

# Loss Aversion Motivates Tax Sheltering: Evidence From U.S. Tax Returns

Alex Rees-Jones\*

The Wharton School, University of Pennsylvania

December 8, 2014

**Abstract:** This paper presents evidence that loss aversion affects taxpayers as they file their annual tax returns. I model the decisions of a loss-averse tax filer who may use tax shelters to manipulate the “balance due” exchanged with the IRS. I use this model to derive distinguishing predictions of loss aversion which facilitate its identification and quantification in the field. Under loss framing, the discretely steeper marginal utility of a dollar motivates greater pursuit of shelters. These motives imply that the post-sheltering distribution of balance due will exhibit a structural shift in the loss domain, due to discretely higher sheltering in this region. Furthermore, the discontinuity in marginal incentives generates excess mass, or “bunching,” at the gain/loss threshold. Using the 1979-1990 IRS Panel of Individual Returns, I document the predicted bunching and shifting in the distribution of balance due, and examine the causes and correlates of these features. The observed distribution is consistent with the framing of tax payments as losses and tax refunds as gains, and is difficult to rationalize with plausible alternative theories. Using two complementary structural approaches—identified from the bunching and shifting predictions, respectively—I estimate substantial potential policy impact of this psychological bias. These results have direct implications for tax policy and public finance.

**Keywords:** Loss aversion, income taxation, tax sheltering, tax avoidance, tax evasion.

**JEL Classification Numbers:** D03, H24.

---

\*Email: [alre@wharton.upenn.edu](mailto:alre@wharton.upenn.edu). For comments and suggestions that improved the project, I am grateful to John Bakija, Dan Benjamin, Jim Berry, Greg Besharov, Aaron Bodoh-Creed, Steve Coate, Stefano DellaVigna, Ori Heffetz, Ben Ho, David Laibson, Rosella Levaggi, Francesca Molinari, Ted O’Donoghue, Daniel Reck, Charlie Sprenger, Ken Whelan, seminar audiences at Boston University SOM, Cornell University, Dartmouth College, Haas School of Business, Harvard Business School, NBER, University of Michigan, and Wharton, as well as conference participants at the NTA Annual Conference on Taxation, the Southern Economics Association annual meetings, the Stanford Institute for Theoretical Economics, and the CEBID Conference on Taxation, Social Norms, and Compliance. I thank the Cornell Institute for the Social Sciences, the National Bureau of Economic Research, and the National Institute on Aging (grant T32-AG00186) for generous research support, and Anthony Hawkins for research assistance. The views expressed in this paper are those of the author, and do not necessarily represent the views of any other individual or organization.

Loss aversion—the tendency to behave according to discretely steeper marginal utility when facing a perceived loss—is a central concept in behavioral economics. This psychological mechanism features prominently in the vast theoretical literature utilizing prospect theory, and has been robustly demonstrated in a wide variety of experimental settings. Despite its prevalence in theoretical and experimental work, there remain relatively few widely accepted demonstrations of loss aversion in the field.<sup>1</sup> As noted by Barberis (2013), there is reason to believe that models incorporating this feature will ultimately take a permanent and significant place in economic field analysis; however, Barberis also notes that the relative shortage of field work permits the interpretation that loss aversion is less relevant outside the laboratory.

In this paper I present evidence that loss aversion is directly relevant in a large-scale field setting of unambiguous economic importance: the manner in which individuals react to income taxes. The rationale for how loss aversion might affect a taxpayer is straightforward. Throughout the year, a taxpayer earns taxable income, takes actions that might be tax advantaged, and makes tax payments based on a forecast of the tax liability that will ultimately be owed. In preparation for tax day, these activities must be precisely documented and reported to the IRS, and the “balance due”—the difference between the total taxes owed and the tax payments already made—must be settled. If the balance due is positive, the tax filer must pay that amount to the IRS, and thus incur a loss in a very literal way. If the balance due is negative, the tax filer collects a refund, yielding a literal gain. Models of loss aversion predict a discretely higher marginal disutility of a dollar taxed when facing a loss. This disutility can influence the taxpayer’s attempts to reduce tax liability through decisions made in the course of tax filing. I will broadly refer to such attempts as “tax sheltering.” Identifying and quantifying the effect of loss aversion on tax sheltering decisions will be the central focus of this paper.

To begin, section 1 presents a theoretical framework that clarifies how loss aversion might

---

<sup>1</sup>Examples of published field tests of loss aversion include studies of taxi-driver labor supply (Camerer, Babcock, Loewenstein, and Thaler, 1997; Farber, 2005; Farber, 2008; Crawford and Meng, 2011), housing prices (Genesove and Mayer, 1999), putting behavior of professional golfers (Pope and Schweitzer, 2010), and behavior in financial markets (reviewed in Barberis, 2013). Related research in progress includes studies of the effect of alternative policies to reduce shopping bag use (Homonoff, 2013), the goal-setting behavior of marathon runners (Allen, Dechow, Pope, and Wu, 2014), and job search (DellaVigna, Lindner, Reizer, and Schmieder, 2014).

be detected in tax data. I consider a model in which taxpayers manipulate their balance due through costly sheltering, and explore the implications of different utility models on the resulting manipulated distribution. If taxpayers' perceived value of a marginal dollar drops discontinuously when losses turn to gains, sheltering motives induce a discontinuity in the balance-due distribution. Individuals in the loss domain (positive balance due) would pursue excess tax sheltering activities relative to individuals in the gain domain (negative balance due). Moreover, a discrete fraction of taxpayers would choose to shelter to precisely the gain/loss threshold, then discontinue pursuit of additional shelters in response to the sudden drop in marginal return.

Section 2 describes the data used to test these predictions, the IRS Statistics of Income 1979-1990 Panel of Individual Returns. This dataset follows a large random sample of taxpayers, and provides a detailed look at the distribution of balance due, as well as the income, deductions, and credits used in the balance-due calculation. Panel sampling allows direct observation of changes in reported taxable behavior over time, which proves useful when contrasting the loss-averse model with alternative theories.

Section 3 tests the identifying implications of loss aversion. As predicted by the model, the distribution of balance due is shifted in a manner consistent with higher sheltering in the loss domain, and significant excess mass is seen precisely at zero balance due. To quantify the additional sheltering that loss aversion motivates, I develop two alternative structural approaches to estimate the effect of loss framing. In line with the theoretical results of section 1, these two approaches are identified from the excess mass at the gain/loss threshold and the distributional shifting in the loss domain, respectively. I present estimates based on each approach, and discuss their comparative strengths and weaknesses.

Section 4 presents additional analysis to assess whether the behavior documented in the previous section can be attributed to psychological sheltering motives. A variety of behaviors indicative of excess tax sheltering are shown to be associated with the documented bunching behavior, and systematically accurate tax forecasting is rendered implausible as an alternate explanation due to the atypical income paths of those reporting zero balance due. Having established the presence of a sheltering response, I consider other potential explanations, such as fixed costs in the loss domain, financial constraints, interactions with tax preparers,

and avoidance of the underwithholding penalty. While there is reason to believe these factors contribute to tax sheltering decisions, they face difficulties as explanations of the primary behaviors documented in this paper.

Section 5 uses the bunching-based and shifting-based structural estimates of section 3 to quantify the potential policy impact of harnessing loss aversion. I estimate the difference in aggregate sheltering predicted to result from a marginal change in the fraction of taxpayers facing a loss. The magnitude of this estimate varies across the two structural approaches used, but suggests the potential for substantial aggregate revenue effects of framing manipulations.

Section 6 concludes by discussing the implications of this study for tax policy, public finance, and behavioral economics. These results inform the optimal implementation of tax withholding, the detection and deterrence of tax evasion, the implementation of bunching-based identification strategies, and the theoretical debate on the correct specification of reference points.

The possibility of reference dependence affecting reactions to the income tax has been the topic of a considerable amount of prior research. Several papers have presented theoretical treatments of loss-averse taxpayers, and have shown that loss aversion can help rationalize a variety of features of our tax system, such as the high rate of voluntary compliance.<sup>2</sup> The presence gain/loss framing effects has also been seen in a number of small-scale surveys and lab experiments.<sup>3</sup> Despite encouraging results from this line of research, direct study of this phenomenon in the field has been limited, presumably due to data constraints and the difficulty of identification.<sup>4</sup> This paper contributes to this literature by presenting new

---

<sup>2</sup>See, for example, Elffers and Hessing (1997), Yaniv (2001), Bernasconi and Zanardi (2004), Kanbur, Pirttilä, and Tuomala (2008), or Dhimi and al Nowaihi (2007, 2010).

<sup>3</sup>See, for example, Chang, Nichols, and Schultz (1987), Copeland and Cuccia (2002), Kirchler and Maciejovsky (2001), Robben et al. (1990), Robben, Webley, Elffers, and Hessing (1990), or Schepanski and Shearer (1995). In contrast, Schadewald (1989) presents experimental results where manipulations of reference points did not have significant effects.

<sup>4</sup>In the vast literature on tax behavior, many papers have examined the correlates of underwithholding. A number of these results are consistent with loss aversion—for example, that underwithholding is positively associated with income underreporting (see, e.g., Clotfelter, 1983) and IRA contributions (see, e.g., Feenberg and Skinner, 1989). However, clear identification of loss-averse tax behavior has rarely been attempted or achieved in the field, with several notable exceptions. Feldman (2010) studies the impact of a change in withholding law on tax-advantaged retirement savings through the lens of a mental accounting model, but carefully considers loss aversion over consumption as an alternative (ultimately rejected) explanation of her results. Engström, Nordblom, Ohlsson, and Persson (2013) present evidence that loss aversion leads to a

implications of loss aversion, tailored to be observable in tax records despite reflecting potentially unobserved sheltering activities. This facilitates the detection of this mechanism in the field, and permits inference on its aggregate impact to tax revenues. The resulting estimates suggest that this setting is among the most economically important applications of loss aversion yet documented, both in terms of the number of individuals affected and the magnitude of the economic consequences.

## 1 Theoretical framework

In this section, I model the sheltering decisions of taxpayers who are in the process of filing their annual tax returns. This model is used to formally characterize the distinguishing observable implications of loss aversion, setting the foundation of the empirical approach pursued in the remainder of the paper.

### 1.1 A simple model of sheltering decisions

In preparation for tax day, taxpayers must complete and submit Form 1040 or one of its variants, formally documenting their tax-relevant information for the year.<sup>5</sup> Completing this form involves identifying oneself, documenting taxable income, claiming credits or deductions to that taxable income due to participation in tax-incentivized behaviors, calculating the total taxes owed, and finally comparing these taxes owed to taxes already paid. This comparison yields the “balance due,” the amount of money that must be exchanged between the taxpayer and the IRS.

In the process of filing Form 1040, taxpayers have the opportunity to manipulate their balance due through legal or illegal tax sheltering. Legal means of tax sheltering typically entail pursuing and reporting behavior that grants a reduction in tax or taxable income. This can include decisions as commonplace as itemizing deductions, deducting business expenses, or investing money in tax-preferred savings accounts. Illegal means for tax sheltering take the form of underreporting taxable income or overreporting tax-advantaged behaviors.

---

higher take-up rate of a specific deduction in Swedish tax code.

<sup>5</sup>Individuals with particularly simple taxable behavior may fill out shortened and simplified versions of this form, 1040A or 1040EZ.

Finding and employing tax shelters is costly. For legal tax shelters, these costs include the effort necessary to find tax benefits for which the taxpayer qualifies, as well as the time and effort needed to document and claim those tax benefits.<sup>6</sup> For illegal tax shelters, these costs can include accounting effort as well as the expected future penalties that will be incurred if evasion is detected. Non-monetary costs associated with evasion—as might be generated from, e.g., psychological stigma—can similarly be incorporated.

Taxpayers thus face a tradeoff between the value of reducing tax payments and the cost that must be incurred to do so. Sheltering decisions made in consideration of this tradeoff can be represented by the simple utility maximization problem:

$$\max_{s \in \mathbb{R}^+} \underbrace{m(-b^{PM} + s)}_{\text{utility over money}} - \underbrace{c(s)}_{\text{cost of sheltering}} \quad (1)$$

In the equation above,  $b^{PM}$  denotes “pre-manipulation” balance due, the balance due the taxpayer would owe prior to the tax sheltering pursued in the course of tax filing.  $b^{PM}$  is assumed to be the realization of a continuous random variable with a continuous PDF  $f_b^{PM}$ .  $s$  denotes the sheltering pursued in an attempt to manipulate balance due.  $m(\cdot)$  denotes the utility from money, while  $c(\cdot)$  denotes the disutility generated from the costly pursuit of sheltering. Assume that  $c(\cdot)$  is increasing and twice continuously differentiable. Further assume that the taxpayer pursues shelters with the lowest marginal cost first, which implies that  $c(\cdot)$  is convex. The balance due reported to the IRS is the final, post-manipulation amount  $b = b^{PM} - s$ , distributed according to the PDF  $f_b$ . For positive values of balance due, the IRS is owed money; for negative values of balance due, a refund is due to the taxpayer.

