners paid an additional 20 percent in central income tax on the part of their income exceeding SEK 317,700 and another 5 percent on the amount exceeding SEK 472,300. Figure 2 shows the marginal tax structure in 2006 for a taxpayer who met the average local tax rate (source: Swedish Tax Agency). Hence, tax rates vary with income due to the progressive income-tax system. However, it also varies across localities as people with the same income face different marginal tax rates in different municipalities.

Final taxes are determined by the income earned during the full year. However, preliminary taxes are paid monthly by employers and they are decided as if the employee had the same salary every month. Moreover, the tax is withdrawn as an integer percentage. Hence, an individual whose income is exactly the same every month and who has a true local tax rate strictly equal to an integer (for example, 31 percent) and who has no loans or other issues that could affect taxation, would end up paying exactly the true tax liability in preliminary taxes and would therefore have a zero tax deficit when it is time to file the tax return.

Deviations from zero are in most cases small and exogenous to the taxpayers. So why may preliminary tax balances deviate from zero at all? The way the local government part of the employment income tax is levied is an important reason.

The actual tax rates are set with two decimals of a percentage (for example, 30.83 percent or 31.12 percent). The tax tables are, however, created assuming integer percentages tax rates at source (for example, 31 percent). Taxpayers who live in municipalities for which the tabled tax rates at source are above the actual rates will tend to pay too much in preliminary taxes. This will increase their preliminary balances. The opposite applies for taxpayers who live in municipalities with actual rates higher than the rates used at source. Figure 3 shows that a higher share of those having the preliminary tax rate rounded downwards have
Figure 2: Marginal tax rates in Sweden during 2006 (Source: Swedish Tax Agency).
preliminary deficits than those having the preliminary tax rate rounded upwards.

The local tax rate is the result of decisions in two separate jurisdictions, the municipality and the county council (covering several municipalities). In 2006 there were 290 municipalities in 20 counties. It would, therefore, be extremely difficult for a single municipality to set the local tax rate strategically trying to use the rounding principles for taxation at source. This would require to correctly predict the behavior and response of the county council. It would be even more difficult for the county council to act strategically in this respect. This would involve predicting the behavior of several municipalities. It would, consequently, also be very difficult for a taxpayer to predict where to move to take advantage of a lower preliminary tax rate. Hence, it should be completely innocuous to assume that the actual tax rate is truly exogenous to the individual taxpayer.

The progressive central government part of the employment income tax and the personal exemption, that depended on employment income, are other possible reasons why the balance may deviate from zero. When preliminary taxes are withdrawn by the employer, it is done according to tax tables, which assume that the income stays the same every month. If income varies a lot during the year, the tax exemption

Figure 3: Relative frequency distributions for taxpayers with preliminary tax rates rounded upwards and downwards.
may be misjudged and preliminary taxes may be higher or lower than final taxes. For taxpayers close to the threshold between the tax brackets, e.g., the threshold for the central income-tax, unanticipated events may create a non-zero balance. Benefits from the social insurance systems are included in employment income but replacement rates are below 100 percent. Absenteeism may therefore affect the preliminary balance and may affect overtime work. A taxpayer with varying employment income over the year will tend not to pay the same amount in preliminary taxes as a taxpayer with the same annual employment income evenly distributed over the year. Having several employers might also result in that not enough taxes will be paid at source. Every employer may withdraw taxes as if they were the only employer, which results in too little taxes due to progressivity.

Deviations from a zero preliminary balance may also arise from capital income taxation. Capital income is taxed at 30 percent at source, but interest on loans are deductible and not automatically taken into account in the preliminary taxes. Taxpayers who pay deductible interest may therefore have preliminary surpluses. Unfortunately, we do not know how much loan each individual has, so we cannot directly control for this in our analysis. This may therefore be an important source of heterogeneity in the initial balance and may explain why a majority start off at the surplus side. Since the vast majority have loans with floating rate it should be almost impossible for individuals to make a precise prediction of how much the deductible interests should affect the total tax liability.

In our analysis we want to see how normal taxpayers behave when they face an exogenously determined preliminary balance. We therefore need to make some restrictions and exclude some taxpayers from our sample. How this is done is presented in the first three columns of Table 1.

<table>
<thead>
<tr>
<th>condition</th>
<th>N</th>
<th>ΔN</th>
<th>percentage drop (marginal)</th>
<th>percentage drop (cumulative)</th>
</tr>
</thead>
<tbody>
<tr>
<td>original sample</td>
<td>4 674 003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no advance payments</td>
<td>4 363 909</td>
<td>-310 094</td>
<td>6.63</td>
<td>6.63</td>
</tr>
<tr>
<td>no missing values for initial deficit</td>
<td>4 363 628</td>
<td>-281</td>
<td>0.01</td>
<td>6.64</td>
</tr>
<tr>
<td>wage income &gt; SEK 100 000</td>
<td>3 627 676</td>
<td>-735 952</td>
<td>16.87</td>
<td>22.39</td>
</tr>
<tr>
<td>wage income &lt; SEK 1000 000</td>
<td>3 610 972</td>
<td>-16 704</td>
<td>0.46</td>
<td>22.74</td>
</tr>
</tbody>
</table>

Our original sample consists of 4.67 million employed taxpayers. We disregard self employed people as they can affect their tax payments to a large extent.

From the original sample, we exclude taxpayers who actively have made advance tax payments directly to the Tax Agency at their own initiative in the beginning of the assessment year, since this may not be
independent from future deductions.

For a few taxpayers, the initial deficit is missing, and since that is crucial to our analysis, we exclude them from the sample.

Finally, only taxpayers with "normal" annual employment incomes are included in our sample. We interpret normal annual employment income to be in the interval SEK 100,000–1,000,000. These selection criteria leave us with a full sample of 3.6 million taxpayers. Descriptive statistics for the full sample are presented in Table 2.

Table 2: Descriptive statistics.

<table>
<thead>
<tr>
<th></th>
<th>Full sample</th>
<th>Maximum bandwidth sample*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>preliminary deficit</td>
<td>preliminary surplus</td>
</tr>
<tr>
<td>number of observations</td>
<td>810,690</td>
<td>2,800,282</td>
</tr>
<tr>
<td>deducting, fraction</td>
<td>0.062</td>
<td>0.044</td>
</tr>
<tr>
<td>deduction, SEK, conditional</td>
<td>4,716</td>
<td>(4,324)</td>
</tr>
<tr>
<td>preliminary balance, SEK, unweighted</td>
<td>-10,757</td>
<td>(51,232)</td>
</tr>
<tr>
<td>employment income, SEK</td>
<td>279,770</td>
<td>(135,523)</td>
</tr>
<tr>
<td>men, fraction</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>age, years</td>
<td>46.8</td>
<td>(12.5)</td>
</tr>
</tbody>
</table>

* Sample selection criterion: weighted preliminary balance in the interval ± SEK 3,000.

Note. mean (standard deviation)

These 3.6 million taxpayers are studied with respect to their deduction behavior. Assessed employment income was reduced by approved deductions for “other expenses for earning employment income” exceeding SEK 1,000 during the income year 2006. The expenses thus did not have to be large and and the specifications were quite vague. Hence, there was room for (small scale) tax non-compliance by the

---

12 This corresponds to an annual income of USD 13,550–135,000 in 2006. Less than 0.5 percent of our sample had incomes exceeding SEK 1,000,000. On the other hand, about 1/6 had an income lower than SEK 100,000, but we have to exclude them for technical reasons. A large share of these low-income earners are students who are working extra.