Two modeling decisions reflected in this approach merit further discussion. The first is my treatment of pre-manipulation balance due. In the broad context of tax-related decision making over time,  $b^{PM}$  is endogenously determined by past decisions regarding labor supply, withholding, and the pursuit of tax-incentivized behaviors. I do not attempt to model

---

<sup>6</sup>Recent research suggests that this “hassle cost” of tax filing is substantial; using a revealed-preference approach based on the amount of “money left on the table,” Benzarti (2014) estimates that itemizing deductions is as painful as working 19 hours at one’s job.

and estimate the structure of this endogenous process—a task that would be exceptionally difficult, and ultimately unnecessary for distilling the primary implications of loss aversion. Instead, I present my theoretical predictions with respect to the distribution of  $b^{PM}$  that this process ultimately generates. The primary predictions I will explore assume that this resulting distribution is endowed with a continuous PDF. There is reason to suspect that this would be a reasonable approximation, particularly for relatively high-income tax filers with complex sources of income. At the time when tax prepayment decisions are made, the final tax liability that will be owed for the year is often uncertain.<sup>7</sup> Forecasting error induced by this uncertainty would naturally be expected to “smooth out” point masses in the balance due PDF. In practice, the empirical distributions examined in section 3 are well approximated by smooth distributions on nearly all of their support, and seem to validate this modeling assumption. Of course, this assumption might fail if some individuals are able to systematically and perfectly target their tax prepayments, naturally resulting in zero balance due year-to-year. This potential problem is evaluated in section 4, and is shown to be unlikely to confound the behavior I document.

An additional abstraction that merits attention is my treatment of  $s$  as a choice from  $\mathbb{R}^+$ . This subtly imposes an assumption that finely manipulable tax shelters are available, which would allow to-the-dollar targeting of total sheltering. Fine manipulation is certainly possible for some sheltering decisions; for example, when a tax evader chooses the precise amount of income to report. However, many tax shelters come in discrete units, and this discreteness could generate imprecision in the targeting of optimal sheltering. Abstracting from this discreteness does not meaningfully affect the main intuitions generated by the loss-averse model. However, this modeling decision will prove important in understanding and interpreting some of the main results. An econometric specification that relaxes the fine manipulability assumption is presented in section 3.3.

---

<sup>7</sup>This uncertainty can be generated from true market uncertainty regarding, e.g., the return on investments or the performance of a small business. It could alternatively be generated from simple inattention to the forecasting problem inherent in withholding decisions, as has been documented in Jones (2012).

## 1.2 Contrasting implications of standard and loss-averse utility

Having established a basic framework for considering sheltering decisions made during the course of tax filing, we may now consider the implications of different utility models in this environment. The goal of this exercise is to generate predictions that could differentiate a “standard” utility model from a loss-averse model, relying only on data observable in tax records. This poses a challenge, since the primary behavior we wish to study—tax sheltering—is notoriously difficult to precisely define, measure, or identify. For the purposes of the model presented above, tax sheltering is broadly defined as the pursuit of modifications to reported tax liability for the explicit purpose of its reduction. A precise measurement of tax sheltering, thus defined, requires knowledge not only of taxpayers’ behaviors, but also of their intentions. Under such a definition, both the quantity of taxes sheltered ( $s$ ) and the pre-manipulation balance due ( $b^{PM}$ ) are fundamentally unobserved in administrative tax data. For the predictions to follow, we will compare alternative sets of assumptions on the structure of utility over money, and consider the implications of these structures on *observed* balance due ( $b$ ).

To model a baseline case without loss aversion, assume that utility over money depends on weakly concave, smooth preferences over final wealth ( $w$ ):  $m(-b^{PM} + s) = u(w - b^{PM} + s)$ . To focus attention on the motives of individuals pursuing tax sheltering, assume the model parameters imply a non-zero, but finite, optimal sheltering solution.<sup>8</sup> A distinguishing feature of this model is the continuity in balance due that it generates, expressed formally in proposition 1 below.

**Proposition 1.** *In the sheltering decision problem of section 1.1, if  $m(-b^{PM} + s) = u(w - b^{PM} + s)$ , where  $u(\cdot)$  is weakly concave and twice continuously differentiable, then  $f_b$  is continuous.*

*Proof.* See appendix A.

Put simply, in a model with a continuously distributed pre-manipulation balance due and a smooth, convex cost function, a standard smooth utility model will not naturally generate

---

<sup>8</sup>Formally, assume that  $c'(0) < u'(w - b^{PM})$  and  $\lim_{s \rightarrow \infty} u'(w - b^{PM} + s) - c'(s) < 0$ .



discontinuities in the distribution of final balance due ( $f_b$ ). This is not meant to suggest that discontinuities in  $f_b$  could not exist. Rather, this suggests that any such discontinuities observed must be attributable to a discontinuity in the underlying pre-manipulation balance-due distribution, or in the structure of marginal sheltering costs.

To illustrate the intuition of proposition 1, consider the case where  $u(\cdot)$  is linear with slope  $\beta$ , thus abstracting from the diminishing marginal utility of wealth. Under this utility structure, optimal sheltering is  $s^* = c'^{-1}(\beta)$ .  $c'^{-1}(\cdot)$  denotes the inverse function of the derivative of  $c(\cdot)$ , which is guaranteed to exist and to be increasing due to the assumed monotonicity and convexity of  $c(\cdot)$ . Since the final balance due is  $b = b^{PM} - s^*$ , the distribution of  $b$  corresponds to the distribution of  $b^{PM}$  shifted by the constant value  $c'^{-1}(\beta)$ . As a result, continuity of  $f_b^{PM}$  will clearly imply continuity of  $f_b$ . Allowing for income effects by reintroducing concavity in  $u(\cdot)$  generates a non-constant optimal sheltering solution; however, the intuition that sheltering motives induce a smooth shift of  $f_b$  still holds.

Now consider instead a loss-averse taxpayer, with utility over money defined as:

$$m(-b^{PM} + s) = (w - b^{PM} + s) + \phi(-b^{PM} + s - r). \quad (2)$$

The first term,  $(w - b^{PM} + s)$ , reflects the value of a sheltered dollar of tax payment in the same manner as the linear example above, and thus abstracts from income effects. The second term,  $\phi(-b^{PM} + s - r)$ , allows for the influence of reference dependence relative to the reference point  $r$ . To capture loss aversion,  $\phi$  is specified according to a piecewise-linear version of the prospect theory value function (Kahneman and Tversky, 1979):

$$\phi(x) = \begin{cases} \eta x & \text{if } x \geq 0 \\ \eta \lambda x & \text{if } x < 0 \end{cases} \quad (3)$$

$\lambda$  is the coefficient of loss aversion, assumed to be greater than 1.  $\eta$  captures the weight on the loss-averse utility component relative to the direct utility component.  $r$  is assumed to be an exogenously determined reference value of balance due. Assumptions on the nature of this reference point will be discussed after results for an arbitrary reference point are established.

The implied optimal sheltering behavior resulting from this model is given by the piecewise solution:<sup>9</sup>

$$s^*(b^{PM}) = \begin{cases} c'^{-1}(1 + \eta\lambda) & \text{if } b^{PM} > c'^{-1}(1 + \eta\lambda) - r \\ b^{PM} + r & \text{if } b^{PM} \in [c'^{-1}(1 + \eta) - r, c'^{-1}(1 + \eta\lambda) - r] \\ c'^{-1}(1 + \eta) & \text{if } b^{PM} < c'^{-1}(1 + \eta) - r \end{cases} \quad (4)$$

In words, a sufficiently large pre-manipulation balance due results in a high level of sheltering,  $c'^{-1}(1 + \eta\lambda)$ . A sufficiently small pre-manipulation balance due results in a low level of sheltering,  $c'^{-1}(1 + \eta)$ . For an intermediate range of pre-manipulation balance due, the level of sheltering chosen will be exactly the amount necessary to offset the pre-manipulation tax bill and reach the gain/loss threshold. For notational convenience, the low and high sheltering values will be denoted as  $s^L$  and  $s^H$ , respectively.

Equation 4 implies that the distribution of final reported balance due can be expressed as:

$$f_b(x) = \begin{cases} f_b^{PM}(x + s^L) & \text{if } x < r \\ F_b^{PM}(r + s^H) - F_b^{PM}(r + s^L) & \text{if } x = r \\ f_b^{PM}(x + s^H) & \text{if } x > r \end{cases} \quad (5)$$

A graphical representation of this solution, and the relationship between the pre- and post-manipulation distributions, is presented in figure 1. The qualitative features of this distribution, which contrast with the smooth distribution predicted in proposition 1, are summarized in propositions 2 and 3 below.

**Proposition 2.** *Consider the sheltering decision problem of section 1.1. If  $m(-b^{PM} + s)$  takes the loss-averse specification of equation 2, and if  $r$  is in the support of  $f_b$ , then  $f_b$  exhibits a point mass at  $r$ .*

Proposition 2 will be referred to as the “bunching prediction,” and summarizes the loss-averse model’s prediction of excess mass at the reference point.

---

<sup>9</sup>A similar optimal sheltering solution was presented in the related framework of Engström, Nordblom, Ohlsson, and Persson (2013).

**Proposition 3.** *Consider the sheltering decision problem of section 1.1. Assume that  $m(-b^{PM} + s)$  takes the loss-averse specification of equation 2, and that  $r$  is in the support of  $f_b$ . Then, denoting  $\tilde{s} \equiv s^H - s^L$ , the balance-due distribution for all non-reference values is:*

$$f_b(x) = \begin{cases} f_b^{PM}(x + s^L) & \text{if } x < r \\ f_b^{PM}(x + s^L + \tilde{s}) & \text{if } x > r \end{cases} \quad (6)$$

Proposition 3 will be referred to as the “shifting prediction.” This prediction clarifies the consequences of the additional sheltering motives a loss-averse taxpayer faces in the loss domain. The resulting manipulation generates a balance-due distribution that can be expressed as a horizontal shift of  $f_b^{PM}(x)$ , with a discretely greater horizontal shift when the balance due implies a loss ( $x > r$ ).

Both propositions follow immediately from the implied balance-due distribution in equation 5. In the analysis to come, each will form the basis of distinct structural approach permitting inference on  $s^H - s^L$ , the excess sheltering pursued when facing a loss.<sup>10</sup>

### 1.3 From individuals to populations

The results of the previous subsection characterize the predictions of an individual decision problem. The empirical analysis of section 3 will assess these predictions using data on a large population of tax filers. In this subsection, I discuss how the structural predictions of the individual decision problem translate into structural predictions for the population.

To begin, consider the implications of the model when we hold utility structure fixed but allow for heterogeneity in the pre-manipulation balance-due distribution. To study such a case, assume the population is split into  $G$  groups. A given group consists of fraction  $p_g$  of the

---

<sup>10</sup>Instead of quantifying the behavioral response in terms of additional dollars sheltered (the approach taken in this paper), one could instead imagine proceeding by estimating the parameters of the utility function directly—e.g., the coefficient of loss aversion. However, it can be seen from equation 4 that the coefficient of loss aversion impacts sheltering in a manner fundamentally intertwined with the shape of the cost function. Without strong parametric assumptions on the cost function, the two are not separately statistically identified by data observed in tax records. In contrast, the excess sheltering that loss aversion generates is statistically identified with relatively weak assumptions, and furthermore is the relevant parameter for forecasting the revenue effects of this behavioral bias. This motivates my focus on this means of quantifying the loss-averse behavior observed.

population, and is endowed with a group-specific pre-manipulation balance-due distribution  $f_{bg}^{PM}$ . In this case, the population distribution of both  $b^{PM}$  and  $b$  can be expressed as a mixture of the group-specific distributions, formally given by

$$f_b^{PM} = \sum_{g=1}^G p_g \cdot f_{bg}^{PM} \text{ and } f_b = \sum_{g=1}^G p_g \cdot f_{bg}, \quad (7)$$

where  $p_g$  indicates the fraction of taxpayers within group  $g$ . For a population sharing the standard utility model considered in proposition 1, we have already established that each group-specific distribution  $f_{bg}$  is continuous. Since a finite mixture of continuous distributions is itself continuous, the implications of proposition 1 translate to the population distribution without modification.

Whether the bunching and shifting predictions apply to a population distribution depends on the manner in which reference points differ across individuals. If reference points are different across many groups or individuals, the heterogeneity in framing can effectively “smooth out” the discontinuities at reference points held by a small fraction of the population. If, instead, the reference point is commonly framed—that is, if all of a population shares the same reference point—then the population-level distribution expressed in equation 7 preserves the structure implied by the individual case expressed in equation 5. As a result, both the bunching prediction and the shifting prediction remain valid.

In the primary empirical analysis of section 3, I will test for the presence of a *subpopulation* who commonly frame balance due relative to a reference point of zero. “Losses” correspond to the literal out-of-pocket losses faced by sending a payment to the IRS. “Gains” correspond to the literal into-pocket gains faced by receiving a tax rebate. A number of laboratory studies of loss-averse tax behavior have tested for reference dependence relative to this reference point, and documented results consistent with this framing (e.g., Shepanski and Shearer, 1995). Reasoning expressed in journals of tax-related thoughts further support this framing by at least some individuals (Carroll, 1992). In short, based on both past evidence and intuitive appeal, zero is a natural candidate for the gain/loss threshold in this environment, thus motivating the tests to come in section 3. However, the theoretical foundations of reference points are the topic of an active, and unresolved, recent literature. In section 6 I

will discuss results relevant to two alternative models of the reference point: a status-quo model and a point-expectation model motivated by Kőszegi and Rabin (2006, 2007).

## 1.4 Summary of theoretical results

This theory as a whole formalizes the notion that loss aversion will motivate individuals facing a loss to do more to reduce taxes, leading to a greater degree of latent manipulation in the loss domain of the balance-due distribution. Furthermore, this theory predicts that an excess mass of tax filers will cease their sheltering efforts at zero balance due as a result of the sudden drop in perceived marginal payoff. The empirical results in section 3 support these basic intuitions and the estimates in section 5 assess their impact on total tax revenue. In practice, it is important to note that aggregate revenue effects of loss aversion will be nearly entirely driven by the shifting of the loss domain expressed in proposition 3. The prediction of excess mass at a single point, while not the primary driver of policy impact, offers a stark and easily observable property that greatly facilitates the identification of this mechanism. The purpose of documenting this excess mass is to permit inference on the underlying data generating process, and to infer the aggregate effects to tax revenue that it would imply. With these goals in mind, we are prepared to turn to tax data and test the predictions of the loss-averse model.

## 2 Data

The data considered in this study come from the the 1979-1990 IRS Statistics of Income (SOI) Panel of Individual Returns, which I obtained from the Office of Tax Policy Research at the University of Michigan. The SOI Panel of Individual Returns is an unbalanced panel which follows a random sample of tax filers. Randomization occurred over social security numbers: five four-digit numbers were drawn, and tax filers whose last four SSN digits matched one of these codes were included in the sample. Not all five codes were sampled in all years; appendix table A.1 illustrates the sampling pattern over time. These data contain many line items from Form 1040 and the relevant supplemental schedules, allowing direct observation of balance due and many steps of its calculation.

In the process of preparing the dataset, I exclude data according to several criteria. First, I restrict my sample to taxpayers in the 50 states or the District of Columbia. Second, I remove a small number of observations which were drawn from a different sampling frame.<sup>11</sup> Finally, I drop any data for filing years before 1979.<sup>12</sup> These exclusions remove 3,051 observations from the raw data, and yield a sample size of 291,275 person-years for 64,027 tax filers.

For most analysis, I will further restrict the data to only individuals with non-zero total tax liability as well as non-zero tax prepayments. This restriction excludes 62,159 observations from the data. Note that individuals without taxable income will often face a balance due of zero for reasons unrelated to loss aversion, potentially confounding the bunching prediction. For individuals with zero tax prepayments, zero balance due aligns with zero total tax, which also presents a potential confound. Excess mass at zero total tax has previously been documented, and can be attributed to non-preference-based discontinuities in the tax environment (Saez, 2010).<sup>13</sup> I exclude these observations to avoid these potential confounds. I refer to the dataset with these individuals excluded as the “primary sample,” in contrast to the “full sample” above. This sample consists of 229,116 tax returns filed by 53,177 taxpayers, and basic summary statistics are presented in table A.2.

### 3 Assessing the predictions of loss aversion

In this section, I test the predictions of the loss-averse sheltering model presented in section 1. After presenting an initial examination of the nominal balance-due distribution, I turn to investigating both the bunching and the shifting predictions individually, and quantifying the additional tax sheltering each of these features imply.

---

<sup>11</sup>This panel was generated by randomization based on groups of social security numbers, but not all groups were sampled in all years of the panel. In years where a given group was excluded from this panel, it was not excluded from other IRS sampling frames. As a result, a small number of those taxpayers were randomly sampled to be part of that year’s IRS tax model file. These observations were subsequently included in this panel, but flagged. I exclude them to preserve a consistent sampling structure.

<sup>12</sup>The small number of such observations available are tardy returns filed during the sampling period.

<sup>13</sup>In particular, bunching at zero total tax might reasonably be expected due to a) the discontinuity in marginal tax, and b) the nature of nonrefundable credits and deductions. Nonrefundable credits cannot be used to generate a net refund for the year relative to total tax payment (the sum of tax prepayments and payments on tax day). Assuming non-zero tax prepayments, the total tax and balance due are distinct, and thus nonrefundable credits *can* be used to generate a refund on tax day. A refund on tax day is sometimes alternatively called a rebate, meant to differentiate from this alternative usage of the term “refund.”

To begin, figure 2 presents a frequency histogram of the nominal balance due reported, with a bin size of \$1. As is visually apparent, this distribution is reasonably smooth and bell-shaped, although more sharply peaked than a standard normal distribution. Consistent with proposition 2, a point mass is seen precisely at zero. Consistent with proposition 3, the distribution of balance due for those owing a tax payment appears shifted to lower values.<sup>14</sup>

To study these features of the data more formally, I fit a symmetric distribution to the histogram of negative values of balance due (the gain domain). I then extrapolate predicted frequencies into the region of positive balance due (the loss domain). Specifically, I model the conditional distribution of negative balance due as

$$f(b|b < 0) = \frac{\sum_{i=1}^3 \frac{p_i}{\sigma_i} \phi\left(\frac{b-\mu}{\sigma_i}\right)}{\sum_{i=1}^3 p_i \left(\Phi\left(\frac{-\mu}{\sigma_i}\right) - \Phi\left(\frac{\underline{b}-\mu}{\sigma_i}\right)\right)}. \quad (8)$$

This equation defines a mixture of normal distributions, with the normal PDF and CDF denoted with  $\phi$  and  $\Phi$ , respectively. A common mean is assumed to ensure the estimated distribution is symmetric.<sup>15</sup>  $p_i$  denotes the mixing probabilities. The denominator ensures that this conditional distribution integrates to 1 on its restricted range.<sup>16</sup> I estimate parameters for this model via maximum likelihood and overlay the predicted frequencies in figure 2. When restricted to the negative values from which it was estimated, this mixture model serves as a reasonable approximation of the balance-due distribution. It exhibits substantial excess mass precisely at zero, and frequencies in the loss domain are substantially lower than would be forecasted from the remainder of the distribution. This figure provides an informal, but visually compelling, first-look at the key data features consistent with a loss-averse model. The following subsections will statistically assess these features and structurally estimate the magnitude of the sheltering response they imply.

---

<sup>14</sup>The appearance of points of diffuse excess mass (approximately at  $-\$200$  and  $-\$400$ ) is driven by the tax returns of individuals with simple taxable behavior, filing forms 1040A or 1040EZ. The distribution restricted to standard 1040s, which compose 66% of my sample, is effectively completely smooth, with the exception of the discontinuity predicted by loss aversion. This distributional smoothness among returns with more complex taxable behavior can also be seen in the higher-income quartiles in figure 4.

<sup>15</sup>A structural approach permitting distributional asymmetry will be discussed in section 3.2.

<sup>16</sup> $\underline{b}$  is the lowest value of balance due considered, set to  $-1700$  in figure 2.

### 3.1 Bunching-based estimates of loss-averse sheltering

To quantify the excess frequency of individuals reporting zero balance due, and to quantify the loss-averse sheltering it implies, I proceed in a manner motivated by the empirical approach of Chetty, Friedman, Olsen, and Pistaferri’s (2011) study of labor-supply elasticity. I fit the frequency distribution of balance due, restricted to a region near the hypothesized gain/loss threshold, as a seventh-order polynomial. I additionally allow for excess mass precisely at zero balance due and a discontinuity when transitioning from a gain to a loss. In contrast to figure 2, where balance due is presented in nominal terms, here balance due is expressed in 1990 dollars to permit the quantification of the increased sheltering in units comparable over time. All such unit conversions are calculated from the consumer price index.

Formally, I estimate

$$C_j = \alpha + \left[ \sum_{i=1}^7 \beta_i \cdot b^i \right] + \gamma \cdot I(b_j = 0) + \delta \cdot I(b_j > 0) + \epsilon_j \quad (9)$$

In this equation,  $j$  indexes each dollar bin of balance due  $b$  from -100 to 100, with corresponding counts  $C_j$ . The polynomial approximates a smooth distribution of  $b$ , although it allows for a discontinuity at zero through the inclusion of  $\delta$ . A indicator variable for zero balance due,  $I(b_j = 0)$ , is also included, and the corresponding coefficient  $\gamma$  estimates the excess frequency of observations in this bin. Estimates of this model are reported in column 1 of table 1, and the predicted distribution is graphed over a frequency histogram of the data in figure 3.

As is predicted by the loss-averse sheltering model, and as is visually apparent in figure 3, the estimate of  $\gamma$  indicates a large and statistically significant excess mass precisely at zero balance due, and the estimate of  $\delta$  indicates a statistically significant downward shift of the distribution in the loss domain. While the degree of the polynomial fit was chosen to match a primary specification of Chetty, Friedman, Olsen, and Pistaferri (2011), the statistical significance of  $\hat{\gamma}$  and  $\hat{\delta}$  is similar or stronger for any polynomial of degree one through ten.

To assist in assessing the economic significance of these estimates, the excess mass at zero may be used to generate bounds on the increase in sheltering induced by loss aversion.



The distribution of balance due generated by the loss-averse sheltering model implies that the mass present at the reference point, estimated by  $\hat{\alpha} + \hat{\gamma}$ , is given by the integral of the “shelter to zero” region in figure 1. The width of the integrated region is  $(s^H - s^L)$ ; as a result, inferring the width of this region is sufficient to calculate the additional sheltering pursued in the loss domain. However, the density which is being integrated is unobserved over this region, and thus the precise parameters of this integral are unidentified. To proceed, I assume that  $f_b^{PM}(\cdot)$  is decreasing on this interval—as is suggested by figure 2—which permits me to generate an informative partial-identification region. As is illustrated in appendix figure A.1, the mass at zero can be no larger than the width of the region times the density on the left, and no smaller than the width of the region times the density on the right. These inequalities define a narrow range of values of  $(s^H - s^L)$  that could rationalize the observed distribution. Formal derivations of these bounds, accounting for the discreteness resulting from rounding to the nearest dollar, are presented in appendix A.

Bunching-based bounds on  $(s^H - s^L)$  are presented in the lower panel of table 1. Since these bounds are nonlinear functions of regression coefficients, standard errors are calculated using the delta method (bootstrapped standard errors are reported in appendix table A.3, and yield similar results). These estimates partially identify  $(s^H - s^L)$  within the narrow range  $[1.37, 1.83]$ , suggesting that loss aversion motivates approximately \$1.5 of additional sheltering among individuals in the loss domain.

The remaining columns of table 1 repeat this exercise while restricting the data to different adjusted gross income (AGI) quartiles. Since the income distribution changes over time, observations are assigned into quartiles based on year-specific income distributions; this practice will be maintained throughout the paper. While a significant amount of excess mass is seen at zero across all four quartiles, it is clear that this bunching behavior is markedly more pronounced among high-AGI tax filers. The distribution of the top year-specific AGI quartile exhibits 247% excess mass at zero balance due relative to the frequency predicted from the gain domain (95% confidence interval:  $[180\%, 315\%]$ ), substantially higher than the excess mass in the first three quartiles (139%, 98%, and 96% respectively).<sup>17</sup> As a result,

---

<sup>17</sup>These estimates are generated from the coefficients in table 1 by calculating the ratio  $\frac{\hat{\alpha} + \hat{\gamma}}{\hat{\alpha}}$ . The confidence interval is calculated with the delta method.

the estimated partial identification region of extra sheltering motivated by loss aversion is notably higher—[2.48, 3.97]—when estimated from the top income quartile. This pattern of heterogeneity is to be expected, since high-income filers have greater access to tax shelters, and particularly to the finely manipulable tax shelters that facilitate the precise targeting of a specific balance-due amount.

### 3.2 Shifting-based estimates of loss-averse sheltering

The analysis of the previous subsection focused on confirming and quantifying the bunching prediction, formalized in proposition 2. We will now focus attention on the shifting prediction, formalized in proposition 3.

Recall that proposition 3 described the precise structure of the balance-due distribution, focusing attention away from the hypothesized reference point:

$$f_b(x) = \begin{cases} f_b^{PM}(x + s^L) & \text{if } x < r \\ f_b^{PM}(x + s^L + \tilde{s}) & \text{if } x > r \end{cases}. \quad (10)$$

$\tilde{s}$  denotes the difference between the high and low sheltering amounts,  $(s^H - s^L)$ . Directly estimating this parameter gives an additional means of quantifying the excess sheltering pursued in the loss domain.