13 Some examples that could be approved 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
means of this deduction.

There are some studies of randomly audited claimed deductions for “other expenses”.\textsuperscript{14} The finding is that 90–95 percent of the claims were not approved. In some cases, the taxpayer may simply not understand the rules and believes that he is entitled to the deduction. The taxpayer may take a chance in other cases. We cannot, therefore, be sure whether it is tax evasion, tax avoidance, or something else going on.

Some claimed deductions are, however, clearly not due to non-compliance. This is likely to be true especially for large claimed deductions. It is difficult to invent very large “other expenses”. Large deductions are more often accurate than small (Persson, 2003). This is an important reason for only including relatively small deductions for “other expenses”. Hence, we concentrate on deductions smaller than SEK 20,000.

It was almost without risk to claim small incorrect deductions in 2007. The audit probability was very low for small claimed amounts. Moreover, the taxpayer had only to pay the increased tax liability if the claimed deduction was audited and not approved. There were no fines for incorrectly claimed deductions of small amounts.\textsuperscript{15} Neither were there any fines for large rejected deductions provided that the taxpayer could prove having had the expenses for which the deduction was claimed. The choice to claim a deduction was, therefore, completely without risk.

To the 3.6 million taxpayers in our full sample, it should be exogenous whether one ends up having a preliminary surplus or a preliminary deficit. The closer the preliminary balance is to zero, the more likely it should be that it is truly random whether one ends up with a surplus or a deficit. We, therefore, focus on taxpayers with a preliminary balance in the interval SEK $\pm$3,000. There are 1.2 million in this sub-sample that we label the \textit{maximum bandwidth sample}.\textsuperscript{16} Our approach is that the deduction behavior of these taxpayers can be viewed as a quasi-experiment.

Just from the descriptive statistics, we notice that a larger fraction with a preliminary deficit than with a preliminary surplus claim the deduction, both in the full and in the maximum bandwidth sample. We also see that, conditional on claiming a deduction, taxpayers with a preliminary deficit claim larger deductions in general), expenses for books and journals related to work for some occupations if not paid for by the employer, and expenses for phone calls related to work if not paid for by the employer (not the phone and not the subscription).

\textsuperscript{14}See RSV (2001) and Persson (2003).

\textsuperscript{15}The tax law has changed since. The threshold value that expenses have to exceed before the tax liability decreases was increased considerably. It is no longer possible to claim deductions of such small amounts.

\textsuperscript{16}Descriptive statistics for the maximum bandwidth sample are also presented in Table 2. This sample is based on the preliminary deficit being weighted by the taxpayer’s employment income for reasons that will be explained in Section 5. Weighting increases the probability of including taxpayers with high employment income in the sample. Employment income and deduction probability are positively correlated. This is the reason why the deduction probabilities reported in Table 2 in Appendix B are slightly higher than those shown in Figure 1.
than those with a preliminary surplus.

4 An illustrative theoretical model

The purpose of this section is to provide predictions for our empirical exercises. We apply loss aversion and reference dependence to a simple model of decision making.

Consider a taxpayer $i$ who is about to file his tax return. He receives information about his preliminary balance in taxes on the tax return. Taxpayer $i$ has taxes due if he has a preliminary deficit, $D_i > 0$, and will receive a refund if he has a preliminary surplus, $D_i \leq 0$. He compares the preliminary deficit to his reference point.

What should be the valid reference point has been widely discussed.\(^{17}\) 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 (Közegi and Rabin, 2006), i.e., how much the taxpayer expects to owe or to get refunded. Schepanski and Shearer (1995), however, find that the current asset position is more relevant as the reference point for taxpayers than the expected asset position. Hence, like in many other areas, status quo seems to be a natural reference point also when filing tax returns. Also Elffers and Hessing (1997) and Yaniv (1999) argue for status quo as the reference point when analyzing tax compliance. Therefore, we assume that a zero balance is the reference point for the taxpayer when filing their return.\(^{18}\)

If the preliminary balance is a deficit, i.e., $D_i > 0$ the taxpayer experiences himself to be in the loss domain. He will, on the other hand, be in the gain domain if he gets a tax refund, i.e., $D_i \leq 0$.

The value function with reference dependence and loss aversion is:

$$
V(D_i) = \begin{cases} 
-vD_i & \text{if } D_i \leq 0, \\
-\lambda vD_i & \text{if } D_i > 0,
\end{cases}
$$

where $\lambda > 1$ is the coefficient of loss aversion.

\(^{17}\)See, e.g., the discussion in Kirchler and Maciejovsky (2001).

\(^{18}\)One could argue that since tax filing is a reoccurring event, previous preliminary balances may be the relevant reference point. In our data, we have unfortunately no information on previous years and our theoretical model is therefore also a one-period model, so in this paper that is a ruled-out possibility. In Section 6.2 we, however, present sensitivity analyses allowing the reference point to be endogenous.
We follow, e.g., Benartzi and Thaler (1995), Schmidt and Traub (2002), and Pope and Schweitzer (2011) and assume a linear value function. This implies constant marginal values. We disregard strict concavity in the gain domain and strict convexity in the loss domain by this assumption. Utility tends, however, to be almost linear on very small intervals. A linear approximation should therefore work well since we limit our analysis to a narrow area around zero.

4.1 Whether to deduct or not

A first step is to predict the probability of claiming a deduction to reduce the tax liability.

The taxpayer’s only choice is whether or not to claim a deduction of a fixed amount, $\delta$, the same for all taxpayers and we analyze what affects the deduction probability. The amount $\delta$ is sufficiently small to ensure that the taxpayer perceives to be at no risk of being audited if he claims the deduction.\footnote{We abstract from detection risks in our model contrary to previous studies of tax compliance. Our main results, however, remain valid even if we include a risky choice of non-compliance.}

Claiming the deduction $\delta$ comes at a certain cost, $c_i$. This cost varies across taxpayers, $c_i \sim U[0, \tilde{c}]$. 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.

The taxpayer compares the value of his preliminary tax balance, $V(-D_i)$, to the value if he claims the deduction, $V(t\delta - D_i) - c_i$. He claims the deduction if the latter exceeds the former. There are three different domains where the taxpayer may end up depending on the sign and size of $D_i$. The three conditions for claiming the deduction $\delta$ are:

A: \[ c_i < vt\delta \] if $D_i \leq 0$,

B: \[ c_i < v[t\delta + D_i(\lambda - 1)] \] if $D_i \in (0, t\delta]$, (2)

C: \[ c_i < \lambda vt\delta \] if $D_i > t\delta$.

The value of the deduction in the three cases above are illustrated in Figure 4.