Estimation of this density may be conducted with reasonably minimal parametric restrictions on the structure of  $f_b^{PM}$ . To see why, notice that for all  $b \notin [s^L, s^H]$ , equation 10 links the unobserved distribution  $f_b^{PM}$  and the observed distribution  $f_b$ . Given  $s^L$  and  $s^H$ , the shape of  $f_b^{PM}$  is nonparametrically identified over most of its support. This matches the intuition graphically expressed in figure 1: we effectively observe the shape of the unmanipulated distribution, *except* for the shaded “shelter to zero” region. Parametric restriction are needed to proceed with estimating this structure, but only to ensure that identifying the distribution for  $b \notin [s^L, s^H]$  is sufficient to identify the full distribution.

To proceed with the estimation of equation 10, I assume that the full distribution can be modeled as a symmetric mixture of normal distributions:  $f_b^{PM}(b) = \sum_{i=1}^2 \frac{p_i}{\sigma_i} \phi\left(\frac{b-\mu}{\sigma_i}\right)$ . Symmetry is imposed by assuming a common mean for both mixing distributions. This

functional form is quite flexible, but sufficiently parameterized to imply a unique structure over the “shelter to zero” region given knowledge of the rest of the plot. I use this framework to fit the frequency distribution of balance due in a manner analogous to the polynomial fit of the bunching-based approach. The resulting nonlinear least squares specification is given by:

$$C_j = Obs \cdot \left[ \sum_{i=1}^2 \frac{p_i}{\sigma_i} \phi \left( \frac{b_j + \tilde{s} \cdot I(b > 0) - \mu}{\sigma_i} \right) \right] + \epsilon_j. \quad (11)$$

In this equation,  $j$  indexes each dollar bin of balance due  $b$  from -4000 to 4000, with zero excluded.  $C_j$  indicates frequency counts.  $Obs$  indicates the total number of observations, and serves to rescale the density to an approximate frequency distribution.

Figure 4 plots this predicted frequency distribution, and table 2 reports the estimated parameters. As seen in the first column, the full-sample parameter estimate of  $\tilde{s}$  is 389 (standard error: 10). This represents a positive and strongly statistically significant shift in the loss domain, consistent with the predictions of the loss-averse sheltering model. Under the interpretation in that model, these results would suggest that taxpayers pursue \$389 of additional sheltering when confronted with loss framing (95% confidence interval: [\$369, \$410]).

As seen in figure 4, the fitted model estimated from the full sample is visually well aligned with the a narrow-bandwidth kernel regression. However, the goodness-of-fit appears comparatively weaker in the early region of the loss domain, which is undesirable for the identification strategy pursued. This problem is resolved when heterogeneity across income levels is permitted. The lower half of figure 4 presents fitted distributions specific to each AGI quartile, and columns 2 - 5 of table 2 report the estimated model parameters. Across all four specifications, the fitted model closely matches the kernel-smoothed distribution, and each specification estimates a statistically significant shift across the loss domain. The excess sheltering pursued in the loss domain is estimated to be \$36, \$70, \$184, and \$586 across the first through fourth income quartiles, respectively. As was seen with the bunching-based estimates, the behavior predicted by the loss-averse model is substantially more pronounced among high-income filers. As seen in figure 4, the predicted models closely match

a narrow-bandwidth kernel fit of the data, although fit is noticeably better after accounting for heterogeneity across income quartiles.

Despite the good observed fit of these models, one might still reasonably question the validity of the assumption of a symmetric distribution. An alternative parametric approach, allowing for skewness, produces similar full-sample results. Appendix table A.5 reports the results of a NLLS specification analogous to 11, but assumes that  $f_b^{PM}$  follows a “skew-normal” distribution. The resulting full-sample estimate of  $\tilde{s}$  is \$406 (standard error: \$9), yielding a precisely estimated result of similar economic magnitude as the symmetric approach. However, the fit of this predicted model is notably worse, due to its overall less flexible structure. For further information, see appendix table A.5 and appendix figure A.2.

Similar to the bunching-based results, these estimates are consistent with loss-averse sheltering, and this behavior is estimated to be more pronounced among high-income taxpayers. However, while these qualitative results are consistent across the two approaches, the estimated magnitudes generated from the shifting-based approach are substantially larger. This discrepancy will be considered in the following subsection.

### 3.3 Comparing the bunching-based and shifting-based approaches

Thus far we have assessed the predictions of the loss-averse sheltering model that were illustrated in propositions 2 and 3. These two propositions led to two complementary empirical approaches, each of which sheds light on structural parameters. The bunching-based and shifting-based approaches each suggest the presence of excess sheltering in the loss domain. However, the magnitude of the estimated effect is notably different. In this subsection, we will compare the advantages and disadvantages of the two different approaches, and explore a leading explanation for the difference in their magnitudes.

First, consider the bunching-based estimates. A crucial distinction of the bunching-based approach is its exclusive focus on data local to the gain/loss threshold. This focus has several advantages. It avoids the need for parametric restrictions on the shape of the pre-manipulation balance-due distribution, instead relying on the weak assumption that this distribution is decreasing across the relevant region. Furthermore, focusing on comparisons local to the reference point helps alleviate concerns that factors outside of the model are

driving the results, as long as those external factors would evolve smoothly with small changes in balance due. These advantages of restricting attention to local comparisons come at a cost. This approach focuses on a small fraction of the total population, which naturally limits the power and precision of statistical estimates, and permits the worry that the estimation sample is non-representative.

In contrast, the shifting-based estimates fundamentally involve fitting a full distribution. This dramatically increases statistical power and precision, and avoids the worry that the estimation sample is non-representative. However, this expanded view comes at a cost: these shifting-based estimates require stronger assumptions regarding the structure and parameterization of the unmanipulated distribution, and are inherently less robust to concerns of omitted model components. The comparative strengths and weaknesses of the two approaches should be kept in mind while comparing their results, and will be important when assessing potential confounding factors in section 4.

What can explain the discrepancy in magnitude between these estimates? A leading candidate is a strong assumption made and discussed in the theory of section 1.1; specifically, the assumption that all taxpayers have access to finely manipulable tax shelters, permitting to-the-dollar targeting of total sheltering. Under the weaker assumption that tax shelters come in discrete units, loss aversion still motivates greater pursuit of shelters among those facing a loss, and the prediction of a distributional shift in the loss domain holds. However, the ability to manipulate balance due exactly to the reference point relies critically on to-the-dollar targeting, and thus the prediction of bunching at zero would be altered. A model with discreteness in available shelters generates a diffusion of excess mass near the reference point, not necessarily a point mass. If relatively few taxpayers have true, to-the-dollar manipulation ability—as can realistically be expected—then estimates identified solely from the excess mass precisely at the reference point could dramatically understate the true behavioral response.<sup>18</sup>

---

<sup>18</sup>While precisely determining which individuals have access to finely manipulable shelters is challenging, several broad classifications are informative. In my primary sample, 34% of taxpayers file 1040A or 1040EZ, simplified forms with fewer categories of potential sheltering available. 27% of tax filers have adjusted gross income consisting of solely their wage or salary income, and thus lack the discretion in reporting practices that comes with more complex income sources. 62% did not itemize deductions. In short, reasonable fractions of the population clearly have few significant sheltering opportunities available.

To inform our understanding of violations of the fine manipulability assumption, I present alternative estimates that simultaneously fit both the shift in the loss domain and the excess mass due to bunching. Unlike in previous sections, I no longer assume the excess mass is precisely at the reference point. Instead, I assume that the excess mass near zero balance due is arbitrarily distributed over the interval  $[-\frac{w}{2}, +\frac{w}{2}]$ , where  $w$  denotes the width of this “bunching range.” Varying the width of this bunching range can be interpreted as varying assumptions on precision of manipulability in available tax shelters. Using this structure, I implement a maximum-likelihood approach to estimating the full balance-due distribution. Outside of the bunching range, this distribution will follow the structure dictated in proposition 3. As in earlier sections, this distribution will be approximated with a mixture of normal distributions, constrained to have a common mean to preserve symmetry. Inside the bunching range, however, the distribution is unknown, but is required to rationalize this region’s empirical mass. When interpreting the resulting fitted model, the difference between the empirical distribution and the predicted distribution inside the bunching range is attributed to imprecise bunching near zero. For a formal derivation of the likelihood function, see appendix A.

Figure 5 plots these fitted models, and reports their implied shifting parameters. Under the alternative assumptions that excess mass must fall within \$100, \$200, \$300, or \$400 of zero, the fitted models produces shifting estimates of \$53, \$141, \$272, and \$382, respectively. Thus, while all of these techniques clearly suggest that meaningful loss-domain shifting is present, choosing a precise estimate of the magnitude of this shift, in the presence of diffuse bunching, is challenging. The magnitude of these estimate depends on the assumed “bunching width,” and little data is available to guide the choice of this parameter. Furthermore, as the bunching width is made large, the fitted model appears to attribute features of the data to loss-averse bunching which might reasonably be thought to be unrelated. For example, the widest bunching-width estimate in figure 5 attributes the sharp peak of the distribution to mistargeted bunching near zero, an implication that seems implausible. However, the narrower bunching-width estimates appear to capture features of the distribution that could reasonably be explained by a simple diffusion near zero, in line with the predictions of loss-averse sheltering.

These results assist in interpreting the differences seen between the bunching-based estimates of section 3.1 and shifting-based estimates of section 3.2. First, these estimates suggest that precise bunching at zero understates the behavioral response by ignoring the bunching “near” zero, which would be expected to be seen due to failures of the fine manipulability assumption. This supports the treatment of the estimates of table 1 as well identified, but extremely conservative, lower bounds on the true effect. In contrast, these diffuse-bunching results suggest that the shifting-based estimates of section 3.2 appear to overstate the true effect of loss framing, and are perhaps best considered an upper bound on the true magnitude of the response. Considerations of possible confounding factors, discussed in section 4, will also support this interpretation.

## 4 Alternative explanations of observed behavior

The results of the prior section demonstrate that the distinctive predictions of a loss-averse model are observed in U.S. tax records. To a degree, this already reflects a success of the loss-averse framework considered in section 1. In light of the results of this paper, and the substantial literature briefly reviewed in the introduction, it appears that the loss-averse model provides a tractable framework for rationalizing certain features of tax reporting behavior. However, the question remains: can the behavior documented in the previous section truly be attributed to the psychological sheltering motivations represented in this model? This section presents supporting analysis to assess that question. In the first subsection, I present evidence that suggests the observed results are driven by sheltering motives, as opposed to, e.g., well targeted tax prepayments. In the second subsection, I consider several alternative channels that could potentially influence sheltering motives. While these additional features of the decision environment likely influence the sheltering decisions I study, ultimately they appear unable to explain the full patterns observed in section 3.

### 4.1 Assessing evidence of a sheltering response

To test the prediction that a positive pre-manipulation balance due was offset with an increase in sheltering activity, we may directly examine the pursuit of observed categories

of tax-reducing provisions. After the calculation of total income on Form 1040, the tax filer may subsequently reduce this value by reporting adjustments to income and by claiming deductions. After the resulting tax is calculated, it may be further reduced by claiming tax credits. The amount claimed in each of these categories is higher for individuals owing a tax payment as compared to those owed a refund (see panel A of table 3). These differences are consistent with loss-averse pursuit of shelters; however, interpretation of this mean comparison is complicated by the non-random assignment of gain/loss status. To generate a sharper test of the theory, I turn to examining the sheltering behavior of those bunching at zero, as compared to similar individuals near zero.

In panel B of table 3, I regress these measure of tax reduction on a polynomial of balance due, while “dummying out” zero. Under an interpretation similar to that of the regressions of section 3.1, a positive estimated coefficient on this dummy variable would indicate an excess prevalence of tax-reduction activity at the reference point. To control for differences in pursuit of tax shelters across years and income levels, I additionally include year-specific fixed effects and a third-order polynomial of the prior year’s AGI. The estimates of columns 1, 3, and 5 indicate that the probability of reporting non-zero amounts in each sheltering category is discretely higher at zero balance due, as compared to the predicted values for those near zero balance due. While a positive effect is detected in all three comparisons, it is only statistically significant at conventional levels for the pursuit of adjustments to income. The estimates of columns 2, 4, and 6 indicate that conditional on the pursuit of the given category of tax reduction, the amount reported is higher for those at zero balance due, although the effect is statistically insignificant for the pursuit of credits. Overall, the evidence supports the prediction that bunching at zero balance due is associated with the pursuit of tax-reducing activities to a greater degree than relevant comparison groups.

Instead of focusing on end-of-the-year tax-reduction motives, one might instead attempt to explain features of the balance due distribution through “prepayments” such as withholdings and estimated tax payments. A subpopulation of conscientious taxpayers with well-targeted tax prepayments could generate an excess mass of filers reporting zero balance due on tax day. To assess this possible confound, I first examine whether the taxpayers reporting zero balance due are systematically well-targeted across the years of the sample. I find that they



are not. I then assess how their reported income aligns with what might be predicted based on other year's data. I find that those reporting zero balance due are having significantly atypical years, which makes to-the-dollar precision in tax withholdings implausible.

Among individuals who report zero balance due in my primary sample, only 6 report zero balance due more than once. For 79% of observations reporting zero balance due, the previous year's balance due is more than \$50 away from zero. Restricting attention to tax filers who report zero balance due at least once, the individual-specific sample average is more than \$50 away from zero for 81%; this number rises to 93% if we restrict attention to individuals with 3 or more observations in the panel. In short, the evidence does not suggest that the bunching behavior documented here is driven by taxpayers who systematically make very accurate tax prepayments.

While taxpayers reporting zero balance due are not systematically well-targeted across years, perhaps the years where they report zero are those where their tax liability is unusually easy to forecast. To explore this possibility, I make use of the panel nature of these data to forecast the growth of adjusted gross income. In table 4, I report estimates from models of the form

$$\Delta AGI_{it} = \alpha + \beta I(b_{it} = 0) + \mathbf{C}_{it}\Gamma + \epsilon_{it} \quad (12)$$

where  $\mathbf{C}\Gamma$  represents the included controls, such as taxpayer and filing-year fixed effects, amount of balance due, and lagged AGI. In all such regressions, standard errors are clustered by taxpayer. If the individuals at zero balance due are simply experiencing years with easy-to-forecast income, one would expect the coefficient on  $I(b = 0)$  to be zero, indicating no deviation from the average forecast. Under the predictions of the loss-averse sheltering model, one would instead expect the coefficient to be positive, representing a positive income shock. If this positive income shock were at least partially unanticipated, tax withholdings or estimated tax payments would be based on a forecast of final tax liability that is too low. As a result, the pre-manipulation balance due faced by the taxpayer would be positive. The loss-averse model predicts that those reporting zero balance due faced a positive pre-manipulation balance due, which they offset with the increased sheltering activity demonstrated above.

The first column of table 4 indicates that relative to all other taxpayers, individuals precisely at zero balance due report an additional \$4,088 of income growth, on average and expressed in 1990 dollars. In a manner similar to that used to study sheltering behavior above, column 2 estimates a model where  $\Delta AGI$  is a smooth function of balance due, but allows for a discontinuity at zero. The third column additionally allows the amount of AGI growth to depend on a third-order polynomial in last-year's AGI, which can capture the notion that a larger absolute amount of year-to-year income growth is expected among higher-income individuals. The fourth through sixth columns repeat these exercises with the inclusion of taxpayer-specific fixed effects. Across these regressions the estimated excess AGI growth ranges from \$3,736 to \$6,321. It is noteworthy that despite the large overall sample size, these estimates are still reasonably imprecise, as they are identified from the comparatively small number of taxpayers precisely at zero. While this imprecision means these reported magnitudes come with a fair degree of uncertainty, it is clear that all estimates are positive and statistically significant. Even the smallest values in their 95% confidence intervals suggest positive income shocks of an economically significant magnitude, giving strong evidence of relatively large income shocks being experienced by these bunching individuals.

In addition to having comparatively unpredictable AGI, individuals reporting zero balance due also have a propensity to earn income from sources that are comparatively difficult to forecast. 35% of zero-balance-due returns have income from schedule C (business income), schedule D (capital gains and losses), schedule E (royalties, partnerships, S corporations, rental real estate, etc.), or schedule F (farm income), all sources that are comparatively more unpredictable than simple wage or salary income. In contrast, only 26% of returns with non-zero balance due have income from one or more of these sources (difference of proportions p-value: 0.002). Regressions analogous to those in table 4, allowing for interactions with income source, indicate that shocks to these income sources are the primary drivers of documented AGI shocks (see appendix table A.6). In addition to supporting the notion that the bunching at zero is difficult to rationalize with accurate tax withholdings, this result also highlights an important potential mechanism. Schedules C through F are known to have comparatively high rates of income underreporting; prior research has explicitly used behavior in these income categories to conduct inference on tax evasion in the absence of

audit data (Feldman and Slemrod, 2007). While precisely identifying tax evasion is not possible in my data, the nature of income sources among bunching individuals is consistent with income underreporting playing a potentially important role.

To summarize, this section has presented direct evidence of a sheltering response on observable categories of tax-reducing activities, and has demonstrated that bunching at zero is associated with having a difficult-to-forecast income path. Overall, rationalizing these patterns with tax forecasting behavior is challenging, whereas an end-of-the-year sheltering response is strongly supported.<sup>19</sup>

## 4.2 Assessing alternative explanations for asymmetric sheltering

*Fixed costs incurred in the loss domain:* The model presented in section 1 considers tradeoffs in marginal sheltering benefits against marginal sheltering costs, and differences in those tradeoffs across the gain and loss domain. One might additionally consider the implications of fixed costs incurred in the loss domain, which introduces a discontinuity to the utility function itself (as opposed to marginal utility). Natural candidates for such a fixed cost include a perception that audit rates change between the gain and loss domain, the annoyance of having to write out a check, or a psychological aversion to any tax payment. Avoidance of fixed costs could generate excess mass precisely at zero balance due, but makes distinguishing predictions about the nature of the distribution in the loss domain. In an individual decision problem, a fixed cost would define a region of potential balance due values,  $[0, b^{max}]$ , for which the taxpayer would choose to shelter to zero to avoid the fixed cost. If pre-manipulation balance due did not fall within this region of sufficiently small tax payments owed, marginal incentives would not be affected, and the distribution would be no different than in the case with no fixed cost. For a population of identical individuals, this theory would suggest a region with no probability mass immediately to the right of zero, and no distributional shifting occurring outside of that region (for an illustration, see

---

<sup>19</sup>While it does not appear that the documented results are reflective of tax withholding behavior, notice that the presence of such behavior could itself be reflective of an alternative loss-averse model. In order to jointly rationalize both excess mass and the observed shifting of the loss domain, a model of withholding targeting must predict that individuals are discretely more likely to attempt precise withholding if they expect to face positive balance due. Such a discontinuity could itself be evidence of anticipation of loss aversion on tax day.

appendix figure A.3). Even in a more complex model, permitting heterogeneity in preference parameters and the presence of some individuals with no fixed costs, the presence of fixed costs would suggest diffuse missing mass immediately to the right of zero. Such a region of missing mass is not apparent in figures 2 through 5, and the fitted models of section 3.3 estimate *excess* frequency of reporting small losses. These results, and the inability of this model to explain the related shift of the loss domain, suggest that perceptions of a fixed cost are not driving the observed behavior.

*Financial constraints:* As explored in Andreoni (1992), financial constraints can incentivize tax noncompliance. For example, in the presence of binding borrowing constraints, tax evasion can serve as a risky substitute for a loan, implicitly trading income now for expected penalties in the future. Financial constraints generate a discontinuity in marginal incentives at the precise point where the borrowing constraint binds. To rationalize excess mass precisely at zero, this would require that a one dollar tax payment makes the difference between hitting the current-period borrowing constraint or not. This possibility, or the possibility that the threshold for the constraint binding is in the immediate vicinity of zero balance due, is perhaps a plausible explanation for the documented behavior in low income quartiles, and might explain some degree of the diffuse excess mass in these income groups. However, it is generally implausible that significant fractions of the high-income filers driving my primary results have a) completely exhausted their cash, liquid assets, and available borrowing technologies, and b) aligned such that the first marginal dollar spent on a tax payment is that which causes the borrowing constraint to bind. Uncertainty over the current period's spending, or smooth heterogeneity in savings or constraint values, is sufficient to rule out the latter condition, casting significant doubt that this mechanism plays a meaningful role in the observed behavior of high-income filers.

*Interaction with tax preparers:* A substantial fraction of tax returns are filed by a paid tax preparer on the taxpayer's behalf. In principle this could complicate the manner in which sheltering decisions are made. A simple reformulation of the baseline model in section 1,

accounting for this complication, is:

$$\max_{s \in \mathbb{R}^+} u_P(\underbrace{m_C(-b^{PM} + s) + c_C(s)}_{\text{utility of client}}) - \underbrace{c_P(s)}_{\text{costs to tax preparer}} \quad (13)$$

The subscripts  $P$  and  $C$  refer to the tax preparer and client, respectively. In this formulation, the tax preparer balances the utility gains from customer satisfaction (capturing concerns such as the increased probability of return service and good recommendations) and the costs of pursuing additional sheltering. If  $u_P$  is an increasing, concave, twice-differentiable function, the qualitative results of section 1 above hold without modification.

Interestingly, the use of tax preparers is particularly prevalent among individuals reporting zero balance due, even when controlling for prior AGI. An estimate analogous to those reported in table 3 indicates an increase in the probability of using a paid tax preparer of 15 percentage points relative to the gain domain, and 3 percentage points relative to the loss domain (standard error: 6 percentage points for both comparisons).

Tax returns filled out by a paid preparer are completed by an individual with significant knowledge and command of the tax code, leading to a richer variety of tax shelters that could be employed. Furthermore, using audited returns from 1979, Erard (1993) demonstrated that tax noncompliance is dramatically higher among individuals with CPA or lawyer-prepared returns, in contrast to self-prepared returns.<sup>20</sup> Under this interpretation, the use of paid tax preparers may be serving as a proxy for high pursuit of sheltering, in which case its discontinuity at zero balance due serves as further evidence in support of the loss-averse model.

*Avoidance of the underwithholding penalty:* The underwithholding penalty, and the discontinuity in the tax schedule it induces, can drive bunching behavior in tax sheltering activity. However, this penalty is not imposed until substantial underwithholding has occurred, exceeding a grace window bounded below by a percentage of total tax. As a result, the bunching behavior induced by this provision would not occur at zero, and cannot rationalize the observed results. Widespread misunderstandings of these withholding requirements,

---

<sup>20</sup>Weighted results from his subsample suggest that 39.2% of self-prepared returns understate their income, with a mean level of noncompliance of \$244. In contrast, 63% of CPA or lawyer-prepared returns understate their income, with a mean level of noncompliance of \$1,786.

such as the incorrect belief that any positive balance due leads to a penalty, could potentially explain the observed bunching and shifting. However, the fact that these behaviors are seen among professionally prepared returns alleviates concerns that misunderstandings of tax law are responsible.

## 5 Aggregate consequences of controlled framing

The results of section 3 provide two complementary approaches for conducting inference on  $s^H - s^L$ , the individual-level additional tax sheltering pursued in the loss domain. In this section, we will consider these estimates' implications for aggregate tax revenue.

The presence of loss-averse sheltering suggests that the IRS can influence aggregate tax sheltering through manipulation of gain/loss framing. Policies increasing the number of individuals facing loss framing will have psychological effects which increase total sheltering, and thus decrease total tax revenues collected on tax day. To illustrate the implied aggregate effects of the primary estimates of this paper, I calculate the aggregate revenue effect each estimate suggests would result from a shift of an additional 1% of all tax returns into the loss domain. Given an estimate of  $\tilde{s} = s^H - s^L$ , a simple calculation of this marginal aggregate effect is given by  $\frac{N}{100} \cdot \tilde{s}$ , where  $N$  denotes the number of returns filed in a given year. Below, I calculate this magnitude using the full-sample bunching-based and shifting-based estimates of section 3.

Since section 3.3 argued that the bunching-based estimate of  $\tilde{s}$  is best interpreted as an extremely conservative lower bound on the true effect, I base my estimates only on the lower bound of its identification region. Across the years of my sample, these estimates imply a lower bound ranging from \$2.18-\$2.68 million per year, measured in 2011 dollars. Given the growth in the tax-filing population that has occurred since 1990, these bunching-based estimates indicate a lower bound on the aggregate marginal effect of \$3.43 million in 2011. Effects of this magnitude are modest when compared with the total annual revenue from personal income taxes; however, as argued in section 3.3, there is reason to believe these bunching-based estimates substantially underestimate the true effect. Furthermore, these estimates are substantially less than those generated by shifting-based estimates. Turning

to the shifting-based estimates of section 3.2, estimates of the aggregate marginal effect range from \$620-\$762 million, again expressed in 2011 dollars. Applying the full-sample shifting estimate to the tax-filing population of 2011 suggests an aggregate marginal effect of \$974 million. Marginal effects on this order of magnitude suggest the potential for meaningful revenue effects resulting from changes to the fraction facing loss framing.

Some consideration of the limitations of these calculations is warranted. First, these calculations are solely considering the psychological response induced by a policy shifting balance due on tax day. Such a policy will certainly have non-psychological costs and effects. For example, if this shift were implemented by increasing or decreasing the rate of tax collection through withholdings, this could influence sheltering at different stages of the tax collection process. These effects are not included in the above calculations, and must be individually assessed for different potential policies meant to change the gain/loss composition. The calculations above are best considered as a psychological component which would be omitted in a standard calculation of such a policy effect. Second, these calculations are focused on small, marginal changes, whereas the scope for significant tax-revenue changes would come from non-local changes in the overwithholding rate.<sup>21</sup> The ability to scale up these marginal effects depends critically on the ability to modify the tax system without offsetting effects on the framing of the reference point. While this possibility is promising, it has yet to be demonstrated in this setting, and will be challenging to cleanly identify in the absence of large-scale experimental manipulation.