The assumption of a uniformly distributed cost makes it straightforward to use the conditions (2) to predict the share of taxpayers deducting at various preliminary deficits. C.p., the higher the tax rate, the larger the share of deducting taxpayers. The threshold cost for claiming the deduction in the loss domain is
lower than in the gain domain as \( \lambda > 1 \). A larger share of the taxpayers should, therefore, claim \( \delta \) in the loss domain than in the gain domain. For those with a positive balance, the share of taxpayers who deduct should be independent of the preliminary balance (because the value function is linear). The same applies for those with a large preliminary deficit (where \( D_i > t\delta \)) who still will have a deficit even after the deduction. It is only in the middle group, \( D_i \in (0, t\delta] \), where the share of taxpayers deducting increases in the preliminary deficit:

\[
\Delta = v [t \delta + D_i (\lambda - 1)],
\]

\[
\frac{\partial \Delta}{\partial D_i} = v (\lambda - 1) > 0.
\]

We then predict the pattern for the share of taxpayers claiming the deduction to be as shown in Figure 5.
The pattern in Figure 5 can be summarized as follows:

**Prediction 1.** The share of taxpayers with a preliminary deficit who will claim the deduction $\delta$ is higher than the share of taxpayers with a preliminary surplus who will.

**Prediction 2.** The probability of claiming the deduction $\delta$, as a function of the preliminary deficit, has the shape as in Figure 5 with kinks at 0 and at $D = t\delta$.

How important is loss aversion in terms of magnitude? Loss aversion implies that the utility (or value) function is steeper for gains than for losses, but how much steeper? The coefficient of loss aversion, $\lambda$, should capture this. There is no overall consensus on how $\lambda$ should be measured: Tversky and Kahneman (1992) originally defined it as the ratio of utilities: $\lambda = -U(-1)/U(1)$. Köbberling and Wakker (2005) instead propose the following definition, which is independent of unit of payment: $\lambda = U'_1(0)/U'_L(0)$, which was also informally suggested by Benartzi and Thaler (1995). The two definitions are equivalent when the value function is linear. Abdellaoui et al. (2007) summarize and discuss various methods of measuring $\lambda$.

Our approach is different. We observe a large number of people in the gain domain and in the loss domain. It is, therefore, possible to estimate an aggregate coefficient of loss aversion. We plot the shares claiming the deduction for each level of preliminary deficits as sketched in Figure 5. The share claiming the deduction in the gain domain is:

$$\int_{c=0}^{vt\delta} f(c) dc = X.$$  \hspace{1cm} (4)

We observe the actual share, $X$, in our data. Since the cost is assumed to be uniformly distributed, (4) is easily solved to yield:

$$\frac{vt\delta}{c} = X.$$  \hspace{1cm} (5)

The share claiming the deduction in the part of the loss domain where $D_i > t\delta$ is:

$$\int_{c=0}^{\lambda vt\delta} f(c) dc = Y \Rightarrow \frac{\lambda vt\delta}{c} = Y.$$  \hspace{1cm} (6)

Equations (5) and (6) then gives $\lambda = Y/X$. Hence, the coefficient of loss aversion is simply the ratio between the two shares claiming the deduction.
**Prediction 3.** If \( c_i \sim U[0, \hat{c}] \), 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 part of the loss domain where \( D_i > t\delta \) and \( X \) is the share deducting in the gain domain.

We will use this prediction to estimate the coefficient of loss aversion, \( \lambda \), in Subsection 7.2.

4.2 The amount deducted

We will now drop the assumption of a constant deducted amount and instead allow the deducted amount to vary.

Previously, we assumed a certain cost \( c_i \) associated with claiming the given deducted amount. We now instead assume a convex cost function, \( c_i(\delta_i) = c_i\delta_i^2 + \varsigma_{\delta>0} \), where \( c_i \sim U[1, \hat{c}] \) and \( \varsigma \) is a constant, the same to everyone. We continue to assume a linear value function, which should be a good approximation in the neighborhood where \( |D_i^p| \) is close to zero.

Regarding deducted amounts, we could think of four qualitatively different solutions:

A: The taxpayer is in the gain domain already from start, i.e., \( D_i \leq 0 \)

B: The taxpayer is in the loss domain from start, but claims a sufficiently large deduction so as to end up in the gain domain afterwards, i.e., \( D_i \leq t\delta_i \)

C: The taxpayer starts out in the loss domain and makes a deduction which exactly brings him to the reference point, i.e., \( D_i = t\delta_i \)

D: The taxpayer remains in the loss domain also after the deduction, i.e., \( D_i \geq t\delta_i \)

The upper part of Figure X shows at what combinations of \( D_i \) and \( c_i \) the taxpayer ends up in which segment.

Figure X around here

A)
Let us start with a taxpayer who has a negative deficit and therefore is in the gain domain already before any deduction ($D_i \leq 0$). The value function if deducting is then

$$V_i = v(t\delta - D_i) - c_i\delta_i^2 - \varsigma|\delta > 0,$$

Maximizing this with respect to $\delta_i$ yields the interior solution $\delta_{iA} = \frac{vt}{2c_i}$ whenever

$$c_i \leq \frac{v^2t^2}{4\varsigma}.$$  

Hence,

$$\delta_{iA} = \frac{vt}{2c_i} \Rightarrow \delta_{iA} \in \left[\frac{2\varsigma}{vt}, \frac{vt}{2}\right].$$  

**B)**

In **B)**, the taxpayer will end up in the gain domain after the deduction, i.e., $D_i \in (0, t\delta_i]$. The maximization at the margin will therefore be the same as in (7) and results in the same $\delta_{iB} = \delta_{iA}$. This means that this solution will occur for initial deficits $D_i \leq t\delta_{iB} = \frac{vt^2}{2c_i}$.

The requirement for an interior solution is that $-\lambda vD_i < v(t\delta_{iB} - D_i) - c_i\delta_{iB}^2 - \varsigma$, which in turn requires that

$$c_i \leq \frac{v^2t^2}{4(\varsigma - vD_i(\lambda - 1))}.$$  

Comparing (8) and (10) we note that the breakpoint cost will differ since the marginal value of a deduction is higher in the loss than in the gain domain. Since $\lambda > 1$ and $D_i > 0$, $\tilde{c}_{iB} = \frac{v^2t^2}{4(\varsigma - vD_i(\lambda - 1))} > \tilde{c}_{iA} = \frac{v^2t^2}{4\varsigma}$. Hence, there is a wider range of deducted amounts in **B)** than in **A)**:

$$\delta_{iB} = \delta_{iA} = \frac{vt}{2c_i} \Rightarrow \delta_{iB} \in \left[\frac{2\varsigma}{vt(2\lambda - 1)}, \frac{vt}{2}\right].$$  

That the breakpoint for an interior solution $\tilde{c}_{iB} > \tilde{c}_{iA}$ guarantees that the probability of claiming a deduction will kink upwards at $D_i = 0$ if $f_c(c) = \frac{dF_c(c)}{dc}$ has support at $c = \tilde{c}_{iA}$, i.e. given $f_c(\tilde{c}_{iA}) > 0$.