## 6 Discussion

In recent years we have seen great interest in transporting the insights of prospect theory into mainstream empirical economics; as argued in Barberis (2013), this enterprise is bearing fruit, but is still in its early stages. The results explored in this paper demonstrate a setting where a key component of prospect theory—loss aversion—productively informs our understanding of a centrally important economic field behavior. Beyond simply highlighting

---

<sup>21</sup>For example, changing withholding policies to eliminate the overwithholding phenomenon, and thus changing the fraction facing a gain on tax day from 78%—the empirical rate of negative balance due in my data—to 50%.

a psychological mechanism in play, the nature of the observed reaction to loss framing has important implications for tax policy, public finance, and behavioral economics. I discuss these applications below, and suggest paths for research moving forward.

The presence of loss-averse sheltering has several direct implications for the design of tax policy. First, these results suggest that changing the distribution of balance due can have a large, psychologically motivated impact on aggregate tax sheltering. In this respect, these results suggest that the phenomenon of overwithholding is even more beneficial to the IRS than previously documented, as it ensures that a comparatively small fraction of tax filers face the additional sheltering motivations associated with a loss. While this has long been recognized as an implication of experimental results supporting loss aversion (see, e.g., Shepanski and Shearer, 1995), this study offers unique field evidence that assists in quantifying these effects. The marginal effects estimated in section 5 indicate that substantially reducing the overwithholding rate could lead to increases in sheltering activity measured in billions of dollars. These psychologically motivated consequences should be taken into account when enacting policies that change the fraction of taxpayers overwithheld.

Second, these results have strong implications for the detection and deterrence of tax evasion. To the extent that illegal means are used to achieve the excess sheltering documented in this paper, targeted auditing of individuals in the loss domain—and especially those reporting zero balance due—assists in enforcing tax compliance.

Third, these results suggest that gain/loss framing can assist in controlling tax morale, and can be employed to reduce evasion or improve the efficacy of tax incentives. Conceptually, it may be possible to manipulate a taxpayer's perception of what constitutes a gain or a loss—potentially through relatively cheap manipulations to phrasing or presentation.<sup>22</sup> Loss framing could be manipulated at the individual level to, e.g., increase the take-up rate of a specific tax-based incentive in targeted populations. Gain framing could be induced to reduce evasion motives among traditionally noncompliant groups, potentially in a cost-effective manner when compared to audits. This promising possibility merits further research. For a recent review of related issues in tax morale, see Luttmer and Singhal (2014).

---

<sup>22</sup>For an attempt to influence the timely payment of UK taxes with gain/loss framing (among other behavioral interventions), see Hallsworth, List, Metcalfe, and Vlaev (2014).



Moving beyond the implications specific to tax policy, the techniques and the results put forth in this study contribute to the recent literature utilizing bunching-based approaches. Analysis of bunching is rapidly becoming a key identification strategy for understanding reference effects. Recently published papers have used such approaches to study the importance of round numbers as goals (Pope and Simonsohn, 2011) and effort provision in the lab (Abeler, Falk, Goette, and Huffman, 2011). Current research in progress uses similar approaches to study the goal-setting behavior of marathon runners (Allen, Dechow, Pope, and Wu, 2014) and job search (DellaVigna, Lindner, Reizer, and Schmieder, 2014). This paper makes a technical contribution to this literature by building upon the work of Chetty, Friedman, Olsen, and Pistaferri (2011) to generate an econometric framework which can detect and quantify latent loss-averse manipulation. The structural approach presented here can generally be applied when a variable is manipulated in response to piecewise-linear marginal incentives, and is easily transportable to other settings.

In the context of the broad study of reference-dependent behavior, results presented here inform an ongoing debate on the precise nature of the reference point. The loose specification of the gain/loss threshold has long been considered an undesirable degree of freedom in reference-dependent models. Recent research has focused on expectations-based reference dependence, which rationalizes a variety of empirical regularities and which successfully “closes the model” by endogenizing the reference point. Some empirical studies have found support for the expectations-based model (e.g., Crawford and Meng, 2011; Ericson and Fuster, 2011), while others have not (Heffetz and List, 2013). In the tax setting considered here, a simple model of the reference point, more in line with Kahneman and Tversky’s original presentation, provides significant insights into the observed behavior—insights which would not be explained by a rational-expectations-based model. However, while this paper has focused on evidence supporting a reference point of zero, evidence consistent with alternative reference points is also present in these data; indeed, diffuse bunching is observed around last years’ balance due (a potential status quo) and the person-specific average balance due (an expectations-based reference point in line with Crawford and Meng (2011)). While this is suggestive that other reference points might be in play, cleanly interpreting bunching along these dimensions as evidence of loss aversion is difficult, since similar behavior over time

is to be expected. Overall, the results of this study and the greater literature suggest the possibility of heterogeneity in the framing of reference points across contexts, and potentially even within context. As we continue to export the insights of prospect theory from the lab into the field, additional attention to the theoretical underpinnings of the reference point, as well as empirical techniques which help differentiate between candidate reference points, will prove essential.

## 7 Works cited

- Abeler, J., Falk, A., Goette, L., & Huffman, D. 2011. "Reference Points and Effort Provision." *American Economic Review*, 101(2): 470-492.
- Allen, E., Dechow, P., Pope, D., & Wu, G. 2014. "Reference-Dependent Preferences: Evidence from Marathon Runners." *Working Paper*.
- Andreoni, J. 1992. "IRS as Loan Shark: Tax Compliance with Borrowing Constraints." *Journal of Public Economics*, 49(1): 35-46.
- Barberis, N. 2013. "Thirty Years of Prospect Theory in Economics: A Review and Assessment." *Journal of Economic Perspectives*, 27(1): 173-196.
- Bernasconi, M. & Zanardi, A. 2004. "Tax Evasion, Tax Rates, and Reference Dependence." *FinanzArchiv*, 60: 422-445.
- Benzarti, Y. 2014. "How Taxing is Tax Filing? Leaving Money On The Table Because Of Compliance Costs." *SSRN Working Paper 2412703*.
- Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *The Quarterly Journal of Economics*, 112(2): 407-441.
- Carroll, J. S. 1992. "How Taxpayers Think About Their Taxes: Frames and Values." In *Why People Pay Taxes: Tax Compliance and Enforcement*, ed. Joel Slemrod, pp. 43-63.
- Chang, O. H., Nichols, D. R., & Schultz Jr, J. J. 1987. "Taxpayer Attitudes Towards Tax Audit Risk." *Journal of Economic Psychology*, 8: 299-309.
- Chetty, R., Friedman, J., Olsen, T., & Pistaferri, L. 2011. "Adjustment Costs, Firm Responses, and Micro vs Macro Labor Supply Elasticities: Evidence from Danish Tax

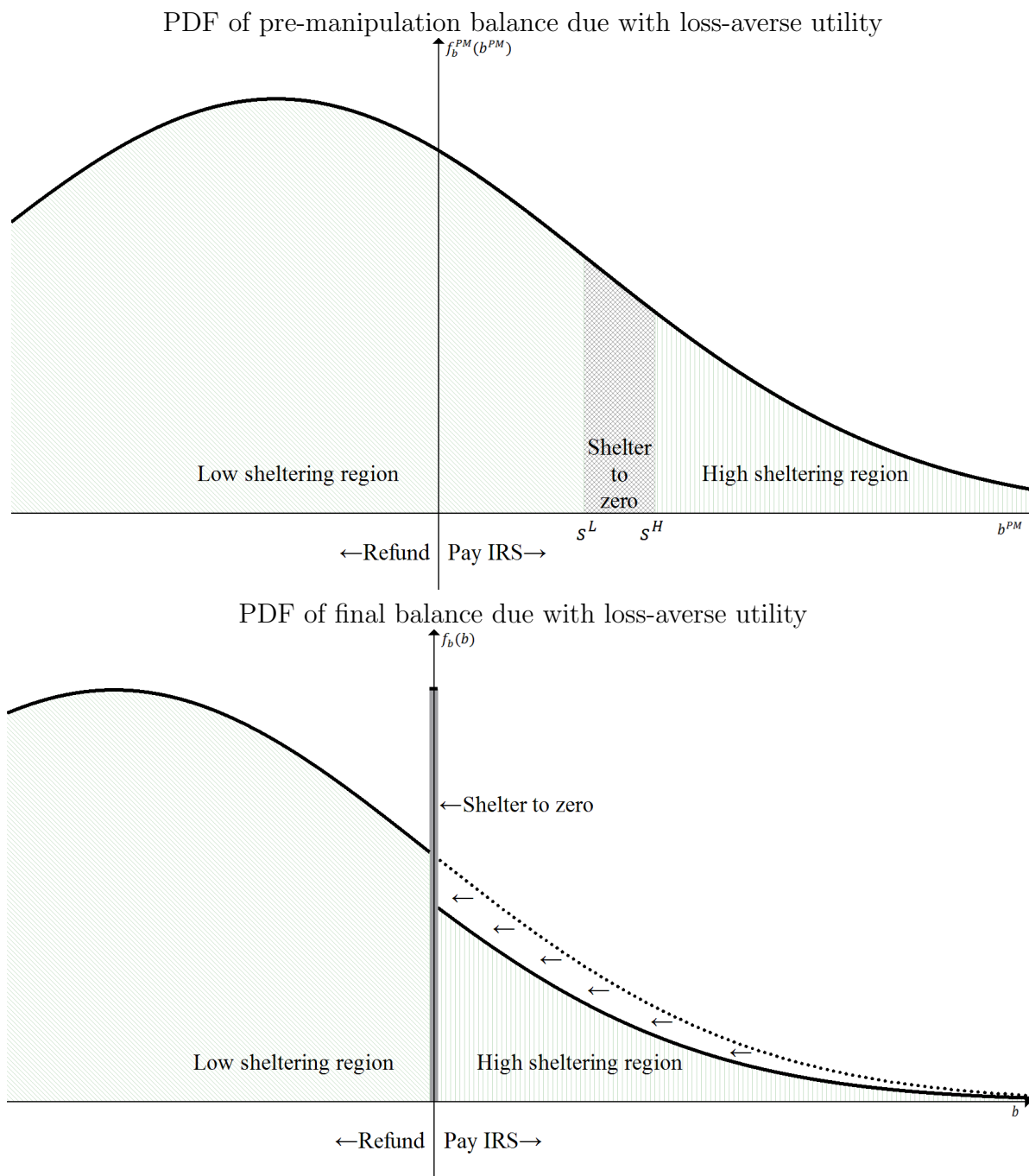
- Records.” *The Quarterly Journal of Economics*, 126(2): 749-804.
- Clotfelter, C.** 1983. “Tax Evasion and Tax Rates: An Analysis of Individual Returns.” *The Review of Economics and Statistics*, 65(3): 363-373.
- Copeland, P. & Cuccia, A.** 2002. “Multiple Determinants of Framing Referents in Tax Reporting and Compliance.” *Organizational Behavior and Human Decision Processes*, 88(1): 499-526.
- Crawford, V. P. & Meng, J.** 2011. “New York City Cabdrivers’ Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income.” *American Economic Review*, 101(5): 1912-1932.
- DellaVigna, S., Lindner, A., Reizer, B., & Schmieder, J.** 2014. “Reference-Dependent Job Search: Evidence from Hungary.” Working Paper.
- Dhmi, S. & al Nowaihi, A.** 2007. “Why Do People Pay Taxes? Prospect Theory versus Expected Utility Theory.” *Journal of Economic Behavior & Organization*, 64(1): 171 - 192.
- Dhmi, S. & al Nowaihi, A.** 2010. “Optimal Income Taxation in the Presence of Tax Evasion: Expected Utility versus Prospect Theory.” *Journal of Economic Behavior & Organization*, 75(2): 313-337.
- Elffers, H. & Hessing, D. J.** 1997. “Influencing the Prospects of Tax Evasion.” *Journal of Economic Psychology*, 18(2-3): 289-304.
- Engström, P., Nordblom, K., Ohlsson, H., & Persson, A.** 2013. “Tax Compliance and Loss Aversion.” Working Paper.
- Erard, B.** 1993. “Taxation with Representation: An Analysis of the Role of Tax Practitioners in Tax Compliance.” *Journal of Public Economics*, 52(2): 163-197.
- Ericson, K. & Fuster, A.** 2011. “Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments.” *The Quarterly Journal of Economics*, 126(4): 1879-1907.
- Feenberg, D., & Skinner, J.** 1989. “Sources of IRA Saving.” *Tax Policy and the Economy*, Volume 3, ed. Lawrence Summers, pp. 25-46.
- Feldman, N. & Slemrod, J.** 2007. “Estimating Tax Noncompliance with Evidence from Unaudited Tax Returns.” *The Economic Journal*, 117: 327-352.