**Proof**

The probability of claiming a deduction is given by $F_c(\tilde{c}(D))$. A positive kink at $D = 0$ thus formally
means

\[
\frac{dF_c(\tilde{c}(D))}{dD}\bigg|_{D=0^+} > \frac{dF_c(\tilde{c}(D))}{dD}\bigg|_{D=0^-} \iff \\
\frac{\delta_c(\tilde{c}(A))}{dD}\bigg|_{D=0^+} > \frac{\delta_c(\tilde{c}(D))}{dD}\bigg|_{D=0^-} \iff \delta_c(\tilde{c}(A)) > 0
\]

Note that at \( D = 0^- \) the A) situation applies and at \( D = 0^+ \) the B) type situation applies (since any positive deduction will take the individual to the surplus side). From (8) and (10) we get

\[
\frac{d\tilde{c}(D)}{dD}\bigg|_{D=0^-} = 0 \\
\frac{d\tilde{c}(D)}{dD}\bigg|_{D=0^+} = -\frac{t^2v^2}{4(\varsigma - Dv(\lambda - 1))^2}(-v(\lambda - 1)) > 0
\]

QED.

Since \( \delta_{iA} = \delta_{iB} \forall \ c_i \leq \frac{\nu^2t^2}{4\varsigma} \) and since \( \frac{d\tilde{c}(D)}{dD}\bigg|_{D=0^+} > \frac{d\tilde{c}(D)}{dD}\bigg|_{D=0^-} \), we get the following prediction:

**Prediction 4.** There is a positive kink in the unconditional amount at the 0 reference point.

In D), the taxpayer remains in the loss domain also after the deduction, i.e., \( D_i > t\delta_i \). The taxpayer maximizes

\[
V_i = \lambda v(t\delta - D_i^p) - c_i\delta_i^2 - \varsigma|_{\delta>0},
\]

with respect to \( \delta_i \). The interior solution gives \( \delta_{iD} = \frac{\lambda vt}{2c_i} \).

The condition for an interior solution where the taxpayer finds it worthwhile to claim the deduction is

\[
c_i \leq \frac{\lambda^2v^2t^2}{4\varsigma},
\]

which in turn implies that

\[
\delta_{iD} = \frac{\lambda vt}{2c_i} \Rightarrow \delta_{iD} \in \left[ \frac{2\varsigma}{\lambda vt}, \frac{\lambda vt}{2} \right].
\]

For those who end up in either of these three solutions, A), B), or C), the amount deducted is independent of \( D_i^p \). This is an artefact of the linear value function. The individual heterogeneity instead comes from variations in the cost function where the deduction is inversely proportional to \( c_i \).
In $B$, where $D_i \leq \frac{vt^2}{2c_i}$ the taxpayer ends up in an interior solution where $\delta_{iB} = \frac{vt}{2c_i}$ and in $D$, where $D_i > \frac{\lambda vt^2}{2c_i}$ the taxpayer ends up in an interior solution where $\delta_{iD} = \frac{\lambda vt}{2c_i}$. The taxpayers where $D^p_i \in \left[\frac{vt^2}{2c_i}, \frac{\lambda vt^2}{2c_i}\right]$, will instead end up in the "bunching" corner solution, $C$, where the deduction is $\delta_{iC} = \frac{D_i}{t}$, i.e., they reach exactly the reference point by claiming the deduction.20

For the taxpayer to find it worthwhile to claim the deduction, total gain needs to exceed total cost:

$$-c_i \delta^2_{iC} - \zeta \geq -\lambda v D^p_i \Rightarrow c_i \leq \frac{t^2(\lambda v D^p_i - \zeta)}{D_i^2}.$$ (15)

This implies that $\delta_{iC} = \frac{D^p_i}{t} \in [\delta_{iB}, \delta_{iD}]$ for any given $c_i$.

Proof:

The marginal cost of claiming $\delta_{iC}$ is $MC_{iC} = 2c_i \delta_{iC}$, which is an increasing function of $\delta$. $\delta_{iC} = \frac{D_i}{t} \Rightarrow MC_{iC} = \frac{2c_i D_i}{2}$ and since $D_i \in \left[\frac{vt^2}{2c_i}, \frac{\lambda vt^2}{2c_i}\right]$ this implies that $MC_{iC} \in [vt, \lambda vt]$.

Assume $\delta_i < \frac{D_i}{t}$. Then the taxpayer will remain in the loss domain where the marginal benefit would be $\lambda vt > MC_{iC}$. Hence, $\delta_i < \frac{D_i}{t}$ cannot be optimal. Instead assume $\delta_i > \frac{D_i}{t}$. Then the taxpayer will enter the gain domain where the marginal benefit is $vt > MC_{iC}$. Hence, $\delta_i < \frac{D_i}{t}$ cannot be optimal either. Hence, $\delta_{iC} = \frac{D_i}{t}$.

The upper part of Figure X shows at which combinations of $D_i$ and $c_i$, the taxpayer ends up in which solution. The lower part shows the deducted amount as a function of $c_i$ in the interior solutions (in the "bunching solution" the amounts are not direct functions of $c_i$, but still $\delta_c \in [\delta_B, \delta_D]$. For a given preliminary deficit an individual with a low cost will deduct an amount sufficiently large to push him into the gain domain, while a person with a high cost will make a smaller deduction and remain in the loss domain also after the deduction. Those with a cost in between will make a "bunching" deduction so that they exactly reach the reference point.

Likewise, those with low costs ($c_i \leq \frac{vt^2}{4c}$) will always deduct, but the amount deducted depends on $D_i$. They claim weakly larger amounts in the loss than in the gain domain.21 Those with $c_i \in \left(\frac{vt^2}{4c}, \frac{\lambda vt^2}{4c}\right]$ only claim a deduction if the preliminary deficit is positive and large enough and those whose $c_i > \frac{\lambda vt^2}{4c}$ never deduct.

20Note that at the limit where $D_i = \frac{vt^2}{2c_i}$ then $\delta_{iB} = \frac{vt}{2c_i} = \frac{D_i}{t}$ and where $D_i = \frac{\lambda vt^2}{2c_i}$ then $\delta_{iD} = \frac{\lambda vt}{2c_i} = \frac{D_i}{t}$.

21In $B$) exactly the same, in $C$) weakly larger (depending on $D_i$) and in $D$) strictly larger
Hence, every single taxpayer would claim a weakly larger deduction in the loss than in the gain domain. From this and from the above proven fact that the probability of making a positive deduction ($F_c(\tilde{c})$) will kink upwards at $D = 0$, given that $f_c(\tilde{c}_A) > 0$ we get the following prediction:

**Prediction 5.** *The unconditional deductions in the loss domain exceed those in the gain domain.*

Concerning conditional deductions, i.e., the amounts deducted conditional on claiming a positive deduction, we cannot come up with an unambiguous prediction to whether they should be smaller or larger in the loss than in the gain domain. On the one hand, for a given, sufficiently low $c_i$, the individual would claim a larger deduction if ending up in the loss than in the gain domain. However, those with a sufficiently high $c_i$ only deduct in the loss domain, and the higher the cost, the smaller the deduction. Hence, there are two counteracting effects concerning the conditional amounts. Both from conditions (9) and (14) and from the lower panel in Figure X, it is obvious that deductions in the loss domain range over a larger span than those in the gain domain. In which domain conditional deductions are the largest depends on the relative parameter values, how large deficits we consider and how strong the degree of loss aversion is. For a given cost function, $c_i$, the conditional amount should kink, not at zero, but in the loss domain where $D_i = \frac{vt^2}{2c_i}$, but since $c_i$ varies across the population, there would be no kink on the aggregate level. Hence,

**Prediction 6.** *The conditional deductions should not kink at the reference point and we cannot generally say whether the conditional deductions are greater in the loss or in the gain domain.*

5 Descriptives

Theory predicts that the share claiming deductions, under certain assumptions, could be illustrated as in Figure 5. We notice that the actual behavior presented in Figure 1 is very similar to what theory predicts. A higher share of those with a preliminary deficit claim a deduction for “other expenses” than those with a preliminary surplus. The share claiming a deduction is independent of the preliminary balance for those with a surplus as predicted by theory. Less than 4.5 percent of those with a preliminary surplus claim a deduction for “other expenses”. The share is, however, increasing in the preliminary deficit for those with a deficit.\(^{22}\)

\(^{22}\)As this kind of deduction was more or less risk less, the deducting share may seem surprisingly low. However, in an international comparison, Swedes have a high degree of tax morale and would not claim something untruthful when the potential gain is only a couple of thousand SEK.
Moreover, theory predicts that more will be deducted in the loss than in the gain domain and that the amounts are more spread in the loss than in the gain domain. From Table 2, we learn that the mean deduction in the loss domain exceeds that in the gain domain. Figure 6 shows the frequency distribution of deductions and it is obvious that many more deductions are claimed by taxpayers who have a deficit than by those who have a surplus. Most deductions are fairly small and those in the gain domain (the dotted line) are more concentrated than those in the loss domain (the solid line).

![Figure 6: Frequency distribution of deducted amounts.](image)

Although the observed pattern is in line with theory, we cannot be certain that it is caused by the preliminary balance; there may be an endogeneity problem. Plotting the distribution of individuals over preliminary deficits is a simple graphical test of this: If the distribution changes around zero we might have problems. Figure 7 shows that the distribution does not seem to kink or jump at zero, so at first inspection there is no indication of any selection or endogeneity problem.

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 look closely at how the covariates evolve around the reference point. Suppose that there is selection based on any of the covariates or selection based on any unobservable factor that is correlated with the covariates. This would show up as a kink or discontinuity around zero. The pattern in Figure 1 could then be due to selection. If, on the other hand, all covariates evolve smoothly around zero it suggests that the deduction pattern is caused by differences in the preliminary balance.

We will here present descriptive tests if the predetermined covariates show kinks or discontinuities at zero preliminary balance. The formal econometric tests are presented in Section 6.

Employment income is a crucial variable. The (blue) diamonds in Figure 8 show the relationship between 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 tax payments at source correctly—the deviations will be scaled up in proportion to the income. Higher employment income will, therefore, move us further away from zero preliminary balance. This is exactly the pattern we see in the unweighted version of the relationship.

The unweighted 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 taxpayer and a low-income taxpayer? It can be argued that the marginal utility of claiming a deduction is higher for a low-income taxpayer. The difference between preliminary and actual local tax rates will also tend to make preliminary deficits and surpluses larger for high-income taxpayers than for low-income taxpayers. There are, therefore, good reasons to instead measure the preliminary deficit weighted by employment income.

Define the weighted preliminary deficit of taxpayer $i$ as:

$$d^p_i \equiv \left( \bar{E}/E_i \right) \ast D^p_i,$$

where $d^p_i$ is the weighted preliminary deficit of taxpayer $i$, $D^p_i$ is the unweighted preliminary deficit of taxpayer $i$, $E_i$ is employment income of taxpayer $i$, and $\bar{E}$ is average employment income. The weight for the individual taxpayer, $\bar{E}/E_i$, is the inverse of the taxpayer’s employment income relative to average employment income. This will move those with employment income above the average closer to the reference point. Those with below-average employment income will be moved away from the reference point. The

Figure 8: Employment income as a function of the preliminary deficit.
kink at zero preliminary deficit disappears when we measure the preliminary deficit weighted by employment income, see the (red) squares 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. It does not matter for this result whether we use unweighted or weighted preliminary deficits.

If we weight our covariates, we should also weight our dependent variable, i.e., the probability of claiming a deduction. In Figure 11, we have added, to the (blue) diamonds from Figure 1, (red) squares showing the weighted version. The patterns are similar. Hence, from the graphs we conclude that the deduction probability changes a lot at the point of zero deficit, while none of the covariates show a similar pattern.

This suggests that it is actually the deficit itself which causes the behavioral difference. However, this is not a formal test, so we continue to the next section, where use econometric methods to formally test causality.

As discussed in Section 4, a zero balance seems a natural reference point. Getting a refund or having to pay is a reasonably salient difference that is likely to matter to people.
Figure 9: Gender as a function of the preliminary deficit.

Figure 10: Age as a function of the preliminary deficit.
Figure 11: The share claiming a deduction as a function of the preliminary deficit.
6 Estimations

6.1 Baseline models

In this section we formally test whether the observed impact of the preliminary deficit on the probability of claiming deductions is causal and not 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):

\[
\Delta_i = \alpha_0 + \alpha_1 I_i + \gamma_1 d_i^p + \beta_1 I_i d_i^p + \epsilon_i,
\]

where \( \Delta_i \) is an indicator of claiming a deduction, \( I_i \) is an indicator for a positive preliminary deficit, \( d_i^p \) is the (weighted) 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 a possible kink at the reference point.

We iterate the estimation over many bandwidths. Bandwidths go from preliminary deficit of SEK \( \pm 3,000 \) at the most and SEK \( \pm 500 \) at the least. We use symmetric bandwidths around the reference point. A bandwidth of SEK 3,000 means that we only include individuals with weighted preliminary balances in the range from SEK -3,000 to SEK 3,000. A large bandwidth gives a larger sample size and, therefore, more precise estimates. A smaller bandwidth, on the other hand, gives sharper identification but less precise estimates.

There are 1.2 million taxpayers in the maximum bandwidth sample (SEK \( \pm 3,000 \)) as previously mentioned. The minimum bandwidth sample consists of 200,000 taxpayers with weighted preliminary balances in the interval SEK \( \pm 500 \).

\( ^{23} \)We use a logarithmic scale when we iterate over bandwidths. This gives relatively more measurement points at lower bandwidths than at higher bandwidths as seen in, for example, Figure 12.
We have argued that a linear model is appropriate very close to zero. It is, however, reasonable to test for more flexible specifications, in particular 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 1. The estimated coefficients of interest are $\beta_1$ for a possible kink and $\alpha_1$ for a possible discontinuity.

The second- and third-order specifications are:

$$\Delta_i = \alpha_0 + \alpha_1 I_i + \gamma_1 d_i^p + \gamma_2 (d_i^p)^2 + \beta_1 I_i d_i^p + \beta_2 I_i (d_i^p)^2 + \epsilon_i,$$

(18)

$$\Delta_i = \alpha_0 + \alpha_1 I_i + \gamma_1 d_i^p + \gamma_2 (d_i^p)^2 + \gamma_3 (d_i^p)^3 + \beta_1 I_i d_i^p + \beta_2 I_i (d_i^p)^2 + \beta_3 I_i (d_i^p)^3 + \epsilon_i.$$

(19)

We also estimate a zero-order specification as a robustness check. This specification focuses exclusively on a possible discontinuity at zero preliminary balance:

$$\Delta_i = \alpha_0 + \alpha_1 I_i + \epsilon_i.$$

(20)

We estimate the four models (17)–(20) for each bandwidth. The optimal non-zero polynomial model for each bandwidth is then determined based on minimizing a Schwarz Bayesian criterion (SBC). This information criterion weighs better goodness of fit against lost degrees of freedom. SBC is a modification of the Akaike information criterion that punishes lost degrees of freedom harder. We then plot the kink- and discontinuity-estimates from the optimal model over a range of symmetric bandwidths.