- Feldman, N.** 2010. "Mental Accounting Effects of Income Tax Shifting." *The Review of Economics and Statistics*, 92(1): 70-86.
- Genesove, D. & Mayer, C.** 2001. "Loss Aversion And Seller Behavior: Evidence From The Housing Market." *The Quarterly Journal of Economics*, 116(4): 1233-1260.
- Hallsworth, M., List, J., Metcalfe, R., & Vlaev, I.** 2013. "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *NBER Working Paper No. 20007*.
- Heffetz, O. & List, J.** 2013. "Is the Endowment Effect an Expectations Effect?" *Journal of the European Economic Association*, 12(5): 1396-1422.
- Homonoff, T.** 2013. "Can Small Incentives Have Large Effects? The Impact of Taxes versus Bonuses on Disposable Bag Use." Working paper.
- Jones, D.** 2012. "Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds." *American Economic Journal: Economic Policy*, 4(1): 158-185.
- Kahneman, D. & Tversky, A.** 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263-291.
- Kanbur, R., Pirttilä, J., & Tuomala, M.** 2008. "Moral Hazard, Income Taxation, and Prospect Theory." *The Scandinavian Journal of Economics*, 110(2): 321-337.
- Kirchler, E. & Maciejovsky, B.** 2001. "Tax Compliance Within the Context of Gain and Loss Situations, Expected and Current Asset Positions, and Profession." *The Journal of Economic Psychology*, 22: 173-194.
- Kőszegi, B. & Rabin, M.** 2006. "A Model of Reference-Dependent Preferences." *The Quarterly Journal of Economics*, 121(4): 1133-1166.
- Kőszegi, B. & Rabin, M.** 2007. "Reference-Dependent Risk Attitudes." *American Economic Review*, 97(4): 1047-1073.
- Luttmer, E. F. P. & Singhal, M.** 2014. "Tax Morale." *Journal of Economic Perspectives*, 28(4): 149-168.
- Pope, D. & Simonsohn, U.** 2011. "Round Numbers as Goals: Evidence from Baseball, SAT Takers, and the Lab." *Psychological Science*, 22(1): 71-79.
- Pope, D. & Schweitzer, M.** 2010. "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes." *American Economic Review*, 101(1):

129-157.

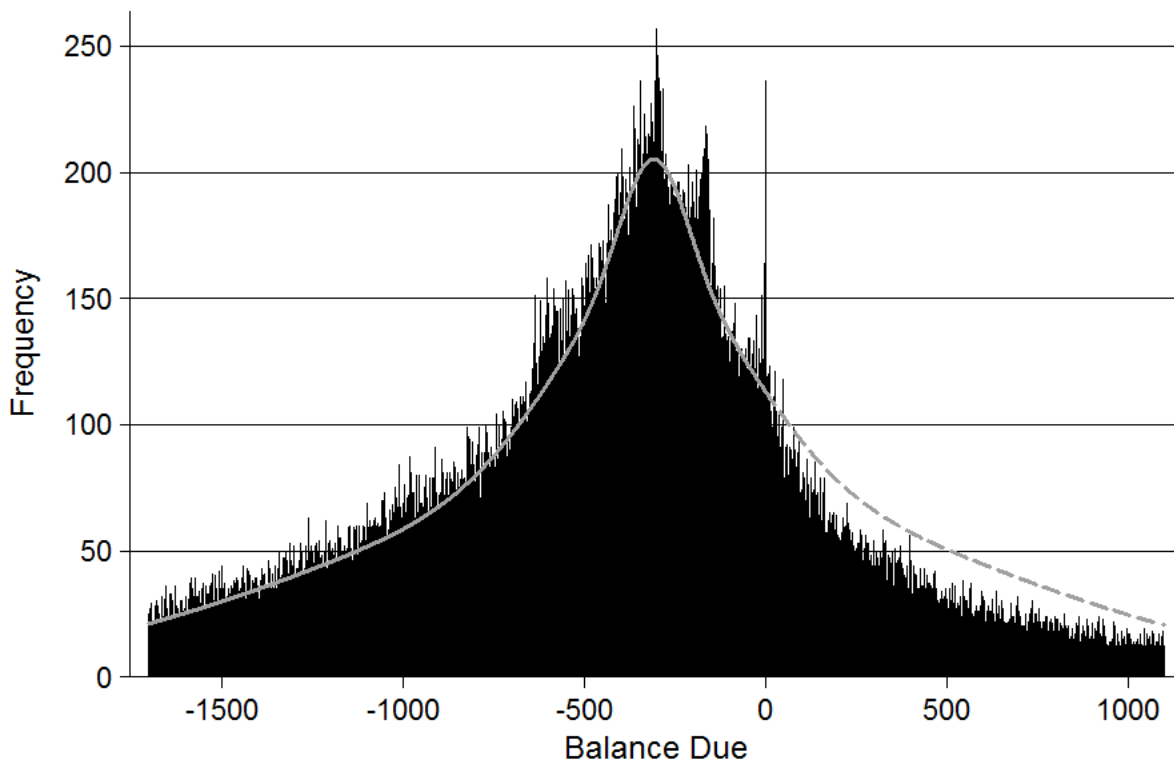
- Robben, H., Webley, P., Elffers, H., & Hessing, D.** 1990. "Decision Frames, Opportunity, and Tax Evasion: An Experimental Approach." *Journal of Economic Behavior and Organization*, 14: 353-361.
- Robben, H., Webley, P., Weigel, R., Wärneryd, K., Kinsey, K., Hessing, D., Martin, F. A., Elffers, H., Wahlund, R., Van Langenhove, L., Long, S., & Scholz, J.** 1990. "Decision Frames and Opportunity as Determinants of Tax Cheating." *Journal of Economic Psychology*, 11: 341-364.
- Saez, E.** 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2: 180-212.
- Schadewald, M.** 1989. "Reference Point Effects In Taxpayer Decision Making." *Journal of the American Taxation Association*, 89(10): 68-84.
- Schepanski, A. & Shearer, T.** 1995. "A Prospect Theory Account of the Income Tax Withholding Phenomenon." *Organizational Behavior and Human Decision Processes*, 63(2): 174 - 186.
- Yaniv, G.** 1999. "Tax Compliance and Advance Tax Payments: A Prospect Theory Analysis." *National Tax Journal*, 52(4): 753-764.

Figure 1: Implications of loss-averse utility for sheltering behavior



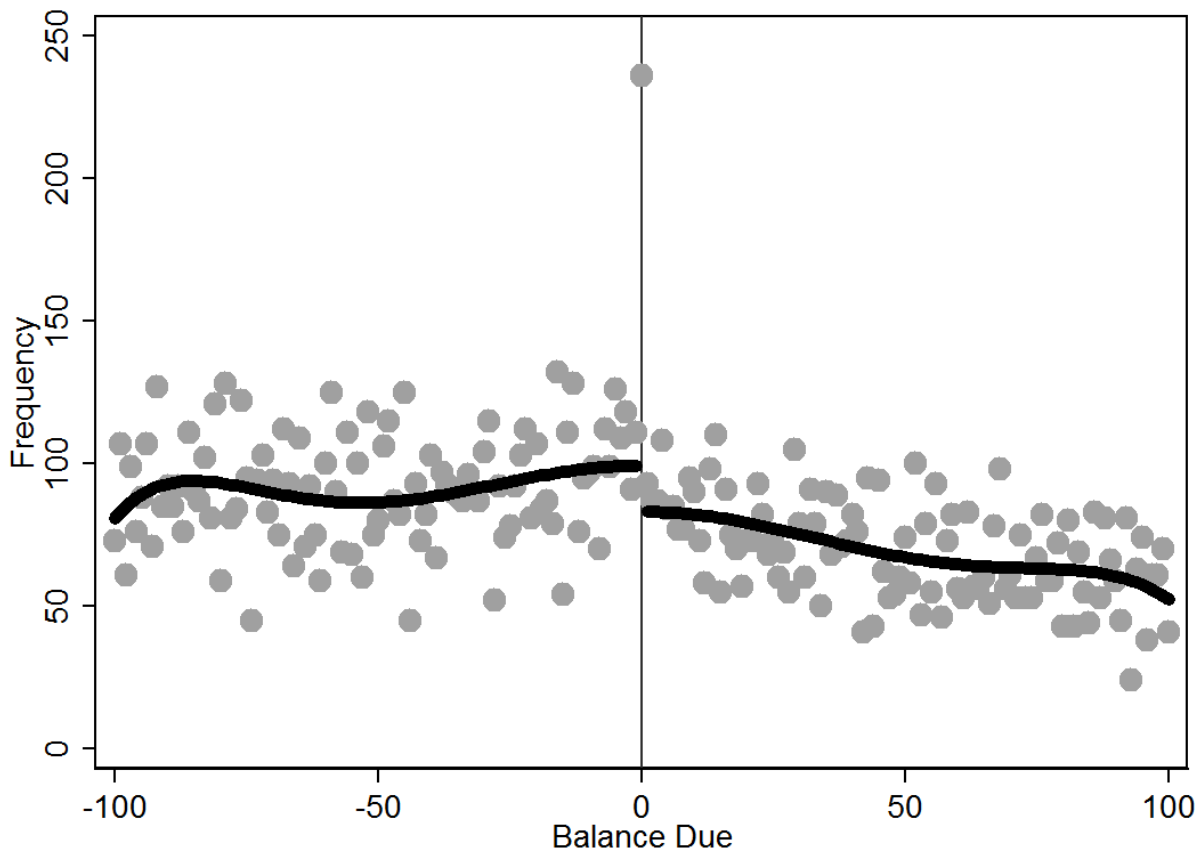
Notes: This figure highlights the central implications of loss-averse tax sheltering for the distribution of balance due. The first panel presents a hypothetical distribution of “pre-manipulation” balance due and indicates the optimal sheltering behavior from equation 4. The second panel indicates the final balance due that would be observed after loss-averse sheltering. The entire distribution is shifted to the left, with a fixed, larger shift for taxpayers with positive balance due. The darkly shaded region of taxpayers all shelter until reaching zero balance due, leading to a point mass in the observed distribution.

Figure 2: Nominal distribution of balance due



Notes: Histogram of balance due in \$1 bins. The graph is centered on  $-300$  with range restricted to  $[-1700, 1100]$ . The solid grey line plots a symmetric distribution fitted to the negative values of balance due. The dashed grey line indicates the extrapolated frequency from extending this symmetric distribution into positive values of balance due. Consistent with the loss-averse model, excess mass is observed at zero, and missing mass is observed in the loss domain (positive balance due). For details of the calculation of the fitted distribution, see equation 8 in section 3.

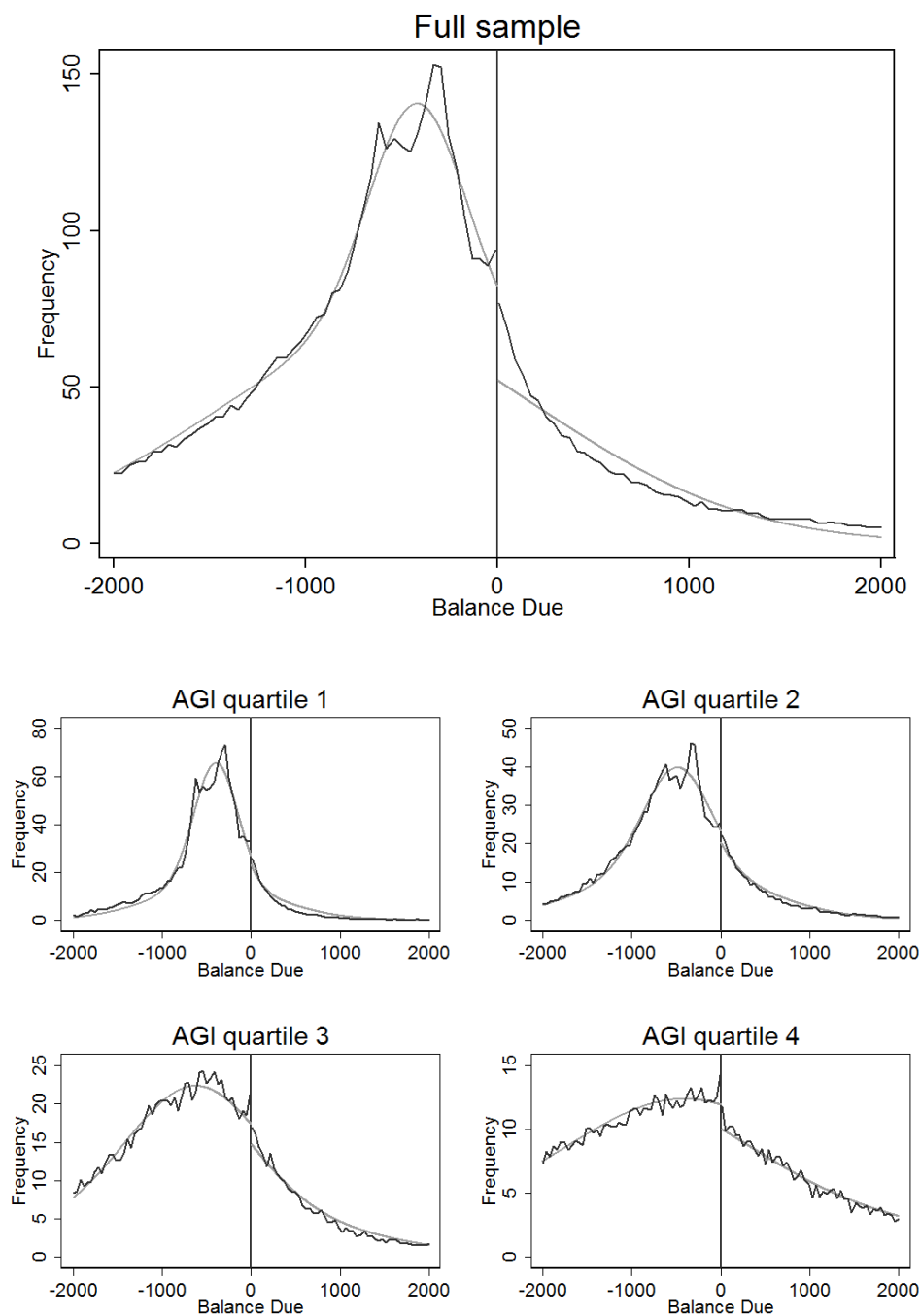
Figure 3: Distribution of balance due in vicinity of zero



Notes: Frequency of reported balance due amounts, expressed in 1990 dollars, in \$1 bins. Range restricted to  $[-100, 100]$ . The black line indicates the fit of a seventh-order polynomial, permitting a shift in the loss domain and a point discontinuity at zero. For details of the calculation of the fitted distribution, see equation 9 in section 3.