Our primary interest is in the $\beta_1$ estimate. We base this on the theoretical predictions from Section 4. The $\beta_1$ parameter measures the change in the derivative at the reference point or, in other words, the kink. We are also interested in if there is a discontinuity at the reference point. This is measured by $\alpha_1$.

Figure 12 reports our estimates of a possible kink at zero preliminary balance. Optimal models are second-order for bandwidths above SEK 2,500, see the (blue) thick bracketed line. First-order models are optimal for smaller bandwidths. The information criterion would, however, select a zero-order model for some of the smaller bandwidths, see the unfilled squares in the figure.
The estimated kink is significant for bandwidths larger 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 decrease when the polynomial order decreases. The estimated coefficients remain 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 13 reports our estimates of this. The estimated coefficients of a discontinuity at zero are statistically significant for all bandwidths except a few bandwidths around SEK 600. It is also clear from the figure that the estimated coefficients increase when the polynomial order decreases. The estimated coefficients decrease when the bandwidth decreases for given order of the model.

We now turn to the corresponding estimates for the three covariates. Figures corresponding to Figure 12 and Figure 13 are presented in Appendix C. The results for employment income show that the kink-estimate is negative and significant for bandwidths between SEK 2,700 and SEK 1,400. Zero polynomial order would, in fact, be optimal for smaller bandwidths. The discontinuity-estimate is only significant for a narrow range of bandwidths between SEK 1,400 and SEK 1,000. We continue with the gender indicator. The optimal polynomial order is always one. There is a significant negative kink for large bandwidths. The discontinuity-estimates are almost always insignificant.

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. The discontinuity-estimates are insignificant for bandwidths below SEK 1,900.

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 the deduction probability show a stable and significant kink for a wider range of bandwidths than for the covariates. The discontinuity-estimates for the deduction variable decrease when the bandwidth becomes smaller. It, however, also remains significant in a much wider range than the estimates for the covariates. Hence, these results indicate that there is no selection effect, but that the observed kink and discontinuity in deduction probabilities are caused by differences in the initial balance.
Figure 12: Kink - estimates of $\beta_1$.

Figure 13: Discontinuity - estimates of $\alpha_1$. 
6.2 Sensitivity analysis – placebo kinks and discontinuities

There is a risk that even the smallest and economically insignificant estimates become statistically significant when the sample size is very large. It is thus 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. We let the reference points vary between SEK -3,000 and SEK 3,000. The bandwidth is fixed at SEK 1,000 in all regressions.

The analysis serves two closely related purposes. The first has to do with causality. The evidence of causal effects becomes stronger if we find that the kinks and discontinuities in deduction probability are more pronounced for the theoretically predicted reference points. The second has to do with selection. It 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 14 shows the estimated kinks in deduction probability for different placebo reference points. The kink-estimates peak slightly below zero at a reference point of SEK -180, which we interpret as zero. There is a concave increase in the deduction probability for those with a preliminary deficit, see Figure 1. This generates moderate negative kinks. The information criterion prefers a zero polynomial order for reference points above SEK 1,700 and below SEK 1,000.

The corresponding discontinuity estimates displayed in Figure 15 mirror the kink estimates. There is a local peak in the discontinuity estimates for reference points close to zero.

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.

We now turn to the corresponding analysis of the covariates. Appendix D has 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, therefore, 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 1. It is hard to come up with a candidate for such an unobservable factor.

\(^{24}\)Incomes are rounded down to the nearest SEK 100 and refunds less than SEK 100 are not paid out (and deficits less than 100 are not claimed), so an initial balance \(\in (-200, 200)\) could be regarded as zero.
Figure 14: Placebo kinks - estimates of $\beta_1$.

Figure 15: Placebo discontinuities - estimates of $\alpha_1$. 

34
Figure 16: The differences in probabilities of claiming a deduction at a SEK 2,000 preliminary deficit with and without estimated kinks and discontinuities.

It is even harder provided that it also needs to be virtually uncorrelated with employment income, gender, and age.

### 6.3 Interpretation and discussion

We have now reached the conclusion that there is a causal and statistically significant effect. Is it, however, also important from an economic point of view? Figure 16 reports the combined effects of estimated kinks and discontinuities for the different bandwidths and reveals the magnitude of the effect. Suppose that we compare the probability of claiming a deduction for a taxpayer with a SEK 2,000 in preliminary deficit with and without 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. The optimal models are here of the second order. The probabilities differences 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 up to 3 percentage point.

The descriptive statistics show that 6.2 percent of those with a preliminary deficit in the maximum band-
width sample claim the deduction for “other expenses”. This gives perspective to the estimated difference in the figure. The estimated effect is thus substantial: Having taxes due implies that the probability of claiming a deduction is approximately 50 percent higher than for those who will receive a refund.

Let us briefly summarize the econometric results so far. We find clear evidence of a kink in the deduction probability at zero preliminary balance for a wide range of bandwidths. Employment income and age do 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. This fraction, however, shows even larger jumps at other reference points. The kink and jump of the dependent variable are substantial in economic terms.

7 Extensions

7.1 An alternative IV-approach

There is a difference between the actual local employment income tax rate and the local employment income tax rate used for preliminary taxation at source as discussed in Section 3. The basic idea in this section is to use this difference as an instrument for a preliminary deficit indicator when estimating probability models for claiming a deduction.

We use two different instruments for the preliminary deficit indicator in the IV-analysis: i) the difference between the preliminary and actual local tax rates and ii) the interaction between the first instrument and employment income. The second instrument can be interpreted as the impact of the rounding on the preliminary balance. The IV-analysis uses the full sample of 3.6 million taxpayers. The OLS and baseline IV-estimate use the standard set of controls, see the notes to Table 3 for details. The models with interacted treatments also include all relevant interacted controls and instruments.

The upper part of Table 3 reports the estimation results for the probability of claiming a deduction. The OLS estimate shows that having a preliminary deficit increases the probability of claiming a deduction by 2.0 percentage points. This estimate cannot, however, be causally interpreted since preliminary deficit may be correlated with the probability of claiming a deduction.

The causal estimate given by the IV-estimation suggests an even stronger effect. We estimate an impact of 5.0 percentage points. The overidentification test does not reject the validity of the instruments. The estimation reported in the third column of the table shows that there is a significant kink in the relationship between the probability of claiming a deduction and the value of the preliminary deficit at zero preliminary
Table 3: The probability of claiming a deduction, IV-approach, full sample.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>IV</td>
<td>IV</td>
<td>IV, gender</td>
<td>IV, age</td>
<td>IV, income</td>
</tr>
<tr>
<td>positive preliminary deficit, indicator</td>
<td>0.020</td>
<td>0.050</td>
<td>-0.027</td>
<td>0.048</td>
<td>0.055</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.006)</td>
<td>(0.046)</td>
<td>(0.007)</td>
<td>(0.010)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>preliminary deficit, amount</td>
<td>-0.000004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>interaction: positive preliminary deficit, indicator *</td>
<td>0.000023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>preliminary deficit, indicator, interacted with:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>man, indicator</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>old (age &gt;= 40), indicator</td>
<td>-0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high income (&gt; SEK 317,700), indicator</td>
<td>0.040</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F stat, first stage</td>
<td>3.387</td>
<td>1.637</td>
<td>1.356</td>
<td>1.388</td>
<td></td>
<td></td>
</tr>
<tr>
<td>overidentification test, p-value</td>
<td>0.798</td>
<td>0.449</td>
<td>0.356</td>
<td>0.708</td>
<td>0.455</td>
<td></td>
</tr>
<tr>
<td>minimum eigenvalue statistic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.33</td>
</tr>
<tr>
<td>( \hat{\lambda} )</td>
<td>1.43</td>
<td>2.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.184)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\lambda} ), men / old / high income</td>
<td>1.97</td>
<td>2.00</td>
<td>2.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.212)</td>
<td>(0.208)</td>
<td>(0.226)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\lambda} ), women / young / low income</td>
<td>2.91</td>
<td>2.66</td>
<td>3.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.404)</td>
<td>(0.382)</td>
<td>(0.626)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value, t-test</td>
<td>0.038</td>
<td>0.130</td>
<td>0.046</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes. Number of observations: 3,610,972.
Heteroscedasticity-corrected standard errors within parentheses.
Baseline instruments: the difference between actual local tax rate and preliminary local tax rate and the tax rate difference times employment income.
Baseline covariates: age, age\(^2\), employment income, (employment income)\(^2\), gender.
Model 3 uses additional instruments that consists of three nonlinear transformations of the baseline instruments:
i) indicator for higher actual local tax rate than preliminary local tax rate;
ii) and iii) consist of the baseline instruments interacted with i).
deficit.

The third column addresses the exclusion restriction by focusing on the relationship above and below zero preliminary deficit. The reported estimates show that there is a significant kink in the relationship between the probability of claiming a deduction and the value of the preliminary deficit at zero preliminary deficit. The kink estimate is somewhat larger and less precisely estimated, but of the same overall magnitude, as the kink estimates reported in the main analysis. The result supports the interpretation that the probability of claiming a deduction increases substantially when an individual exogenously locate in the loss region. This model includes three endogenous regressors and we only have two baseline instruments. We solve this technically by adding nonlinear transformations of the two baseline instruments (see table for details). As a result, the strength of the instruments is substantially reduced; Stock and Yogo (2005) report a bias of around 30 percent for a minimum eigenvalue statistic of 4.3, in a model with three endogenous regressors and five instruments.

The remaining columns in Table 3 report IV-estimations that test for possible differences between different sub-groups in our sample. We conclude that there are no significant gender differences from the estimation reported in column 4. Column 5 in the table presents the results from a corresponding estimation where we test if the response of those older than 40 years differ from the estimated response of younger taxpayers. The result suggest that there are no such differences. We do find, on the other hand, that there is a significant difference in response between those with employment income higher than SEK 317,700 (the threshold for the central government income tax) and those below this threshold. Those above the threshold have a marginal tax rate which is 20 percentage points higher than those below it. The response of those with the higher tax rate is (both economically and statistically) significantly stronger, which is in line with theory, since the marginal value of the deduction is higher above the threshold.

7.2 Estimating the coefficient of loss aversion

The coefficient of loss aversion ($\lambda$) can be expressed in a simple way according to Section 4. It is the share of deducting taxpayers in the loss domain (with sufficiently large preliminary deficits) divided by the share of deducting taxpayers in the gain domain. We use the IV-estimation reported in Table 3, column 2, to calculate these shares. Our estimates of $\lambda$ based on the IV-estimations ignore that the share that claims a deduction does not immediately move to a higher stable level when entering the loss domain. This will
give a downward bias to our $\lambda$-estimates. The reason is that the average deduction probability in the whole deficit region is slightly lower than the average deduction probability on the level further to the right. This is clear from Figure 5. The results from Table 3 are, therefore, conservative estimates of the coefficient of loss aversion.

We estimate the probability of claiming a deduction for a taxpayer with a preliminary surplus and average values of the covariates to 4.23 percent in our full sample of 3.6 million taxpayers. The standard error of this estimate is 0.001. The corresponding probability of claiming a deduction for an average taxpayer with a preliminary deficit is 9.19 percent, with a standard error 0.005. Given these assumptions, our estimate of loss aversion is:

$$\hat{\lambda} = \frac{\hat{Y}}{\hat{X}} = \frac{9.19}{4.23} \approx 2.17. \quad (21)$$

We calculate the standard error of this ratio using the delta method. The resulting estimate of the standard error is 0.184.

The lower part of Table 3 presents our estimates of the coefficient of loss aversion, $\hat{\lambda}$. We estimate $\lambda$ to be higher for women than for men, 2.91 compared to 1.97. This difference is statistically significant at the 5 percent level according to a $t$-test. This test takes potential covariances between the estimated parameters into account. It is possible to do this as we have estimated the response of men and women in the same model.

The estimated $\lambda$ for young is higher than that of old. This difference is, however, not statistically significant at the 5 percent level according to the $t$-test. There is, on the other hand, a statistically significant difference in the estimated $\lambda$’s between those with relatively high income and those with relatively low income. The estimated $\lambda$ of taxpayers with low income is much higher. This is surprising as the estimated effect of preliminary deficit was significantly higher for the high-income group. But the $\lambda$-estimate is a ratio so we need to consider differences in both the numerator and the denominator. The high $\lambda$-estimate for the low-income group is because the deduction probability for those with a preliminary surplus is only about 1.7 percent.

Our estimates are all in the range which has been reported in the literature. Different definitions and different questions and samples have, however, given rise to quite disperse estimates of $\lambda$. The most frequently cited estimate is by Tversky and Kahneman (1992). They report $\hat{\lambda} = 2.25$ using experimental data. Each
individual was observed in both the gain domain and the loss domain. Our overall estimate, \( \hat{\lambda} = 2.17 \), is very close to this. Abdellaoui et al. (2007) cite various studies where the estimates of \( \lambda \) range from 1.4 to 4.8 in risky settings. Gächter et al. (2007) estimate average \( \hat{\lambda} = 2.6 \) for the risky setting and \( \hat{\lambda} = 2.0 \) for the riskless.

Our finding that women are more loss averse than men is similar to that of Schmidt and Traub (2002). Gächter et al. (2007), on the other hand, find no significant gender difference once controlling for other covariates. We have access to very few covariates. This might explain why our results differ.

8 Concluding remarks

We study the vast majority of Swedish working age taxpayers’ behavior when filing their tax returns for the income year 2006. The research method is quasi-experimental using a regression kink and discontinuity approach. We find behavior to be consistent with loss aversion. Taxpayers 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. We find a significant change at zero preliminary deficit. The empirical analysis also makes clear that none of the covariates shows a similar evolution around zero preliminary deficit. We can, therefore, rule out selection. Loss aversion is the obvious candidate for explaining the result.

We also use an alternative IV-approach to strengthen the causal interpretation. Actual local tax rates on employment income are set with two decimals of a percentage (for example, 30.83 percent). Preliminary taxation at source, on the other hand, is based on actual tax rates being rounded to the closest integer percentages tax rates (for example, 31 percent). Taxpayers will tend to pay too much in preliminary taxes if they live in municipalities where tax rates at source are above the actual rates. The preliminary balances will, therefore, tend to be in surplus. The opposite applies to taxpayers who live in municipalities with actual rates above the rates used at source. We use the difference between the actual and the preliminary rates as an instrument for a preliminary deficit indicator. This indicator is used to estimate probability models for claiming a deduction. Our previous results are confirmed using this approach.

We also estimate the coefficient of loss aversion. Our estimate, \( \hat{\lambda} = 2.17 \) for the full sample, is very close to the estimate reported by Tversky and Kahneman (1992), \( \hat{\lambda} = 2.25 \). We have a different approach to estimating this coefficient compared to previous studies. This literature usually uses experimental data from samples of students while we use real-world data for the majority of taxpayers in a country. We still arrive
at results very close to those previously reported.

The main contribution is that we have shown that taxpayers do not act in a way consistent with neo-classical theory when filing their tax returns, but that they tend to be loss averse. We had the opportunity to study a certain deduction, namely the deduction for “other expenses for earning employment income”, which is not very widely used, but which tend to be less correlated with income, gender and other covariates than, e.g., deductions for travel costs. Only about 4 percent of our full sample claim the deduction, but we can still find a clear indication of loss aversion. When the tax file indicates that they have taxes due they are 50 percent more likely to claim deductions than if they would get a refund. The effect is even stronger for those in a higher marginal tax rate bracket.

The implication for tax policy is that a slight over-withholding of preliminary taxes would enhance compliance. This would not only increase tax revenue and strengthen tax morale, it would also reduce tax auditing costs. Over-withholding should not be be made to a too large extent, however. Taxpayers could feel wrongly treated if too much is always withheld. A systematic over-withholding could also shift the reference point from being zero to being a certain positive amount, so that the effect would disappear.

Per Engström, Henry Ohlsson: Department of Economics, Uppsala University and Uppsala Center for Fiscal Studies (UCFS), Department of Economics, Uppsala University

Katarina Nordblom: UCFS and Department of Economics, University of Gothenburg

Annika Persson: The Swedish Tax Agency
Appendix A  More on the institutional framework

The Tax Agency provides a lot of information to taxpayers. The information is very detailed as is clear from this excerpt from SKV (2009).

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’. M any 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 (kontrolluppgifter) 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.

<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</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>
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?

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

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

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

**DONE!**

**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
- have specified a bank account for refund of tax.

**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.

---

<table>
<thead>
<tr>
<th>June</th>
<th>August - September</th>
<th>December</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>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.</strong></td>
<td><strong>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.</strong></td>
<td><strong>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.</strong></td>
</tr>
</tbody>
</table>
The tax account: interest rates

The taxpayer’s tax account is interest bearing. During the assessment year 2007, a surplus on the tax account earned a tax exempt interest of 1.35 percent from 13 February 2007. This interest rate was increased to 1.8 percent from 1 July 2007. The increase was decided 25 June 2007. Interest was earned until the surpluses was refunded.

The taxpayer has to pay interest on deficits on the tax account. The interest paid on tax account deficits is not deductible. The interest rates were as follows during the assessment year 2007:

<table>
<thead>
<tr>
<th>from</th>
<th>deficit</th>
<th>interest rate, percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 February</td>
<td>&lt;=20000</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>20001-30000</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>&gt;30001</td>
<td>3.75</td>
</tr>
<tr>
<td>4 May</td>
<td>&lt;=30000</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>&gt;30000</td>
<td>3.75</td>
</tr>
<tr>
<td>1 July</td>
<td>&lt;=30000</td>
<td>4</td>
</tr>
<tr>
<td>(rate increase)</td>
<td>&gt;30000</td>
<td>5</td>
</tr>
<tr>
<td>due date</td>
<td>&lt;=10000</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>&gt;10000</td>
<td>19</td>
</tr>
</tbody>
</table>

Preliminary employment income taxes at source

There were nine tables for preliminary employment income taxes at source for integer percentages local government income tax rates from 29 percent to 37 percent in 2006. Tax payers facing local government income tax rates with decimals of a percentage ≤ .50 pay preliminary income taxes according to the closest
downward integer percentage table, while taxpayers facing local government income tax rates with decimals of a percentage > .50 pay preliminary income taxes according to the closest upward integer percentage table.
### Appendix B  Descriptive statistics

Table 4: Descriptive statistics, full sample.

<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, conditional</td>
<td>4,716 (4,324)</td>
<td>3,559 (3,342)</td>
<td>3,894 (3,692)</td>
</tr>
<tr>
<td>preliminary balance, SEK, unweighted</td>
<td>-10,757 (51,232)</td>
<td>6,553 (6,762)</td>
<td>2,667 (26,017)</td>
</tr>
<tr>
<td>employment income, SEK</td>
<td>279,770 (135,523)</td>
<td>269,296 (115,198)</td>
<td>271,648 (120,140)</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 (12.3)</td>
</tr>
</tbody>
</table>

*Note.* mean (standard deviation)
Table 5: Descriptive statistics, maximum bandwidth sample.

<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>370,929</td>
<td>822,686</td>
<td>1,193,615</td>
</tr>
<tr>
<td>deducting, fraction</td>
<td>0.061</td>
<td>0.042</td>
<td>0.048</td>
</tr>
<tr>
<td>deduction, SEK, conditional</td>
<td>4,212 (3,813)</td>
<td>3,714 (3,403)</td>
<td>3,910 (3,578)</td>
</tr>
<tr>
<td>preliminary balance, SEK, unweighted</td>
<td>-1,289 (1,140)</td>
<td>1,611 (1,153)</td>
<td>710 (1,767)</td>
</tr>
<tr>
<td>preliminary balance, SEK, weighted</td>
<td>-1,263 (856)</td>
<td>1,582 (847)</td>
<td>698 (1,567)</td>
</tr>
<tr>
<td>employment income, SEK</td>
<td>277,908 (130,544)</td>
<td>276,876 (119,073)</td>
<td>277,197 (122,753)</td>
</tr>
<tr>
<td>men, fraction</td>
<td>0.49</td>
<td>0.50</td>
<td>0.49</td>
</tr>
<tr>
<td>age, years</td>
<td>44.9 (12.5)</td>
<td>43.1 (12.3)</td>
<td>43.7 (12.4)</td>
</tr>
</tbody>
</table>

Notes: mean (standard deviation)
Sample selection criterion: weighted preliminary balance in the interval ± SEK 3,000.
Figure 17: Number of observations at different bandwidths.
Appendix C  Estimates for covariates
Figure 18: Kink - estimates for employment income.

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

Figure 21: Jump - estimates for gender.
Figure 22: Kink - estimates for age.

Figure 23: Jump - estimates for age.
Appendix D  Placebo estimates for covariates
Figure 24: Kink - placebo estimates for employment income.

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

Figure 27: Jump - placebo estimates for gender.
Figure 28: Kink - placebo estimates for age.

Figure 29: Jump - placebo estimates for age.
References