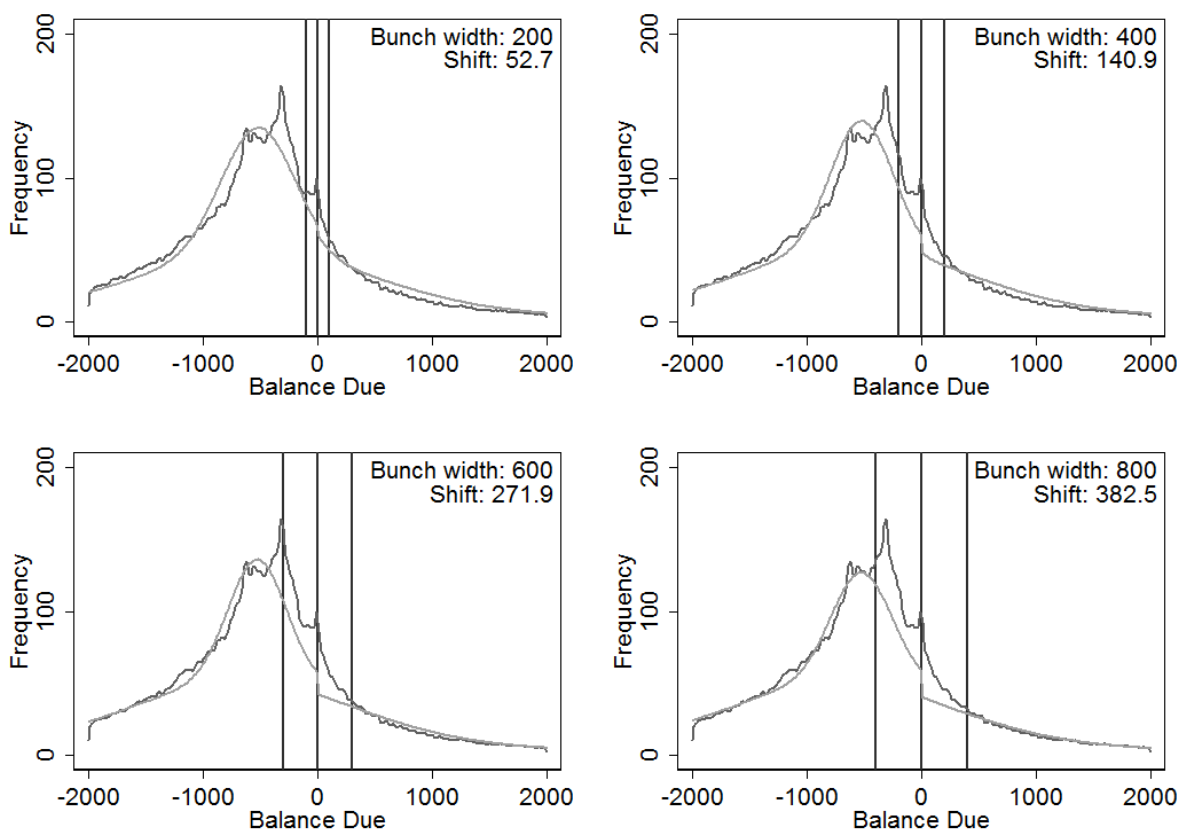


Figure 4: Fit of predicted mixture models



Notes: Plots of distributions fitted to the balance due frequency histogram. Balance due expressed in 1990 dollars, and rounded to \$1 bins. Grey lines indicate the estimated models from table 2, fitting symmetric distributions with a shift in the loss domain. For comparison, black lines indicate local-average kernel regressions (bandwidth: 10, kernel: Epanechnikov). Range of plot restricted to  $[-2000, 2000]$ , with zero excluded.

Figure 5: Fit of mixture models that permit diffuse bunching



Notes: Plots of the fit of the estimated mixture models of section 3.3. Resulting densities are multiplied by the sample size to rescale to frequency estimates. The fitted models assume the distribution takes the functional form predicted by loss-averse sheltering (see equation 5), but permit the excess mass to be arbitrarily distributed within  $\left[-\frac{\text{bunch width}}{2}, +\frac{\text{bunch width}}{2}\right]$ . For comparison, black lines indicate kernel density estimates (bandwidth: 10, kernel: Epanechnikov). This model interprets the difference between the grey and black lines within the bunching window as diffuse excess mass due to bunching near zero. Range of plot restricted to  $b \in [-2000, 2000]$ . Estimation sample for predicted models:  $b \in [-10000, 10000]$ . Balance due expressed in 1990 dollars.

Table 1: Estimates of excess mass at zero balance due

	(1)	(2)	(3)	(4)	(5)
All AGI groups	1st AGI quartile	2nd AGI quartile	3rd AGI quartile	4th AGI quartile	4th AGI quartile
$\gamma : I(\text{balance due} = 0)$	136.43*** (18.46)	46.57*** (8.25)	26.79*** (6.95)	21.06*** (5.66)	42.01*** (4.15)
$\delta : I(\text{balance due} > 0)$	-16.26* (9.41)	-3.50 (4.21)	-4.20 (3.54)	-3.42 (2.89)	-5.14** (2.12)
$\alpha : \text{Constant}$	99.57*** (5.45)	33.43*** (2.44)	27.21*** (2.05)	21.94*** (1.67)	16.99*** (1.23)
Balance due polynomial	X	X	X	X	X
$N$ : Bins in histogram	201	201	201	201	201
Observations	16348	5725	4553	3602	2468
$R^2$	0.490	0.479	0.259	0.209	0.489
Estimates of excess sheltering					
Bunching-based upper bound	1.83*** (0.28)	1.68*** (0.34)	1.34*** (0.35)	1.32*** (0.36)	3.97*** (0.60)
Bunching-based lower bound	1.37*** (0.22)	1.39*** (0.29)	0.98*** (0.29)	0.96*** (0.29)	2.48*** (0.34)

Notes: Standard errors in parentheses. Standard errors for excess sheltering estimates calculated with the delta method. Balance due expressed in 1990 dollars. Regression sample limited to observations with balance due  $\in [-100, 100]$ . For details of these regressions, see equation 9 in section 3. For a recreation of this table with bootstrapped standard errors, see appendix table A.3. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 2: NLLS estimates of shift in loss domain

	(1)	(2)	(3)	(4)	(5)
Full sample	1st AGI quartile	2nd AGI quartile	3rd AGI quartile	4th AGI quartile	4th AGI quartile
$\tilde{s}: s^H - s^L$	389.37*** (10.41)	35.64*** (6.45)	70.05*** (8.50)	184.22*** (15.51)	585.93*** (44.25)
$p_1$	0.80*** (0.00)	0.45*** (0.01)	0.58*** (0.02)	0.52*** (0.04)	0.64*** (0.05)
$\mu$	-419.08*** (2.53)	-392.41*** (1.66)	-487.87*** (3.19)	-621.93*** (7.62)	-412.11*** (24.49)
$\sigma_1$	1058.15*** (7.97)	702.44*** (17.74)	937.59*** (20.25)	1585.22*** (57.69)	2610.20*** (129.95)
$\sigma_2$	252.97*** (4.31)	246.84*** (3.28)	375.22*** (7.58)	741.72*** (20.03)	1218.48*** (51.65)
$N$ : Bins in histogram	8000	8000	8000	8000	8000
Observations	216227	57163	56799	55272	46993

Notes: Standard errors in parentheses. Balance due expressed in 1990 dollars. Reported are nonlinear least squares estimates of the model  $C_j = Obs \cdot \left[ \sum_{i=1}^2 \frac{p_i}{\sigma_i} \phi \left( \frac{b_j + \tilde{s} \cdot I(b > 0) - \mu}{\sigma_i} \right) \right] + \epsilon_j$ . In this equation,  $j$  indexes each dollar bin of balance due  $b$  from -4000 to 4000, with zero excluded.  $C_j$  indicates frequency counts.  $Obs$  indicates the total number of observations within the AGI group (including those with  $b \notin [-4000, 4000]$ ). To constrain the probabilities to  $(0, 1)$ , they are estimated in logistic form:  $p_1 = \frac{\exp(\theta_p)}{\exp(\theta_p) + 1}$ ,  $p_2 = \frac{1}{\exp(\theta_p) + 1}$ . To constrain standard deviations away from zero, they are estimated as  $\sigma_i = 10 + \exp(\theta_{\sigma_i})$ . To constrain the shift parameter to non-negative values, it is estimated as  $\tilde{s} = \exp(\theta_{\tilde{s}})$ . The table above converts estimates of  $(\theta_p, \theta_{\sigma_1}, \theta_{\sigma_2}, \theta_{\tilde{s}})$  into their implied values for  $(p_1, \sigma_1, \sigma_2, \tilde{s})$  and calculates standard errors with the delta method. For a recreation of this table with bootstrapped standard errors, see appendix table A.4. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 3: Assessing pursuit of tax reductions

	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Adjustments</i>		<i>Itemized Deductions</i>		<i>Credits</i>	
	> 0	Amount	> 0	Amount	> 0	Amount
Panel A: Differences in means of tax-reduction activities across gain and loss domain						
Mean in gain domain	0.198	2965.99	0.366	11506.70	0.330	310.47
Mean in loss domain	0.269	3061.25	0.448	11748.04	0.333	457.15
p-value of difference	0.000	0.040	0.000	0.000	0.256	0.000
Panel B: Regressions assessing pursuit of tax-reduction activities when reporting zero balance due						
I(Balance due = 0)	0.09*** (0.03)	1138.38* (619.59)	0.01 (0.03)	2015.49* (1112.42)	0.01 (0.03)	535.50 (493.06)
I(Balance due > 0)	0.05*** (0.00)	259.35*** (76.24)	-0.00 (0.00)	429.42*** (99.31)	-0.01*** (0.00)	27.97 (29.76)
Filing-year fixed effects	X	X	X	X	X	X
Balance due polynomial	X	X	X	X	X	X
Lagged AGI polynomial	X	X	X	X	X	X
<i>N</i>	148325	33935	148325	62441	148325	54223

Notes: This table assesses the pursuit of three classes of tax-reducing activities that can be claimed on Form 1040: adjustments to income, itemized deductions, and credits. For each of these variables, I separately consider the propensity to have any non-zero amount reported (expressed with a zero-one dummy variable), and the amount reported among those with positive pursuit. Monetary quantities are expressed in 1990 dollars. Panel A reports the mean of these variables separately for those reporting a gain (negative balance due) and a loss (positive balance due). I report p-values of two-sided difference-in-proportion tests in columns 1, 3, and 5, and two-sided t-tests in columns 2, 4, and 6. Panel B reports OLS regressions predicting pursuit of these tax-reducing activities. Standard errors are clustered by taxpayer. Xs indicate the presence of filing-year fixed effects, a third-order polynomial in lagged AGI, or a third-order polynomial in balance due interacted with  $I(\text{balance due} > 0)$  to allow for discontinuity at zero. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 4: Estimates of AGI shocks at zero balance due

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable : $\Delta$ AGI					
Balance due = 0	4088*** (1573)	3736** (1576)	6321*** (1665)	4235** (1959)	5389*** (1983)	5136*** (1913)
Balance due > 0		-636*** (137)	965*** (130)		664*** (170)	1511*** (139)
Filing-year fixed effects	X	X	X	X	X	X
Balance due polynomial		X	X		X	X
Lagged AGI polynomial			X			X
Taxpayer fixed effects				X	X	X
<i>N</i>	148325	148325	148325	148325	148325	148325

Notes: Standard errors, clustered by taxpayer, in parentheses. Monetary quantities expressed in 1990 dollars. Xs indicate the presence of filing-year or taxpayer fixed effects, a third-order polynomial in lagged AGI, or a third-order polynomial in balance due interacted with  $I(\text{balance due} > 0)$  to allow for discontinuity at zero. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .