NBER WORKING PAPER SERIES

HEURISTIC PERCEPTIONS OF THE INCOME TAX: EVIDENCE AND IMPLICATIONS FOR DEBIASING

Alex Rees-Jones
Dmitry Taubinsky

Working Paper 22884
http://www.nber.org/papers/w22884

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2016

We thank Doug Bernheim, Stefano DellaVigna, Emmanuel Farhi, Xavier Gabaix, Alex Gelber, Jacob Goldin, Alex Imas, Damon Jones, Erzo F.P. Luttmer, Kyle Rozema, Richard Zeckhauser, Jonathan Zinman and many seminar participants for helpful comments and advice. We thank Sargent Shriver and Linda Yao for excellent research assistance. We gratefully acknowledge research funding from the Pension Research Council/Boettner Center for Pension and Retirement Research at the Wharton School of the University of Pennsylvania, and from the Urban Institute. The opinions expressed in this paper are solely the authors', and do not necessarily reflect the views of any individual or institution listed above, nor the views of the National Bureau of Economic Research.

Heuristic Perceptions of the Income Tax: Evidence and Implications for Debiasing
Alex Rees-Jones and Dmitry Taubinsky
NBER Working Paper No. 22884
December 2016, Revised 2017
JEL No. D03,D04,D1,H0,K34

ABSTRACT

Using an incentivized tax forecasting task, we estimate the prevalence of previously discussed heuristics for constructing mental representations of nonlinear incentive schemes. We find strong evidence for “ironing” (linearizing the tax schedule using one’s average tax rate), no evidence for “spotlighting” (linearizing the tax schedule using one’s marginal tax rate), and we identify features of the remaining misperceptions that are not captured by existing models. We then embed these misperceptions in a standard model of income taxation and study their welfare consequences. We find that our estimated misperceptions increase social welfare because they are helpful in achieving redistributive goals.

Alex Rees-Jones
Operations, Information, and Decisions Department
The Wharton School
University of Pennsylvania
553 Jon M. Huntsman Hall
3730 Walnut St.
Philadelphia, PA 19104
alre@wharton.upenn.edu

Dmitry Taubinsky
Department of Economics
Dartmouth College
Rockefeller Hall
Hanover, NH 03755
and NBER
dtaubinsky@gmail.com

Financial incentives are often nonlinear and complex. As explored in models of energy markets, health insurance, and taxation, nonlinear incentives can be desirable in the presence of moral hazard, adverse selection, or the need for redistribution to the poor. Despite their benefits, nonlinear incentives can be cognitively demanding to process, and may invite misperception. Individuals may adopt heuristic approximations that are simpler to process, and rely on these approximations when decisions must be made. The full evaluation of the effects of a nonlinear incentive scheme must therefore incorporate a detailed empirical understanding of how incentives are perceived, since it is *perceived incentives* that ultimately guide individual behavior.

In this paper, we conduct such an evaluation in the context of income taxation. The nonlinear structure of the U.S. income tax schedule is designed to raise revenue progressively while minimizing distortions to labor supply. Empirical studies suggest that this complex system is imperfectly understood. Surveys of taxpayers demonstrate that many have inaccurate beliefs about key tax parameters, and that the average beliefs typically exhibit bias (Fujii and Hawley, 1988; Blaufus et al., 2013; Gideon, 2015). Studies of taxpayer behavior demonstrate that taxpayers modify their income in response to professional advice about tax incentives (Chetty and Saez, 2013) and in response to changes in lump sum transfers that do not affect marginal incentives (Liebman and Zeckhauser, 2004; Feldman et al., 2016), but systematically fail to modify their income to account for kinks in tax schedules (Saez, 2010; Chetty et al., 2013). Moreover, taxpayers appear to misreact to reforms occurring in complex systems, as demonstrated in lab experiments (Abeler and Jäger, 2015) and in field studies of changes to tax credits (Miller and Mumford, 2015).

The complexity of the tax code is often lamented because of the misunderstanding that it appears to induce (for a discussion, see Slemrod and Bakija, 2008). Yet whether taxpayer misunderstanding indeed harms social welfare remains an open question. To address this question, it is necessary to advance beyond broad demonstrations of taxpayer confusion. Instead, the analyst must have detailed knowledge of the alternative schedule that taxpayers believe is in place (as in, e.g., Farhi and Gabaix, 2015).

This paper reports the results of an experiment that provides the necessary direct quantification of individuals' misperceptions of the U.S. income tax schedule. We show that these misperceptions are driven by widespread adoption of simplifying heuristics, and that they are in fact substanitally welfare-enhancing in a standard model of taxation.

We administered our experiment in the tax season of 2015, recruiting 4,828 taxpayers to complete a series of incentivized questions about the tax that would be owed by a hypothetical taxpayer. This hypothetical taxpayer was constructed to be nearly identical to the experimental participant, except that the hypothetical taxpayer's income was varied across questions. This design identifies misperceptions of the complete tax schedule, and how these misperceptions vary with the individual's income, average tax rate (ATR), and marginal tax rate (MTR).

We find that, on average, taxpayers believe that the tax schedule is "flatter" than it truly is.

Our respondents overestimate the taxes owed by comparatively low-income filers and underestimate the taxes owed by comparatively high-income filers.[1] Evidence of flattening occurs for forecasts across the full tax schedule, and also for forecasts within subjects' own tax brackets—translating to an underestimation of marginal tax rates.[2]

To better understand these reduced-form results, we structurally decompose aggregate tax perceptions into the heuristics that generate them—specifically, the "ironing" and "spotlighting" heuristics of Liebman and Zeckhauser (2004). When applying these heuristics, the taxpayer approximates the nonlinear schedule by drawing on a single, salient feature of his tax bill. When applying the ironing heuristic, the taxpayer approximates the tax schedule as linear, with slope equal to his own ATR. When applying the spotlighting heuristic, the taxpayer linearly approximates the change in taxes induced by a change in income, applying his own MTR to the difference. Disentangling these heuristics from loosely specified misunderstanding of the shape of the tax schedule—as we do—would not be feasible with datasets that contain only local tax perceptions. However, our approach is identified and well-powered when estimated from perceptions of the full tax schedule.

Adoption of linearizing heuristics is prevalent in our experimental population. Estimates from our preferred specifications suggest that the ironing heuristic is adopted by 29-43% of tax filers. We estimate the prevalence of the spotlighting heuristic to be effectively zero, suggesting that this much-discussed heuristic is relatively unimportant in this environment. Among the remaining tax filers, we find that perceptions of the tax schedule are still "flatter" than the truth, preserving some of the qualitative predictions of ironing without being generated by over-reliance on the ATR.

Our results are broadly robust to a variety of issues that arise in survey research. Our empirical results persist when analyzing taxpayers with comparatively high incentives for accurate tax knowledge, such as the employed or those who complete their returns without outside assistance. They are also robust to alternative assumptions about respondents' ability to disentangle federal income taxes from other taxes that appear on pay stubs. They are robust to restrictions of the sample to those who do and do not perfectly resemble our hypothetical taxpayer. Finally, our results are robust to alternative means of screening inattentive respondents and outlier forecasts.

In the second part of the paper, we analyze the implications of these mistaken perceptions for social welfare. We show theoretically that a social planner would often avoid correcting the misperceptions that we document. This is for two reasons. First, the social welfare loss due to

[1] This result is consistent with the findings of Blaufus et al. (2013), who ask German respondents about income tax liability of unmarried German individuals who either have very high income (300,000 or 2,000,000 EUR) or low income (40,000 EUR or 10,000 EUR). Blaufus et al. (2013) find underestimation of the ATR faced by the very high income individuals. This is also consistent with Gideon (2015), who reports on a survey question from the Cognitive Economics Study that elicits taxpayers' perceptions of MTRs in the highest income tax bracket in the US, and finds underestimation. Gideon (2015) also reports on the other two tax questions in the Cognitive Economics Study, which elicit individuals' perceptions of their own marginal tax rates and average tax rates. Although overall underestimation of progressivity is not formally tested, a graphical summary of his results appears to be consistent with our finding.

[2] This result is consistent with Fujii and Hawley (1988), who conduct a survey that directly asks each individual about his marginal tax rate and find underestimation of MTRs.

individual misoptimization is offset by a fiscal externality: the additional tax revenue raised due to incorrect beliefs. Second, because ironing leads high-income tax filers to underestimate tax rates the most in absolute terms, correcting existing misperceptions constitutes a regressive policy. Consequently, the planner's optimal level of misperceptions is non-zero, and nudges that reduce biased behavior will not be utilized unless they can be applied in a targeted manner to low-income tax filers.

We examine the predicted impact of debiasing interventions in a simulation exercise and find that the quantitative impact on welfare is substantial. Across a wide range of parameter values for labor supply elasticities and the value of public funds, we find that the social planner would willingly pay between 0.9 and 4.4% of total tax revenue to avoid eliminating these misperceptions. These estimates translate to 32-156 billion dollars when applied to 2016 projected tax revenue.[3] While this range can be influenced by alternative modeling decisions, the qualitative presence of large welfare losses persists across the many variants of modeling assumptions that we explore.

This paper relates and contributes to recent advances in model-based evaluation of nudges. Economic evaluation of nudges—and the mistakes that nudges correct—has typically focused on direct analysis of behavior. In such applications, it is taken as given that, e.g., healthier eating or the cessation of smoking is desirable, and nudges are evaluated according to their success in achieving that goal. Recent papers have begun to explore more complete evaluations of the welfare consequences of nudges, taking into account considerations such as the psychic costs of being nudged (Allcott and Kessler, 2015), or the interactions of mistakes with broader economic objectives. For example, Chetty et al. (2009) consider the welfare consequences of average inattention to sales taxes, and find that the reduction in distortion it induces can substantially reduce the efficiency costs of taxation.[4] In applications beyond taxation, recent studies demonstrate that behavioral frictions in health insurance markets can play a crucial role in combating adverse selection (Handel, 2013; Handel and Kolstad, 2015; Handel et al., 2015; Spinnewijn, Forthcoming), or moral hazard (Baicker et al., 2015), rendering nudges undesirable. Our findings contribute to the growing understanding that nudges can be welfare-reducing—at times, dramatically so—despite their elimination of individually harmful mistakes.[5]

The rest of this paper proceeds as follows. Section 2 describes the experiment designed to measure tax misperceptions. Section 3 presents our empirical results. Section 4 theoretically analyzes the redistributive implications of the misperceptions we find, and presents simulated estimates of the impact of debiasing policies. Section 5 concludes. Supplemental analyses are available in the

[3]We apply the Tax Policy Center's projection of a total revenue of $3,525,179,000.
Source: http://www.taxpolicycenter.org/statistics/amount-revenue-source.

[4]However, in our own recent work (Taubinsky and Rees-Jones, 2016) we document that the inefficient sorting induced by heterogeneous mistakes can substantially offset this average effect.

[5]Relatedly, Bordalo et al. (2015) show how reminders can backfire by making some dimensions of a decision overly salient. This provides a psychological channel by which seemingly innocuous nudges can decrease the efficiency of consumer choice.

online appendices.

1 Experimental Design

We administered our experiment during the tax season of 2015. From March 15th through May 17th, subjects were recruited for a brief[6] web survey hosted on the Qualtrics platform, with recruitment targeting similar sample sizes in all weeks of this sampling window. Subject recruitment was managed by ClearVoice Research, a market research company that maintains a large, national population of subjects willing to take brief online surveys.[7] Subjects were recruited based on demographic data previously provided to ClearVoice, allowing us to generate a sample with demographics that approximate the national age, income, and gender distribution found in the U.S. census records (for tabulations of demographics in our sample and the census, see appendix table A1).

1.1 Experimental Protocol

The Qualtrics survey featured 4 modules. Screenshots of the full experiment are available in the web appendix; we summarize the contents here.

Introductory Module: The first module elicited basic information about subjects' tax filing behavior in order to facilitate the creation of a similar hypothetical tax filer in the forecasting module. Subjects were asked if they had already filed their tax return; who completed (or would complete) that tax return; their filing status; their exemptions claimed; if they claimed the standard or itemized deduction; their total income; if they filed each of schedule B through F; if they used TurboTax or similar software; if they or their spouse were born before January 2, 1950; and if they claimed the Earned Income Tax Credit.[8] Additionally, subjects were asked their degree of confidence in the key parameters determining their tax: their filing status, their exemptions, their deduction status, and their income. Confidence in these parameters was high. Given ratings options of "very confident," "somewhat confident," and "not confident at all," 96% of subjects were "very confident" in their filing status; 89% of subjects were "very confident" in their number of exemptions; 90% of subjects were "very confident" in their deduction status; 71% of subjects were "very confident" that their total income reported was within $1000 of being correct.

Forecasting Module: The primary questions were contained in the forecasting module. Subjects were presented with a variant of the following prompt, describing a hypothetical taxpayer whose filing behavior was very similar to their own:[9]

[6] Median completion time: 16 minutes. Interquartile range: 11-25 minutes.

[7] For other economic research making use of the ClearVoice panel, see Benjamin et al. (2014) or Taubinsky and Rees-Jones (2016).

[8] Subjects who claimed the Earned Income Tax Credit completed an additional brief battery of questions regarding their understanding of this tax provision.

[9] For subjects who had not yet completed their tax return, the verb tense was changed from past to future as

> This next group of questions is about Fred, a hypothetical taxpayer who is very similar to you. Fred is your age, and has a lifestyle similar to yours. Fred filed his 2014 Federal Tax Return claiming [own exemptions] exemption(s) and [own status] filing status, like you did. Fred also claimed the standard deduction, like you did.[10] However, Fred's tax computation is particularly simple, since all of his taxable income comes from his annual salary. He has no other sources of taxable income, and is not claiming additional credits or deductions.
>
> For the following questions, we will ask you to estimate the total federal income tax Fred would have to pay for different levels of total income. To help motivate careful thought about these questions, we are providing a monetary reward for correct answers. At the end of the survey, one of these questions will be chosen at random. If your answer to that question is within $100 of the correct answer, $1 will be added to your survey compensation.

Following this preamble, subjects made 16 forecasts of taxes due under different amounts of income, given the following prompt:

> If Fred's total income for the year were $[X], the total federal income tax that he has to pay would be:

The amounts of income substituted into the prompt above were drawn according to three sampling schemes. Ten forecasts were drawn from what we refer to as the "primary sampling distribution." This is a range of income values spanning all but the top of the national income distribution, sampling uniformly from $0 up to a point partially through the fourth tax bracket. This sampling pattern differs by filing status, leading us to present estimates separately by filing status in some of our analysis. Four forecasts were drawn uniformly from the "high-income sampling distribution," starting from the top of the primary income distribution and ranging to approximately $500,000. Finally, two draws were included that guarantee the presence of some forecasts "close" to the respondent's own income. One draw substituted the respondent's own reported income for X above, while the second substituted in that income plus a random perturbation taking a value between 0 and 1000.[11] When assessing respondents' knowledge of the tax schedule local to their own income, we will restrict data to the "local distribution" consisting of these two forecasts as well as any of the random forecasts that happen to fall in the respondent's own tax bracket. However, when we are not assessing questions about local tax perceptions, we exclude these two forecasts to preserve a random sampling structure.

---

appropriate.

[10] For filers not claiming the standard deduction, this sentence read: "Unlike you, Fred claimed the standard deduction."

[11] This value fell within the respondent's own tax bracket for all but 89 respondents.

Miscellaneous Questions: After the forecasting task, subjects faced a brief battery of miscellaneous questions. These included an elicitation of the salience of their income tax, assessments of their health and savings behaviors, an elicitation of their elasticity of charitable giving, the "big three" financial literacy questions of Lusardi and Mitchell (2014), an attention check, and a test of knowledge of their sales tax rate.

Incentives: On the final screen, one of the respondents' 16 tax forecasts was randomly selected for incentivization. They were told the correct answer, reminded of their own answer, and awarded the bonus payment if their response was within $100 of the truth.

1.2 Sample for Analysis and Dataset Preparation

Over the course of our sampling period, we collected 4,828 complete responses. We exclude responses according to several criteria. First, we exclude 5 responses with missing data on one or more of the tax forecasts. Second, we exclude 73 responses from individuals forecasting either 0 tax or 100% tax for all forecasts, as we believe this reporting pattern indicates either misunderstanding of the prompt or represents an attempt to quickly click through the survey without meaningfully responding to questions. Third, we restrict our sample to individuals reporting income ranging from zero to $250,000, excluding 117 respondents whose self-reported incomes are outside the typical range of the panel. Finally, we exclude 436 respondents who failed the attention check included in the miscellaneous questions module. To limit the influence of extreme tax forecasts, we conduct a rolling Winsorization of tax forecasts to the 1st and 99th percentile values in each $10,000 income bin.

This set of restrictions results in a final sample of 4,197 respondents, and a total of 58,758 forecasts of tax liability for randomly drawn incomes. In section 2.3, we analyze the robustness of our empirical results to these dataset construction decisions.

2 An Empirical Assessment of Tax Misperceptions

2.1 Reduced-Form Analysis of Aggregate Misperceptions

Graphical Summary of Perceived Schedules: To present an initial, non-parametric summary of income tax perceptions, figure 1 plots a kernel-smoothed estimate of individual tax forecasts from our two primary filing-status groups: single and married filing jointly. The top panels present estimates restricted to the data from the ten income draws of the primary sampling distribution. The lower plots extend the support to include the four income draws from the high-income sampling distribution.

The top two panels reveal several initial patterns. First, on average, the perceived tax schedule is qualitatively similar to the true tax schedule, though it displays some systematic error. Over the primary sampling range, respondents overestimate the tax burden by \$679 (clustered s.e.: \$185) on average, or 3.2 percentage points (clustered s.e.: 0.003pp) in effective tax rates.

Second, and perhaps more importantly, these plots also demonstrate that the sign of the average misperception depends on the amount of income that is being taxed. In both plots, the average perceived tax schedule appears more linear than the true schedule, with a tendency towards overestimation of the tax burden for low amounts of income and underestimation of the tax burden for high amounts of income. In the lower two plots of figure 1, which expand the income support to include high-income forecasts, this underestimation of taxes on high incomes becomes even more pronounced. This pattern indicates a general underappreciation of the degree of progressivity in the current U.S. tax code.

To explore the differences in perceived schedules as a function of respondents' own income, figure 2 summarizes the forecasting bias across the tax schedule by income quartile. Presented are the fitted values from an estimate of the regression model

$$(\tilde{T} - T)_{i,f} = \sum_{b,q}\sum \alpha_{b,q} * I(\text{income}_f \in \text{bin}_b) * I(\text{income}_i \in \text{quartile}_q) + \epsilon_{i,f}.$$

In this regression, we predict the difference between the perceived tax ($\tilde{T}$) and the true tax (T) for person i's assessment of Fred scenario f on an income-quartile (denoted q) specific mean forecast error for each \$5,000 bin (denoted b).[12] The primary pattern described above—overestimation of low tax burdens and underestimation of high tax burdens—persists across all four income groups. Despite this consistent pattern, tax perceptions are significantly different across income quartiles: Wald tests reject the joint equality of the four income-quartile-specific estimates of α for each income bin (all p-values <0.014).[13] A key pattern revealed in the plots is that the "crossing point" where overestimation turns to underestimation occurs at higher income values for higher income respondents.

Testing for "Flattening" of Tax Schedules: Both figures 1 and 2 provide visual demonstrations of the key systematic misperception that will drive our theoretical analysis: underestimation of the slope of the tax schedule. This is visually apparent in the "flattening" of the estimated schedules in figure 1, and in the negative slope of the bias functions in figure 2. We quantify these patterns in table 1.

To formally test for underestimation of the slope of the tax schedule, we estimate fixed-effect OLS regression models of the form $\tilde{T}_{i,f} = \beta T_{i,f} + \nu_i + \epsilon_{i,f}$. The object of interest in this analysis is β, which measures the scaling of the tax schedule implicit in the subjects' responses. By including respondent-specific fixed effects (ν_i) we identify β from the effective slope of the tax schedule reported within-subject. We test the null hypothesis of $\beta = 1$, the value that would be estimated if

[12]Note that the estimation sample is restricted to cases with tax burdens in the range [0,55000).

[13]Standard errors in the regression forming these estimates are clustered at the respondent level.

respondents indicated a rate-of-increase of taxes consistent with the true tax schedule. An estimated value over 1 would indicate an implicit steepening of the schedule, and an estimated value under 1 would indicate an implicit flattening.

The first panel of table 1 presents estimates of β derived from the 14 random draws of the primary and high-income sampling distributions. The parameter estimate of 0.62 (clustered s.e.=0.010) indicates substantial and statistically significant underestimation of the steepness of the income tax. The 2nd-5th columns of the table provide estimates of this same parameter when the sample is restricted to respondents in each of the four income quartiles. All four groups significantly underestimate the steepness of the tax schedule. The degree of underestimation is most severe among the lowest-income respondents: estimates range from 0.53 (clustered s.e.=0.020) for the lowest-income respondents to 0.76 (clustered s.e.=0.017) for the highest-income respondents. The second panel identifies β only off of the primary income draws, and again demonstrates that $\beta < 1$, with the difference increasing in the respondent's own income.

In the third panel we restrict the estimation sample to the "local distribution," consisting only of the two locally sampled income draws and any of the randomly sampled income values that happen to fall in the respondent's own tax bracket. Under this restriction, the reduced-form analysis directly tests if people correctly perceive the marginal tax rates that they face. We find that people underestimate the marginal tax rates in their own tax-bracket ($\beta = 0.81$, clustered s.e.=0.043, p<0.001). When examining these estimates by income quartile, we find that the effect remains statistically detectable for respondents in the top two income quartiles. For respondents in the bottom two income quartiles, we cannot reject correct perception of the local slope of the tax schedule. However, the standard errors of these estimates are sufficiently large that we can not reject that these respondents underestimate (or overestimate) their MTRs by a meaningful degree.

An interesting feature of the reduced form results is that while β is increasing in income in the top two panels, it appears to be decreasing in income when using only local draws. We will show below that these patterns are consistent with the models of heuristics that we estimate.

2.2 Disentangling Heuristic Use

What generates the misperceptions observed in the previous section? And more generally, how do taxpayers form their forecasts of the tax consequences of different actions? In this section, we explore the possibility that taxpayers apply simplifying heuristics to to react to a complex, nonlinear schedule. We focus our empirical analysis on the ironing and spotlighting heuristics of Liebman and Zeckhauser (2004), and additionally quantify the features of remaining misperceptions not captured by these existing models.

Defining Candidate Heuristics: We begin by formally defining our models of the heuristics described in Liebman and Zeckhauser (2004), and present a simple illustration of these heuristics in figure 3.

The first heuristic, *ironing*, is applied by individuals who know the average tax rate they face, and forecast tax liability by applying their average tax rate to all incomes. Using the ironing heuristic, the forecasted tax at income z is given by $\tilde{T}_I(z|z^*, \theta) = A(z^*|\theta) * z$, where z^* denotes the individual's own income, θ denotes all individual-specific characteristics that determine the applicable tax schedule, and $A(z^*|\theta)$ denotes the individual's average tax rate. This heuristic has the practical benefit that it leads to reasonably accurate beliefs about the *levels* of taxes when considering small deviations from one's current income. Thus for decisions about how to budget one's annual income, this heuristic leads to minimal errors.

However, for a variety of other decisions, the application of this heuristic meaningfully deviates from correct forecasting. When applied to forecasts of non-local income amounts, this heuristic leads to overestimation of the tax burden for comparatively low incomes and underestimation of the tax burden for comparatively high incomes, consistent with the qualitative patterns seen in figure 1. Most importantly, this heuristic leads to inaccurate beliefs about marginal tax rates: because the tax schedule is convex, average tax rates are systematically smaller than marginal tax rates, and thus the application of this heuristic generates a "flattening" of perceived schedules, consistent with the patterns observed in the reduced form results. Moreover, because the difference between marginal and average tax rates is largest for the top income quartiles, this heuristic is consistent with the qualitative results in panel 3 of table 1.[14] At the same time, because the ATR is higher for higher-income individuals, the perception of marginal tax rates in tax brackets other than one's own will be increasing in an individual's income, producing patterns consistent with panels 1 and 2 of table 1. We will demonstrate in section 3 that ironing leads to suboptimal labor supply decisions.

The second heuristic, *spotlighting*, is applied by individuals who know their own tax and own *marginal* tax rate, and forecast tax liability by applying their marginal rate to the difference between their own income and the income amount under consideration. Using the spotlighting heuristic, the forecasted tax at income z is given by $\tilde{T}_S(z|z^*, \theta) = T(z^*|\theta) + MTR(z^*|\theta) * (z - z^*)$, where z^* again denotes the individual's own income, $MTR(z^*|\theta)$ denotes the marginal tax rate at that income, and $T(z^*|\theta)$ denotes the true tax due at that income. Within one's own tax bracket, this heuristic leads to correct beliefs about the level and slope of the tax schedule; as a result, this heuristic is a good short-cut to determining optimal labor supply decisions in the short-run. Formal study of the use of this heuristic in tax settings has been limited, since these properties imply that the use of this heuristic is hard to identify from local tax perceptions or from labor supply decisions. The forecasts of this heuristic deviate from accurate tax forecasting only when considering income amounts outside of the taxpayer's own tax bracket, at which point this heuristic leads to underestimation of tax rates both for the comparatively rich and the comparatively poor. The aggregate patterns in the data seem inconsistent with this heuristic. However, the reduced-form

[14] Average difference between individual MTR and ATR by income quartile: quartile 1) 3.1pp, quartile 2) 6.8pp, quartile 3) 8.0pp, quartile 4) 8.9pp.

results cannot rule out that at least some people may be relying on this heuristic, a quantitative question we turn to next.

Estimating Heuristic Propensity: To provide quantitative estimates of the propensity of heuristic use, we estimate a structural model of misperceptions that incorporates these two heuristics, as well as any residual misperceptions that cannot be accounted for by these heuristics. We present results generated from two estimating equations:

$$\tilde{T}_{f,i} = (1 - \gamma_I - \gamma_S)T(z_{f,i}|\theta_i) + \gamma_I \tilde{T}_I(z_{f,i}|z_i^*, \theta_i) + \gamma_S \tilde{T}_S(z_{f,i}|z_i^*, \theta_i) + \epsilon_{f,i} \quad (1)$$

$$\tilde{T}_{f,i} = (1 - \gamma_I - \gamma_S)(T(z_{f,i}|\theta_i) + r(T(z_{f,i}|\theta_i))) + \gamma_I \tilde{T}_I(z_{f,i}|z_i^*, \theta_i) + \gamma_S \tilde{T}_S(z_{f,i}|z_i^*, \theta_i) + \epsilon_{f,i} \quad (2)$$

In these equations, $\tilde{T}_{f,i}$ denotes the forecasts of the taxes due by the hypothetical taxpayer. Individual respondents are indexed by i, and iterations of the hypothetical taxpayer question are indexed by f. We model tax forecasts as a convex combination of three possible models of tax perceptions. We include the ironing and spotlighting forecasts as defined above, each evaluated at the hypothetical income assigned to Fred ($z_{f,i}$), but using the average tax rate or marginal tax rate determined by the respondents' own income (z_i^*). In equation (1), we estimate a model in which aggregate tax forecasts are formed by a mixture of these two heuristics and the true tax liability ($T(z_{f,i}|\theta_i)$).[15] In equation (2), this latter term is augmented to $T(z_{f,i}|\theta_i) + r(T(z_{f,i}|\theta_i))$, denoting the true tax due plus a *residual misperception function.* By including this term and estimating it with a flexible functional form, we can separately identify our candidate heuristics from general misperceptions of the tax schedule not attributed to the models defined above. In the estimates we present below, we model the residual misperception function as a fifth order polynomial. Similar results are obtained with any polynomial of orders 1 through 10 (see appendix table A2).

The dataset obtained from our experiment is unique in allowing us to estimate equations (1) and (2). To see why, imagine that we had data on individual's beliefs about only their own MTRs or ATRs—the perceptions typically identified from observable marginal decisions. In such a setting, the forecasts of the spotlighting model do not differ from the true tax schedule, making it

[15] Our estimates of the true tax owed by a respondent contain some measurement error if the respondent's tax returns are more complex than Fred's (e.g, due to additional tax credits claimed by the respondent or due to an itemized instead of a standard deduction). However, our design ensures that we can compute the tax owed by Fred without any measurement error, and thus if respondents were answering all questions correctly, we would correctly compute that $\gamma_I = \gamma_S = r(T) = 0$. That is, our design ensures that there is no measurement error in our classification of unbiased responses.

However, if the measurement error generates unaccounted variance between respondents and/or leads our estimates of average and marginal tax-rates to be higher than they are, then that would attenuate our estimates of our ironing and spotlighting propensity *if any exists.* This would render our estimates conservative. As we discuss in section 2.3 and elaborate in appendix C, all of our results continue to hold—and are slightly larger in magnitude for ironing–when restricting our sample to the 1357 respondents with simple tax returns for which we can compute the exact tax owed with any potential measurement error.

impossible to separately identify. Furthermore, the prediction of the ironing heuristic is a "flattening" of the local tax schedule, but this feature could alternatively be generated by, e.g., a simple misunderstanding of progressivity that is identical for both low-income and high-income consumers. A unique prediction of ironing that can be tested with our experiment, however, is that holding fixed Fred's income, a respondent's beliefs about the tax owed by Fred should be increasing in the survey respondent's income, to a degree quantitatively determined by the mapping of incomes to average tax rates. Spotlighting, in contrast, predicts that an individual's beliefs about Fred's tax liability will be increasing in the individual's income if the individual's income is lower than Fred's, but *decreasing* in the individual's income if the individual's income is higher than Fred's. These sharp and quantitative predictions of ironing and spotlighting allow us to identify the propensity of each heuristic, and to disentangle these heuristics from misperceptions of the tax schedule that are invariant to one's own income.

Table 2 presents non-linear least squares estimates of the models (1) and (2), with standard errors clustered at the respondent level. Columns 1 and 3 present estimates of model 1, whereas columns 2 and 4 present estimates of model 2. Focusing first on the first two columns, which restrict the estimation sample to only the 10 iterations of the primary sampling distribution, we see that substantial weight is placed on forecasts of the ironing heuristic. In column 1, the point estimate implies 21% weight on the ironing heuristic in the convex combination model. However, the point estimate on the spotlighting forecast is negative 9%—outside the range of valid probability values, and marginally significantly so. We view the estimation of invalid probabilities for heuristic propensity as evidence of model mispecification, and a demonstration of the difficulty of inference in this setting when non-income-dependent misperceptions are not accommodated.[16] Illustrating that point, when the residual misperception function is included in this estimation in column 2, this odd result disappears. Weight on the spotlighting heuristics drops to what we view as a precisely estimated zero, while weight on the ironing heuristic increases to 29%. The top panel of figure 4 plots the estimate of the residual misperception function generated in column 2, and demonstrates a systematic tendency to overestimate taxes across the entire primary sample distribution. The contrast of columns 1 and 2 demonstrates the importance of allowing for residual misperceptions when estimating the propensity of these heuristics: since these heuristics can change the level of aggregate tax forecasts, their identification can be confounded with level effects when residual misperception is not accommodated.

In columns 3 and 4, we repeat the estimation exercise of columns 1 and 2 while additionally including the high income sample of tax forecasts. In these specifications, we again find a high weight on the ironing forecast (47% and 43% across the two models) and effectively zero weight on the

[16]To rationalize this result, and the confound introduced by excluding controls for residual misperceptions, recall that we found systematic overestimation of the taxes due across the primary income distribution. This feature is not predicted by any of the forecasting rules included in model 1. Since the spotlighting heuristic generates underestimation of taxes due outside of one's own bracket, placing negative weight on this heuristic is a simple way for the model to approximate our finding of systematic overestimation of tax levels.

spotlighting forecast (-3% and -2%). Furthermore, the estimated residual misperception function again exhibits a systematic overestimation of the tax burden across the primary sample range of tax values. These residual misperceptions change to systematic underestimation of comparatively large tax burdens, although this latter range of the distribution is imprecisely estimated due to the sparse sampling pattern over high-income forecasts.

Individual-Level Estimates: The results of table 2 suggest that aggregate tax misperceptions can be rationalized by placing significant weight on the ironing forecast. This is perhaps most naturally interpreted through a heterogeneous model in which some individuals have accurate beliefs (or accurate beliefs up to the perturbation of the residual misperception function) and some individuals employ the ironing heuristic. In such a model, our estimated coefficients may be interpreted as the propensity of use for each of the candidate forecasting rules. However, in principle our results could alternatively be rationalized by a homogeneous population all following a decision rule that places some weight on the truth and some weight on the ironing heuristic. While we believe that such a model would be difficult to psychologically motivate, we present individual-level estimates of our regression model to help distinguish between these possibilities.

To begin, we estimate tax perception model 1[17] at the individual level for each of the 3552 respondents facing a non-zero tax rate.[18] Figure 5 plots a kernel-density estimate of the distribution of estimated individual classifications.[19] In general, this distribution is quite diffuse. This is to be expected, since estimating two parameters from only 14 tax forecasts would lead to a distribution of point-estimates that is a convolution of both the true parameters but also substantial measurement error. Notice, however, that the resulting estimates yield a sharply bimodal distribution, with peaks of mass centered on pure ironers and pure "true tax" forecasters. While this analysis cannot rule out the existence of some intermediate cases, this distribution is consistent with a substantial population of pure ironers, in the sense intended by Liebman and Zeckhauser's (2004) presentation of the heuristic.[20]

We can also perform the same estimation exercise while restricting individual-level parameter estimates to valid probabilities. To do so, we numerically optimize over a grid of probability values,

[17]Similar figures can be generated estimating equation 2, and by restricting the estimation sample to only the primary income distribution. For replications of figure 5 applying the restrictions for each of the columns in table 2, see appendix figure A3.

[18]Notice that for individuals facing zero tax, the ironing and spotlighting heuristics yield the same forecast, and are thus not separately identified. For all individuals, the lack of within-subject variation in the MTR and ATR eliminates the possibility of separately identifying our heuristics from residual misperceptions, although the estimated residual misperception functions of table 2 may be included as an analogous control.

[19]For plots of the marginal distributions of both the ironing and the spotlighting coefficients, see appendix tables A1 and A2.

[20]When incorporating the residual misperceptions appendix figure A3, we use the residual misperception function estimated in 2 and estimate, for each individual, the weights γ_I and γ_S that they place on the ironing and spotlighting forecasts. Even with a residual misperception term, we find a large mass at $\gamma_I = 1, \gamma_S = 0$, indicating that our inference about the existence of pure ironers is not confounded by the residual misperception function. However, the density around $\gamma_I = 0, \gamma_S = 0$ is reduced by half.This can be rationalized by interpreting the residual misperception function not as a literal forecasting rule by a pure "type," but as an approximation to heterogeneous remaining misperceptions beyond the heuristics we study.

and assign individuals to the grid point that minimizes the mean squared error of their forecasts. Table 3 presents the frequency of individual level estimates, and again we see strong support for the existence of pure types. In this classification approach, 33% of respondents are classified as "true tax" forecasters, 33% are classified as pure ironing forecasters, 2% are classified as pure spotlighters, and the remainder are estimated at intermediate values.[21]

Summary: Taken together, our data provide substantial evidence of widespread use of the ironing heuristic. Evidence of at least some improper reliance on average tax rates has been shown in laboratory studies (de Bartolome, 1995) and in empirical studies of individuals' responses to changes in tax credits (Feldman et al., 2016); furthermore, use of this heuristic has been shown to extend to consumer perceptions of non-linear energy-pricing schedules (Ito, 2014). Our elicitation of perceived tax schedules provides a uniquely direct test that facilitates precise estimates of the prevalence of this heuristic. At the same time, we are able to additionally and simultaneously test for the presence of spotlighting, as well as for residual mistakes not explained by these models. Our estimated model establishes that while spotlighting seems to be a negligible bias, a non-negligible portion of misperceptions arises from channels beyond ironing, and appears most consistent with a systematic under-appreciation of the progressivity of the tax code. In the remainder of this paper, after a discussion of robustness, we explore the theoretical and quantitative consequences of the misperceptions that we estimate.

2.3 Robustness

Persistence of Heuristic Use and Exposure to Tax Decisions: While tax knowledge is important to all people for, e.g., determining which policies and politicians to support or for budgeting spending, economic analyses often hinge on knowledge among specific groups: workers (in models of labor-supply) or individuals completing their own tax returns (in models of compliance). We continue to find significant prevalence of the ironing heuristic among both the employed and the unemployed (see appendix table A6), those who completed their own tax return and those who did not (A7), those who use tax preparation software and those who do not (A8), and those who completed the survey before or after tax day (A9). Furthermore, we find substantial prevalence of the ironing heuristic among both financially literate and financially illiterate tax filers, as classified by performance on the "Big Three" financial literacy measures (A10). Finally, we find that adoption of ironing persists among both above- and below-median age respondents (A11; median age: 51), suggesting that the misperceptions we document are not eliminated by the typical experience with tax payments incurred over a lifetime. Across all these groups considered, we additionally reproduce the reduced-form finding of underestimation of the slope of the tax schedule.[22] In short,

[21] Similar results are obtained by repeating analogs of this exercise for the sets of model and data restrictions from each column of table 2. See appendix tables A3-A5.

[22] Across the 36 hypothesis tests conducted (sample splits on 6 variables, testing for flattening among 3 estimation samples), the sole exception to this finding is a failure to reject the null hypothesis of correct slope perception for

the incorrect perceptions that we estimate appear pervasive across groups, and persists even among the most economically interesting respondents in our sample.

Inclusion of Other Taxes: In practice, the federal income tax is not the only tax on income; for most respondents, state taxes and FICA taxes also apply. Our experimental exercise specifically asked respondents to make forecasts about their federal income tax; however, a confused respondent could make forecasts that incorporate additional tax components. Since the inclusion of these extra taxes increases both the aggregate MTR and ATR, the presence of confusion of this sort would render our estimates of the degree of underestimation of the steepness of the tax schedule conservative. Thus, this confusion cannot account for our central reduce-form results. Moreover, this confusion could not account for our evidence of ironing, since it would not explain why for a fixed Fred income, a respondent's estimate of Fred's tax liability is increasing in his own income.

However, such confusion could affect the actual point estimate of ironing, as well as our estimate of the residual misperception function. To examine the sensitivity of estimates to these concerns, we reestimate our primary heuristic model presented in table 2 under three alternative assumptions. In each iteration, we reestimate table 2 assuming that the true tax, ATR, and MTR are all based on an aggregate tax schedule that additionally includes FICA tax, state tax, or both.[23]

Results are presented in appendix tables B1-B3. In our preferred specifications (columns 2 and 4 of these tables), we find that our quantitative estimates of ironing and spotlighting propensities are only minorly affected by these potential sources of confusion; however, when the residual misperception function is not included, quantitative results are dramatically different, and generally implausible. This contrast demonstrates an advantage of our empirical approach. The apparent misperception of tax amounts that would result from the contraindicated inclusion of additional taxes takes a form that can be closely approximated by the residual misperception function ($r(t)$). Absent the presence of a residual misperception function, this type of confusion could be incorrectly attributed to heuristic forecasting; with a residual misperception function included, this class of forecasting errors is correctly classified as alternative phenomena.

Similarity of Actual and Hypothetical Tax Filers: Our experiment focused on a hypothetical taxpayer, constructed to approximate the respondent. While the hypothetical taxpayer had the same filing status and number of exemptions as the respondent, he was built with intentionally simple taxable behavior: only wage income, with no additional schedules, credits, or deductions. This design element resolves an important difficulty present in other surveys of tax knowledge: uncertainty about the complete details shaping tax liability. While this design does eliminate the measurement error inherent from that lack of knowledge, and thus allows us to incentivize the forecasts, it has one undesirable feature: respondents with filing behavior more complex than pure

local draws by the financially literate.

[23] We approximate state tax liability by applying the state's single or married-filing-jointly schedule to the federal adjusted gross income. Note that across states there are often small differences in the calculation of the tax base, which we necessarily abstract from due to data limitations. In analysis including state tax approximations, we exclude 34 respondents that we are unable to match to a state.

wage income are making forecasts regarding a tax schedule that imperfectly approximates their own. Our description of Fred precisely matches the returns submitted by 1357 (32%) of our respondents, and the remaining 2840 respondents have some element of their tax return—such as schedule B-F, an itemized deduction, or a claim to the EITC—that renders the approximation imperfect. In appendix tables C1, C2, and C3, we conduct our main analyses restricted to both those respondents whose simple tax return behavior that perfectly matches our description of Fred in the forecasting exercise, as well as those with additional complexity. In both groups, we replicate the flattening of perceived tax schedules, the presence of ironing, the lack of spotlighting, and the general shape of the residual misperception function.

Importance of Data Restrictions: While most of our data restrictions described in section 1.2 are standard and affect few responses, two decisions might be viewed as non-standard and affect larger groups. First, note that we exclude 436 respondents (9% of our initial sample) who failed the attention check included in the miscellaneous questions module. In appendix tables $D1$ and $D2$, we show that our primary findings of flattening of tax schedules and the prevalence of ironing persist when these respondents are reincluded. While this exclusion has little effect on the final results, we implement it as a matter of principle. Prior to running analyses, we worried that forecasts of respondents that do not carefully read instructions would necessarily be imperfect, and that the imperfection resulting from their inattention would not generate an externally valid measurement of the misperceptions of interest.

Second, we employ a Winsorization strategy as a means of controlling extreme forecasts. When deploying a unconstrained-response survey to thousands of respondents, at least a small number of wildly unreasonable forecasts are to be expected. To present an illustrative example, one respondent indicated that the tax due for an income of $823 is $96,321, when in fact it is zero. Even if most respondents have reasonably accurate tax perceptions, a small number of such extreme forecasts can significantly impact both parameter estimates and power. Furthermore, we believe the extremity of such forecasts does not approximate any externally valid forecasting problem, but rather is an indication of unusual confusion or experimental noncompliance. This motivated our choice to Winsorize tax forecasts at the 1st and 99th percentile forecasts within each $10,000 bin. As we demonstrate in appendix tables $D3$-$D6$, alternative means of Winsorization have little impact on our quantitative estimates. Furthermore, our basic results persist even with the complete omission of outlier control, although estimates become notably less precise.

3 Implications for Social Welfare

In this section, we embed the misperceptions that we estimate in a standard model of labor supply with heterogeneous and hidden skills in the tradition of Mirrlees (1971). We study whether a social planner with redistributive motives would want to correct these misperceptions.

3.1 Model Set-Up and Definitions

Economic Setting: Individuals have a utility function $U(c, l) = G(c - \psi(l))$, where c is consumption, l is labor, and $\psi(l)$ is the disutility of labor. Each individual produces $z = wl$ units of income for every l units of labor, where the wage w is drawn from an atomless distribution F. We assume $\psi'(0) = 0$ and $lim_{z\to\infty}\psi'(z) = \infty$.

The government cannot observe earnings potential w, and is thus restricted to setting taxes $T(z)$ as a function of earnings z. For a given tax schedule T, we let $z^*(w)$ denote the earnings choice of a type w individual. The social welfare function is given by $\int U(z^*(w) - T(z^*(w)), z^*(w)/w)dF + \lambda \int T(z^*(w))dF$, subject to $\int T(z^*(w))dF \geq 0$, where λ is the marginal value of public funds. We let $g(z) = G'(z-T(z)-\psi(l))/\lambda$ denote the social marginal welfare weight at the current tax system.

Tax Perceptions and Labor Supply Choice*:* Let $\tilde{T}_w(z|z^*)$ be the perceived tax schedule of a type w individual earning z^*. In our model, tax perceptions influence welfare through their effect on labor-supply decisions. To close the model, we formalize the relationship between labor-supply behavior and misperceptions through a solution concept we refer to as *Misperception Equilibrium.*

Definition 1. Choice $z^*(w)$ is a *Misperception Equilibrium (ME)*, if $z^*(w) \in \text{argmax}\{U(z - \tilde{T}_w(z|z^*(w)), z/w\}$.

The ME concept requires that z^* is a fixed point of a decision process in which misperceptions are possibly shaped by z^*, while at the same time z^* is a perceived optimum given those misperceptions.[24] The ME concept may be reformulated as a special-case of Berk-Nash equilibrium (Esponda and Pouzo, 2016) and, as such, can be microfounded as a steady state of a dynamic process in which individuals follow a myopic best-response strategy while learning through a misspecified model.[25] However, a fully dynamic model would be impractically unwieldy in many applications, and inappropriate for the kinds of canonical static labor supply models that we consider. We view our static solution concept as a tractable and useful approximation to a more nuanced and dynamic psychological process.

To study the implications of our empirical findings using the ME framework, we put the following structure on misperceptions $\tilde{T}$:

Assumption A: $\tilde{T}(z'|z) = (1 - \gamma_I - \gamma_r)T(z') + \gamma_I A(z)z' + \gamma_r \tilde{\tilde{T}}(z')$, where $\tilde{\tilde{T}}$ is everywhere flatter than T.

Note that the functional form imposed by Assumption A corresponds to the functional form in the

[24]Farhi and Gabaix (2015) implicitly use this solution concept in their study of optimal income taxation. We formalize the solution concept and characterize existence or uniqueness for the types of misperceptions that we estimate.

[25]To formally embed our model in the Berk-Nash framework, we must re-interpret $\tilde{T}(z|z^*)$ as the mean of the individual's belief, while allowing the individual to have a sufficiently diffuse prior so that no outcomes are "surprises."

empirical model:

$$\begin{aligned}\tilde{T}(z'|z) &= (1-\gamma_I)T(z') + \gamma_r(\tilde{\tilde{T}}(z') - T(z')) + \gamma_I A(z)z' \\ &= (1-\gamma_I)\left(T(z') + r(T(z')\right) + \gamma_I A(z)z',\end{aligned}$$

where $r(T(z')) = \frac{\gamma_r(\tilde{\tilde{T}}(z') - T(z'))}{1-\gamma_I}$. For simplicity, we assume that γ_I and γ_r do not vary by earnings ability w, an assumption that is consistent with our empirical results.

3.2 Implications for Behavior

Before moving on to studying implications for policy, we first show that ME labor supply is well-behaved under assumption A, and can be tractably studied in optimal tax settings. Proofs of all propositions are available in appendix E.

Proposition 1. *Suppose that $T(z)$ is continuous and convex. Then*

1. *There exists a unique ME z^*.*

2. *z^* is continuous and increasing in w.*

3. *z^* is continuous and increasing in γ_I and γ_r.*

Part 1 of the proposition states that there exists a unique ME as long as individuals believe that the tax system is progressive. Part 2 of the proposition shows that the biases we study preserve the monotonicity property of the standard model that higher wage workers choose to earn more.

Part 3 presents a key result, characterizing how these biases affect behavior. In the standard model—assuming perfect tax perception—a lower marginal tax rate increases the return to labor, and thus leads the taxpayer to work and earn more. The two forms of misperception captured under Assumption A (measured by γ_I and γ_r) both create the misperception that marginal tax rates are lower than they actually are, and thus induce greater earnings. This result—that the misperceptions we estimate induce higher earnings, and thus higher tax revenue—will be the key channel through which our estimated misperceptions influence welfare calculations in the results that follow.

3.3 Implications for Social Welfare

In this section we present our two key results summarizing the implications of debiasing for social welfare, and then discuss the general intuition driving these results. For the sake of conciseness, we state the results here for everywhere-differentiable tax functions T, which lead to earnings choices z^* that are everywhere differentiable in the model parameters. However, identical results can be generated for misperceptions with a finite number of points of non-differentiability (with derivatives replaced by left and right derivatives where appropriate), and we provide these results in the proofs in the appendix.

We begin by establishing a qualitative result about the optimal level of bias in the population, showing that at least some bias is beneficial.

Proposition 2. *Let* (γ_I^*, γ_r^*) *be the values of* (γ_I, γ_r) *that maximize social welfare. Then* $\max(\gamma_I^*, \gamma_r^*) > 0$.

Proposition 2 illustrates a provocative tension faced by the social planner. If the social planner were able to choose the parameters governing misperceptions, the planner would not chose to set them to zero, despite the fact that the bias leads individuals to make personally suboptimal decisions.[26]

To help develop intuitions for this result, and to extend these findings to situations where a baseline level of misperceptions exists, we now characterize the social welfare consequences of "nudges" applied to existing misperception levels. We introduce nudges to our model as an object, denoted n, that changes individuals' perceptions (for example, by decreasing γ_r or γ_I parameters) and thus, by Proposition 1, influences earnings. For the general results here, we make no assumptions about which individuals are affected more or less by the nudge. This formulation of a nudge can alternatively be used to characterize the effects of *increasing* taxpayer bias; for example, by increasing γ_r or γ_I parameters. We consider both of these cases in the results below.

Proposition 3. *1. If* $\frac{dz^*(w)}{dn} < 0$ *for all* w *then*

$$\frac{dW}{dn}/\lambda \leq \int \frac{dz^*(w)}{dn} \left(T'(z^*(w))(1 - g(z^*(w))\right) dF(w) \tag{3}$$

Moreover, if $\left|\frac{dz^*(w)}{dn}\right|$ *is nondecreasing in* w *and* $E[g(z)] \leq 1$*, then*

$$\frac{dW}{dn}/\lambda \leq 0$$

2. If $\frac{dz^*(w)}{dn} > 0$ *for all* w *then*

$$\frac{dW}{dn}/\lambda \geq \int \frac{dz^*(w)}{dn} \left(T'(z^*(w))(1 - g(z^*(w))\right) dF(w) \tag{4}$$

Moreover, if $\frac{dz^*(w)}{dn}$ *is nondecreasing in* w *and* $E[g(z)] \leq 1$*, then*

$$\frac{dW}{dn}/\lambda \geq 0$$

[26] See also Goldin (2015) for an analysis in this same spirit, illustrating that a policymaker's optimal choice of sales tax salience would make consumers at least partially inattentive to the tax.

This proposition shows that for the kinds of biases that we study, it is possible to generate appropriate bounds that are expressed solely in terms of how behavior responds to the nudge in question and the policy-maker's social preferences (given by the social marginal welfare weights). Assuming that behavior is biased by the misperceptions we have documented, the conditions of part 1 of the proposition ($\frac{dz^*(w)}{dn} < 0$) represent a debiasing nudge, and equation 3 presents an upper bound for the benefit of this nudge. Analogously, the conditions of part 2 of the proposition represent a nudge that *exacerbates* the mistake caused by our estimated misperceptions, and equation 4 provides a lower bound on the welfare benefit of this nudge. These formulations demonstrate that, while exact quantitative estimates of γ_r and γ_I are needed to point-identify the welfare effects, informative bounds can be constructed based only on the qualitative features formalized in Assumption A.

Using these bounds, we can sign the effect of the nudge under two additional assumptions: that the nudge has a larger absolute effect on higher ability (and thus higher-income) consumers ($\left|\frac{dz^*(w)}{dn}\right|$ is nondecreasing in w) and that the marginal value of public funds is not too small ($E[g(z)] \leq 1$). Under these assumptions, an analog of proposition 2 is obtained when considering deviations from preexisting misperception levels: nudges that reduce bias harm social welfare on the margin, and nudges that exacerbate bias help social welfare on the margin.

Propositions 2 and 3 are driven by a similar intuition. The social welfare effect of nudging a particular person earning z^* at wage w is given by

$$g(z^*) \cdot (\underbrace{\tilde{\tau}_w - T'}_{\text{Misperceived MTR}})\frac{dz^*}{dn} + \underbrace{\frac{dz^*}{dn}T'}_{\text{Fiscal externality}} \tag{5}$$

where $\tilde{\tau}_w$ is the perceived marginal tax rate for the biased person. There is a simple "price metric" interpretation for $\tilde{\tau}_w - T'$: if this person were rational, $\tilde{\tau}_w - T'$ is the amount by which his marginal tax rate would have to decrease for him to choose the same labor that he chooses when biased.[27] This term, therefore, is a money-metric measure of the cost of individual misoptimization, which the social planner weights according to the individual's social marginal welfare weight $g(z)$. But balanced against this individual welfare cost is a fiscal externality. As established in proposition 1, the misperceptions we estimate would induce the taxpayer to work too much, and thus pay more in taxes. When the social planner "nudges away" these misperceptions, tax revenue—and thus the funding for public goods—falls. Note, however, that nudging an individual for whom $g(z^*)(\tilde{\tau}_w - T') > T'$ actually increases social welfare, and thus it is not generally true that all possible nudges lower social welfare.

The assumptions imposed to derive unambiguous signs on nudge effects are both intuitive and relatively weak. The first assumption—that $\left|\frac{dz^*(w)}{dn}\right|$ is increasing in w—will typically hold in

[27] See Lockwood and Taubinsky (2015) for analogous results—making use of the price-metric approach—about commodity taxation in the presence of redistributive motives and nonlinear income taxes.

environments with constant elasticities and with with smooth and convex tax schedules. Use of the ironing heuristic generally leads higher-income consumers to have larger misperceptions of marginal tax rates than lower income consumers; and empirically, the third panel of table 1 confirms that higher-income consumers underestimate their marginal tax rates more in a proportional sense, and thus in an absolute sense. Consequently, a nudge that reduces taxpayers' biases will have a larger absolute impact on the marginal tax rate misperceptions of the comparatively rich. If the structural elasticity of labor supply with respect to the marginal tax rate is constant (or increasing) across the income distribution—as is commonly assumed—then this implies that higher-income taxpayers are more responsive to a marginal change in their perceived tax rates *in absolute terms*. Thus, under the assumption of constant elasticities and constant proportional effects of the nudge, $\left|\frac{dz^*}{dn}\right|$ would increase very steeply in w and satisfy the restrictions of the proposition. Applying this line of logic, the necessary inequality may be expected to hold so long as the poor are not significantly more elastic than the rich, and so long as the effects of nudges aren't extremely well-targeted to the poor.[28]

The second assumption—that $E[g(z)] \leq 1$—is effectively a requirement that the marginal value of public funds (λ) is sufficiently large. In our model, misperceptions generate extra tax revenue, and that tax revenue is applied to public funds. If public funds are not valued, this extra tax revenue is not valued, and thus the preservation of misperceptions is not valued. While this does suggest that bias-reducing nudges may be beneficial in situations where government spending is extremely ineffective, we note that the condition $E[g(z)] > 1$ necessarily implies that the social planner cannot view the tax schedule as optimal.

3.4 Simulating the Effects of Debiasing Nudges

To move beyond qualitative results and to gain a better understanding of magnitudes, we now study the quantitative impact of a nudge that completely corrects the misperceptions that we have measured. While we do not believe such an effective nudge literally exists, this thought experiment serves as a means of quantifying the aggregate welfare costs of the misperceptions we have observed.

In the formal implementation of this thought experiment, we parameterize individual utility according to the functional form $U(z) = log(z - T(z) - \frac{(z/w)^{1+k}}{1+k})$, a commonly used specification in optimal tax studies (e.g., Atkinson 1990; Diamond 1998; Saez 2001).[29] In this model, assuming correct tax perceptions, the labor supply elasticity is determined by $\frac{1}{k}$. When tax rates are misperceived, the elasticity with respect to wages must be augmented by the term $\frac{1-\tilde{T}'}{1-T'}$, which takes an average value of 1.01 in our data—thus, while the formal calculation of elasticity is not identical,

[28] An important caveat to this logic is that the monotonicity assumption on $\left|\frac{dz^*}{dn}\right|$ is necessarily violated near kinks in a tax schedule. However, these comparatively small local violations of monotonicity do not change the key comparisons of differences between low and high income taxfilers that we discuss above.

[29] Our quantitative results are robust to alternative assumptions on the strength of redistributive preferences. See appendix table A12 for a reproduction of table 4 conducted by replacing the log utility with alternative CRRA parameters.

quantitatively the difference is negligible. In simulations, we will vary the parameter k across values from 1 to 5, capturing elasticities ranging from approximately 0.2 to 1. In a recent meta-analysis of labor-supply elasticity estimates, Chetty et al. (2011) report microeconomic estimates ranging from 0.26 to 0.82 (across different labor supply elasticity concepts and methods of inference), with a preferred estimated of the intensive-margin Hicksian elasticity of 0.33. Our range of k values spans this range, and our midpoint of $k = 3$ approximately corresponds to Chetty et al.'s preferred estimate.

Contingent on these parameters, we calculate the implied wage parameter for each experimental subject. Under the assumption that observed behavior is a ME, an individual's wage parameter can be calculated according to the equation $w = (\frac{z^k}{1-\tilde{T}'})^{\frac{1}{1+k}}$. Since this calculation depends on the individual's perceived marginal tax rate $\tilde{T}$, we forecast this value at the individual level based on the model estimated in column 4 of table 2. With these estimated values, we then re-solve the individual's utility maximization problem contingent on the true tax schedule and their estimated wage parameter, calculate total government revenue under both the biased and debiased regimes, and calculate social welfare under the assumption that all government revenue funds a public good with marginal value λ. We conduct simulations across three alternative values of λ, benchmarking this parameter against the marginal utility of a dollar for different members of our experimental population. In the "low λ" simulations, we set λ to be equal to the 10th percentile of marginal utilities in our population. In the "high λ" simulations, we set λ to be equal to the 90th percentile of marginal utilities in our population. As an intermediate case, we set λ to be equal to the 50th percentile of marginal utilities in our population.

Table 4 presents the results of these simulations. The first column shows that across the range of considered elasticities, the impact of debiasing on total government revenue is substantial. Correcting the misperceptions we estimate, and thereby decreasing earnings by increasing the perceived marginal tax rate, is forecasted to reduce total revenue by 1.0-4.4%. In the last 3 columns of the table, we quantify the welfare effect of a full debiasing nudge—balancing the lost tax revenue against the elimination of individual optimization errors—by calculating the amount of government revenue the social planner would pay to avoid it. This amount ranges from a minimum of 0.9% in the most inelastic specifications to a maximum of 4.4% in the most elastic specifications. In short, accounting for the welfare impact of the correction of individual mistakes only minimally offsets the substantial welfare loss arising from the reduction of public funds.

3.4.1 Robustness of Welfare Estimates

Integrating State Taxes and FICA: A weakness of the analysis presented in Table 4 is that it abstracts from the reality that optimal labor supply decisions should additionally depend on taxes paid through other channels. For completeness, we thus perform the same exercise while additionally accounting for taxes collected through state income taxes and FICA.[30] Including these features in the simulation requires assumptions regarding the baseline level of misperception of these taxes, and the effect of the nudge of their perception. We consider two boundary cases. At one extreme, we consider a situation in which these components are already correctly understood and thus a nudge has no influence on their perception. Under this assumption, we find that the inclusion of additional taxes has minimal effect on our quantitative estimates. Complete debiasing is estimated to reduce federal tax revenue by 1.1-4.9%, and the social planner would pay 1.0-4.9% of current government revenue to avoid implementing the nudge. At the other extreme, we consider a situation where State and FICA taxes are completely ignored prior to the nudge, but the nudge completely corrects their misperception. Under this assumption, the nudge eliminates a substantially greater misunderstanding of the tax code than previously considered, and thus welfare effects are significantly inflated. Complete debiasing is estimated to reduce federal tax revenue by 5.5-22.8%, and the social planner would pay 2.6-21.5% of current government revenue to avoid implementing this nudge. We believe these dramatic social welfare losses overstate the true effect—both because we do not believe State Income Tax and FICA are fully ignored, and we do not insist that debiasing nudges for the federal income tax must necessarily completely debias misperceptions of other taxes. However, this contrast of assumptions provide a reasonable approximation of the best-case and worst-case debiasing scenarios when additional taxes are included, and demonstrate that the estimates in table 4 are conservative. For full results, see appendix table B4.

Locally Estimated Tax Perceptions: In our simulation exercise we assign perceptions of marginal tax rates according the the estimated model in column 4 of table 2. This model is estimated based on subjects' global knowledge of the tax schedule. As an alternative specification, we may substitute the estimates of marginal tax rates based on the locally estimated parameters presented in the bottom panel of table 1. Results of this exercise are presented in appendix table A13. Across these two approaches, estimated welfare losses from complete debiasing range from to 1.6 to 8.0% of total revenue, and are systematically higher than the estimates derived from our heuristic misperception model. In summary, our finding of substantial welfare losses of debiasing does not hinge on our use of non-local forecasts for heuristic identification, and can be replicated even with the most local and individually relevant forecasts included in our exercise.

Omission of Very-High-Income Filers: Due to our sampling structure, our within-sample income distribution closely approximates U.S. income distribution, with the caveat of being truncated at $250,000. This truncation is not innocuous for welfare estimates. While filers above this income threshold account for only 2% of tax returns, they pay 46% of all federal income tax revenue.[31] Their exclusion influences our estimates in two important ways.

[30] As previously noted in footnote 23, we approximate state tax liability by applying the state's tax schedule to the federal adjusted gross income. In analysis including state tax approximations, we exclude 34 responses that we are unable to match to a state.

[31] See https://www.irs.gov/uac/soi-tax-stats-individual-income-tax-returns#prelim.

First, notice that if top tax filers exhibit the underestimation of marginal tax rates documented in this paper, the welfare losses associated with debiasing nudges become more dramatic. As illustrated in equation 5, the social planner down-weights individual taxpayers' misoptimization costs by their social marginal welfare weights, which are typically assumed to tend to zero for sufficiently rich filers. The welfare-relevant consequence of debiasing a top-2-percent filer would therefore be nearly entirely driven by the fiscal externality component of the equation, guaranteeing that this taxpayers' individual contribution to the welfare effect of the nudge would be negative (this point is further explored in the following two sections). We believe that our focus on within-sample analysis provides the most principled and conservative approach to approximating welfare costs, as it does not rely on untested assumptions that the absolute richest filers exhibit the same misperceptions measured in our population. However, if they do, their effect would serve to accentuate the welfare costs.

Second, however, notice that in several of our calculations in table 4, we benchmark revenue losses or welfare effects against total government revenue. The lack of top-2-percent tax filers in our sample would naturally lead our within-sample revenue forecasts to underestimate true total revenue. Since the omitted range of returns pays 46% of total taxes, rescaling columns 2-4 of table 4 by 0.54 corrects for their omitted revenue. This correction rescales the welfare costs to range from 0.5-2.5% of total revenue, still representing a large welfare loss compared to common interventions.

3.4.2 Who Do Nudges Really Help?

Having established that the social welfare losses to universal debiasing would be large, we now show that a significant reason for this is that the nudge is highly regressive. We do this by calculating, for each person, the equivalent variation in income that would induce the same change in utility as the debiasing nudge does, and we normalize it by the per capita loss of tax revenue due to the nudge. Importantly, such money metric approaches enable transparent comparisons between the effects of different types of policy instruments—both in terms of their progressivity and in terms of how efficiently they raise public funds. In a standard model with optimizing agents, the effects of any local tax reform can be decomposed into two effects: the impact on public funds and the mechanical effect on each taxpayers' total wealth.[32] The progressivity of the tax reform can then be judged by the wealth effects on the rich vs. the poor, per unit change in the public funds. Our equivalent variation metric provides us with an analogous decomposition that can be compared to the effects of any tax reform, in a way that enables a transparent judgment about which is more progressive.

In figure 6 we present a local-polynomial estimate of the ratio of the EV of bias reduction to per-capita revenue loss.[33] Focusing first on this value for the lowest income filers, we see that this

[32]The envelope theorem ensures that it is not necessary to track how changes in taxpayers' behavior affect their utility.

[33]Note that for a fixed tax schedule, the value of of this ratio takes a single value for each point on the income

ratio takes values of essentially zero; for the poor, the individual benefit from debiasing is negligible in comparison to the per-capita loss of public funds. In contrast, for individuals exceeding income of approximately $175,000, the ratio takes values exceeding 1; for the rich, the individual benefit from debiasing is greater than the per-capita loss of public funds. In this sense, the nudge policy is clearly regressive: it uses costly public funds to effectively make the rich richer, while generating virtually no benefit to the poor. A converse of this statement is that the biases that we estimate help make the tax burden more progressive: they help the government collect tax revenue in such a way that the incidence falls more heavily on the rich than on the poor.

The intuition for this result derives from equation 5 above, which decomposed the individual social welfare effect into the misoptimization and fiscal externality components. The degree of misoptimization is determined by the difference between perceived and actual marginal tax rates. For individuals adopting the ironing heuristic, these deviations are comparatively small when earned income is low. However, once the amount of income crosses into higher tax brackets, ironing forecasts and true tax burdens can differ substantially. In short, tax misperceptions have little effect on those with low tax burdens; however, for those with large tax burdens, tax misperceptions generate significant individual misoptimization cost and significant fiscal externalities.

3.4.3 Implications for Targeted Debiasing

Given that an untargeted nudge is extraordinarily regressive, we now examine whether a more socially desirable policy could be constructed by applying the nudge in a targeted manner. In figure 7 we plot the individual contribution to total social welfare loss from debiasing taxpayers of different incomes, according to the assumptions about the social planner's preferences imposed by our model.

The figure shows that when income is comparatively low, both mistakes and fiscal externalities are small in magnitude. Consequently, the marginal impact of debiasing a low-income individual has a near-negligible contribution to social welfare. As progressively higher-income filers are considered, the welfare loss of debiasing becomes more significant in magnitude. As discussed in the prior section, these high-income individuals have significantly biased forecasts, which generates a comparatively large fiscal externality. However, when applying a social welfare function with preferences for redistribution, the individual welfare loss shouldered by the rich is heavily discounted by their low social marginal welfare weight. The social welfare impact of debiasing these individuals is therefore effectively entirely determined by the resulting fiscal externalities. Thus, debiasing each high-income filer generates social welfare losses exceeding $1,000.

distribution, and thus the effective averaging that occurs through a local polynomial estimate would be unnecessary. However, in the U.S. tax system, individuals at the same income level face different tax schedules due to differences in filing status, number of exemptions, and other relevant inputs. Our local polynomial results may be interpreted as assessing the relevant ratio averaged across the empirical joint distribution of income and schedule-relevant parameters.

While theoretical considerations point to the possibility of welfare gains from targeted debiasing of the poor, these analyses suggest that the welfare gains of such policies are extremely small in magnitude when compared to the welfare losses of untargeted policies.

4 Discussion

A rapidly growing literature has suggested that heuristics and biases can meaningfully influence the implementation of income taxes.[34] We contribute to this literature by combining a precise measurement of misperceptions of income tax schedules with a theoretical framework for evaluating these misperceptions' welfare consequences. We find that the welfare consequences of these misperceptions are dramatic: debiasing would lead to substantial reductions in both tax revenue and social welfare, and the incidence of these changes would be highly regressive. These considerations would lead a social planner to avoid corrective nudges.

While we have focused our theoretical analysis on policies directly designed to correct misperceptions, we believe that some of the most practical implications of these results arise from their interaction with policies that indirectly resolve confusion. Debates on large-scale tax policy changes often involve suggestions of substantial simplification of the bracket structure, at times taken to the extreme of suggesting a "flat tax." While such alternative policies have obvious direct effects on the progressivity of the tax system, our results imply that they would have important indirect effects through the policy change's interaction with misperceptions. If it were the case that these simplified structures naturally reduced or eliminated misperceptions of marginal rates, their implementation would inherently involve giving up the "free" redistribution currently occurring because of people's misunderstanding of the tax code. Our simulations suggest that this effect is meaningful in magnitude.

While we have focused on the implications of our empirical estimates for debiasing, our findings can also provide quantitative guidance to the recent work of Farhi and Gabaix (2015), who study the implications of misperceptions for optimal tax design.[35] We caution, however, that applying our empirical estimates to questions of optimal schedule design requires important, and potentially strong, assumptions. In order to solve for the optimal schedule, the analyst requires knowledge not only of the current structure of misperceptions, but also of how misperceptions would change if the tax schedule were perturbed. The models of ironing and spotlighting heuristics that we estimate are parsimonious theories of how individuals will perceive a tax schedule, and readily map new tax schedules to predicted patterns of misperception. In contrast, we have neither theoretical nor empirical guidance on how our estimated residual misperceptions would react to changes in the

[34] E.g., Spicer and Hero (1985); McCaffery and Baron (2006); Rees-Jones (2014); Engström et al. (2015); Lockwood and Taubinsky (2015); Lockwood (2015); Kessler and Norton (2016); Benzarti (2016).

[35] For related analysis, combined with empirical measurement of tax biases based on subjective well-being data, see Gerritsen (2016).

tax schedule. We leave the important challenge of theoretically grounding our estimated residual misperceptions for future research.

Moving beyond applications specific to the design of optimal tax policy, we highlight that our empirical estimates are broadly relevant to our understanding of tax incentives. When analyzing decisions to adopt tax-preferred behaviors, to invest in human capital in the hopes of raising future wages, or to make financial investments that will only accrue at a future date, our experimental design provides a unique view into the tax perceptions that should shape such decisions.

As a final note, we wish to caution readers against a particular interpretation of our primary theoretical results: that nudges should never be used to correct tax-related misperceptions. We emphasize that our evaluation of nudges is focused on correcting misperceptions of the tax rates of the federal income tax schedule for wage income. We have abstracted from misunderstanding regarding various other components of the schedule, such as the take up of credits and deductions. Particularly for the poor, failure to optimize along these dimensions is costly (Currie, 2006). For example, the Earned Income Tax Credit is the largest cash transfer program for low-income U.S. families. But despite it's importance as a redistributive instrument, substantial fractions of eligible claimants fail to request their credit, and those who are eligible often demonstrate confusion about the incentives the EITC induces (Liebman, 1998; Romich and Weisner, 2000). While naturally occurring variation in knowledge demonstrates some responsivity to the EITC's incentives (Chetty et al., 2013), interventions to nudge program participation often have limited effects (Chetty and Saez, 2013; Bhargava and Manoli, 2015). We believe that continued efforts to intervene in these environments are worthwhile, as they provide means for implementing tax-nudge policies that avoid the undersirable redistributive consequences that we have considered in this paper.

References

Abeler, Johannes and Simon Jäger, "Complex Tax Incentives," *American Economic Journal: Economic Policy*, 2015, *7* (3), 1–28.

Allcott, Hunt and Judd Kessler, "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons," *Working Paper*, 2015.

Atkinson, Anthony Barnes, "Public Economics and the Economic Public," *European Economic Review*, 1990, *34* (2-3), 225–248.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein, "Behavioral Hazard in Health Insurance," *The Quarterly Journal of Economics*, 2015, *130* (4), 1623–1667.

Benjamin, Daniel J., Ori Heffetz, Miles S. Kimball, and Nichole Szembrot, "Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference," *American Economic Review*, September 2014, *104* (9), 2698–2735.

Benzarti, Youssef, "How Taxing Is Tax Filing? Leaving Money on the Table Because of Hassle Costs," *Working Paper*, 2016.

Bhargava, Saurabh and Day Manoli, "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment," *American Economic Review*, 2015, *105* (11), 3489–3529.

Blaufus, Kay, Jonathan Bob, Jochen Hundsdoerfer, Christian Sielaff, Dirk Kiesewetter, and Joachim Weimann, "Perception of Income Tax Rates: Evidence from Germany," *European Journal of Law and Economics*, 2013, *40* (3), 457–478.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Memory, Attention, and Choice," *Working Paper*, 2015.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber, "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, May 2011, *101* (3), 471–475.

_ , Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

_ and Emmanuel Saez, "Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients," *American Economic Journal: Applied Economics*, 2013, *5* (1), 1–31.

_ , John N Friedman, and Emmanuel Saez, "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings," *American Economic Review*, 2013, *103* (7), 2683–2721.

Currie, Janet, "The Take-Up of Social Benefits," in Alan J. Auerbach, David Card, and John M. Quigley, eds., *Public Policy and the Income Distribution*, Russell Sage Foundation, 2006, pp. 80–148.

de Bartolome, Charles A. M., "Which Tax Rate do People Use: Average or Marginal?," *Journal of Public Economics*, 1995, *56* (1), 79–96.

Diamond, Peter A, "Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Marginal Tax rates," *American Economic Review*, 1998, *88* (1), 83–95.

Engström, Per, Katarina Nordblom, Henry Ohlsson, and Annika Persson, "Tax Compliance and Loss Aversion," *American Economic Journal: Economic Policy*, November 2015, *7* (4), 132–164.

Esponda, Ignacio and Demian Pouzo, "Berk-Nash Equilibrium: A Framework for Modeling Agents with Misspecified Agents,," *Econometrica*, 2016, *84* (3), 1093–1130.

Farhi, Emmanuel and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," *NBER Working Paper No. 21524*, 2015.

Feldman, Naomi E., Peter Katuscak, and Laura Kawano, "Taxpayer Confusion: Evidence from the Child Tax Credit," *American Economic Review*, 2016, *106* (3), 807–835.

Fujii, Edwin T and Clifford Hawley, "On the Accuracy of Tax Perceptions," *The Review of Economics and Statistics*, 1988, *70* (2), 344–347.

Gerritsen, Aart, "Optimal Taxation When People Do Not Maximize Well-Being," *Journal of Public Economics*, 2016, pp. –.

Gideon, Michael, "Do Individuals Perceive Income Tax Rates Correctly?," *Public Finance Review*, 2015.

Goldin, Jacob, "Optimal Tax Salience," *Journal of Public Economics*, 2015, *131*, 115–123.

Handel, Ben, Jon Kolstad, and Johannes Spinnewijn, "Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets," *Working Paper*, 2015.

Handel, Benjamin R, "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *The American Economic Review*, 2013, *103* (7), 2643–2682.

— **and Jonathan T Kolstad**, "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, 2015, *105* (8), 2449–2500.

Ito, Koichiro, "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing," *American Economic Review*, February 2014, *104* (2), 537–563.

Kessler, Judd and Michael Norton, "Tax Aversion and Labor Supply," *Journal of Economic Behavior and Organization*, 2016, *123*, 15–28.

Liebman, Jeffrey B., "The Impact of the Earned Income Tax Credit on Incentives and Income Distribution," in "Tax Policy and the Economy, Volume 12" NBER Chapters, National Bureau of Economic Research, Inc, June 1998, pp. 83–120.

— **and Richard Zeckhauser**, "Schmeduling," *Working Paper*, 2004.

Lockwood, Benjamin, "Optimal Income Taxation with Present Bias," *Working Paper*, 2015.

Lockwood, Benjamin B. and Dmitry Taubinsky, "Regressive Sin Taxes," *Working Paper*, 2015.

Lusardi, Annamaria and Olivia S. Mitchell, "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 2014, *52* (1), 5–44.

McCaffery, Edward and Jonathan Baron, "Thinking about Tax," *Psychology, Public Policy, and Law*, 2006, *12*, 106–135.

Miller, Benjamin and Kevin Mumford, "The Salience of Complex Tax Changes: Evidence From the Child and Dependent Care Credit Expansion," *National Tax Journal*, 2015, *68* (3), 477–510.

Mirrlees, James A, "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, 1971, pp. 175–208.

Rees-Jones, Alex, "Loss Aversion Motivates Tax Sheltering: Evidence from U.S. Tax Returns," *Working Paper*, 2014.

Romich, Jennifer L. and Thomas Weisner, "How Families View and Use the EITC: Advance Payment versus Lump Sum Delivery," *National Tax Journal*, 2000, *53* (4), 1245–1266.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *The Review of Economic Studies*, 2001, *68* (1), 205–229.

— , "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

Slemrod, Joel and Jon Bakija, *Taxing Ourselves, 4th Edition: A Citizen's Guide to the Debate over Taxes*, 1 ed., Vol. 1, The MIT Press, 2008.

Spicer, Michael W. and Rodney E. Hero, "Tax Evasion and Heuristics: A Research Note," *Journal of Public Economics*, 1985, *26* (2), 263–267.

Spinnewijn, Johannes, "Heterogeneity, Demand for Insurance and Adverse Selection," *American Economic Journal: Economic Policy*, Forthcoming.

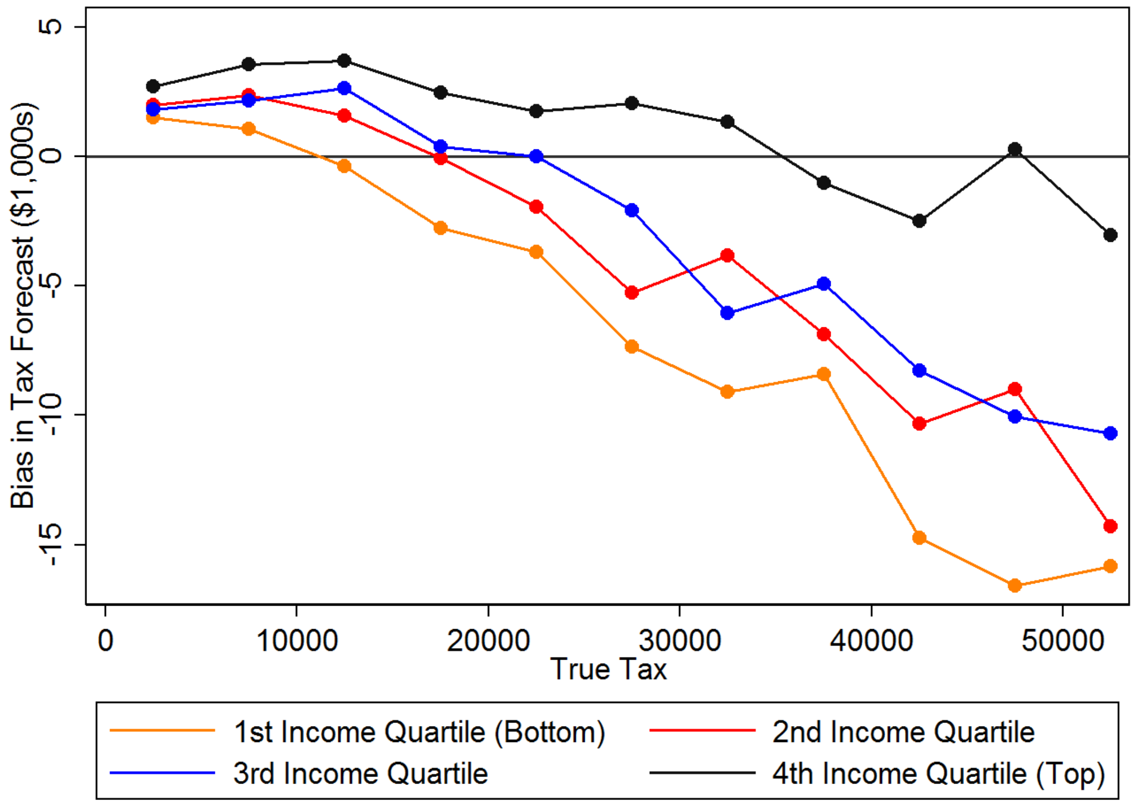
Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *NBER Working Paper No. 22545*, 2016.

Villas-Boas, J. Miguel, "Comparative Statics of Fixed Points," *Journal of Economic Theory*, 2016, *73*, 183–198.

Figure 1: Local-Polynomial Approximations of the Perceived Tax Schedule

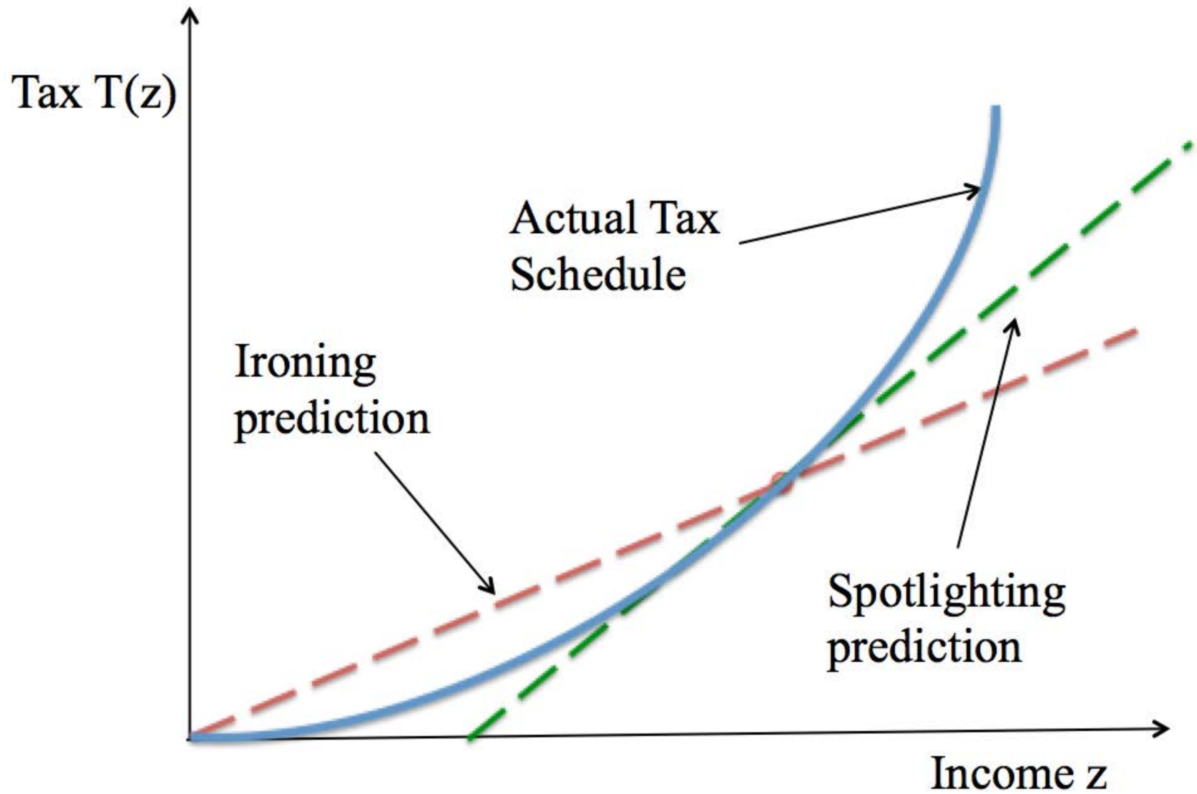
Notes: This figure presents local-polynomial approximations of the perceived relationship between the income earned and taxes owed. Results are plotted separately for single and married-filing-jointly tax filers, as incomes considered in the forecasting task were drawn from filing-status-specific distributions. The first row of figures presents estimates derived from the primary sampling distribution, while row two additionally incorporates perceptions of the high-income sampling distribution. Bandwidth: 10,000. Degree of polynomial: 2. Kernel: Epanechnikov.

Figure 2: Bias in Tax Perceptions, by Respondents' Income Quartile



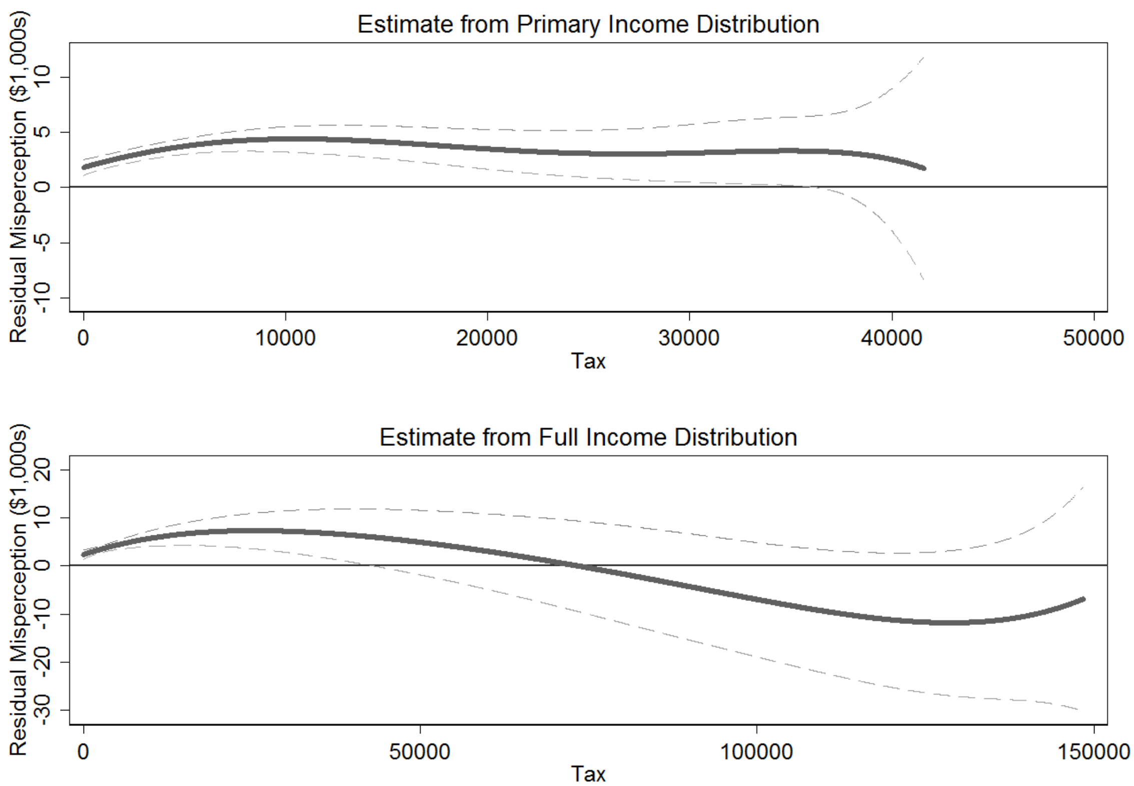
Notes: This figure plots the average bias in tax forecasts as a function of the true tax owed by the hypothetical tax payer. To explore how misperceptions of the tax schedule vary depending on the forecasters' own income, we plot this relationship separately by the income quartile of the respondent. Presented are the estimated coefficients from the regression $(\tilde{T} - T)_{i,f} = \sum \sum_{b,q} \alpha_{b,q} * I(\text{income}_f \in \text{bin}_b) * I(\text{income}_i \in \text{quartile}_q) + \epsilon_{i,f}$, predicting average bias conditional on income quartile and the true tax owed, rounded into $5,000 bins.

Figure 3: Ironing and Spotlighting Heuristics



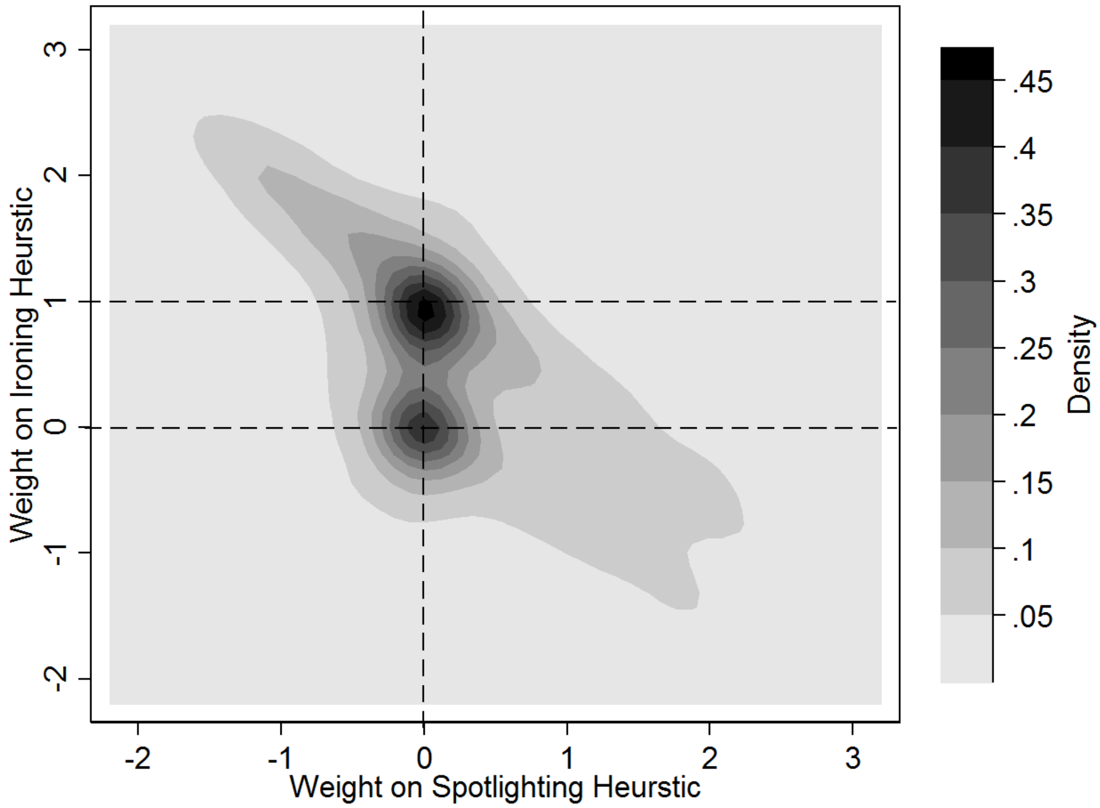
Notes: This figure presents an illustration of the ironing and spotlighting heuristics applied to a generic convex schedule. When using these heuristics, the taxpayer linearizes the convex schedule according to parameters local to his own position on the schedule, indicated by the red dot. Under the ironing heuristic, the taxpayer forecasts by applying his average tax rate at all points, resulting in the observed secant line. Under the spotlighting heuristic, the taxpayer forecasts by applying his marginal tax rate to the change in income that would occur, resulting in the observed tangent line.

Figure 4: Estimates of Residual Tax Misperception



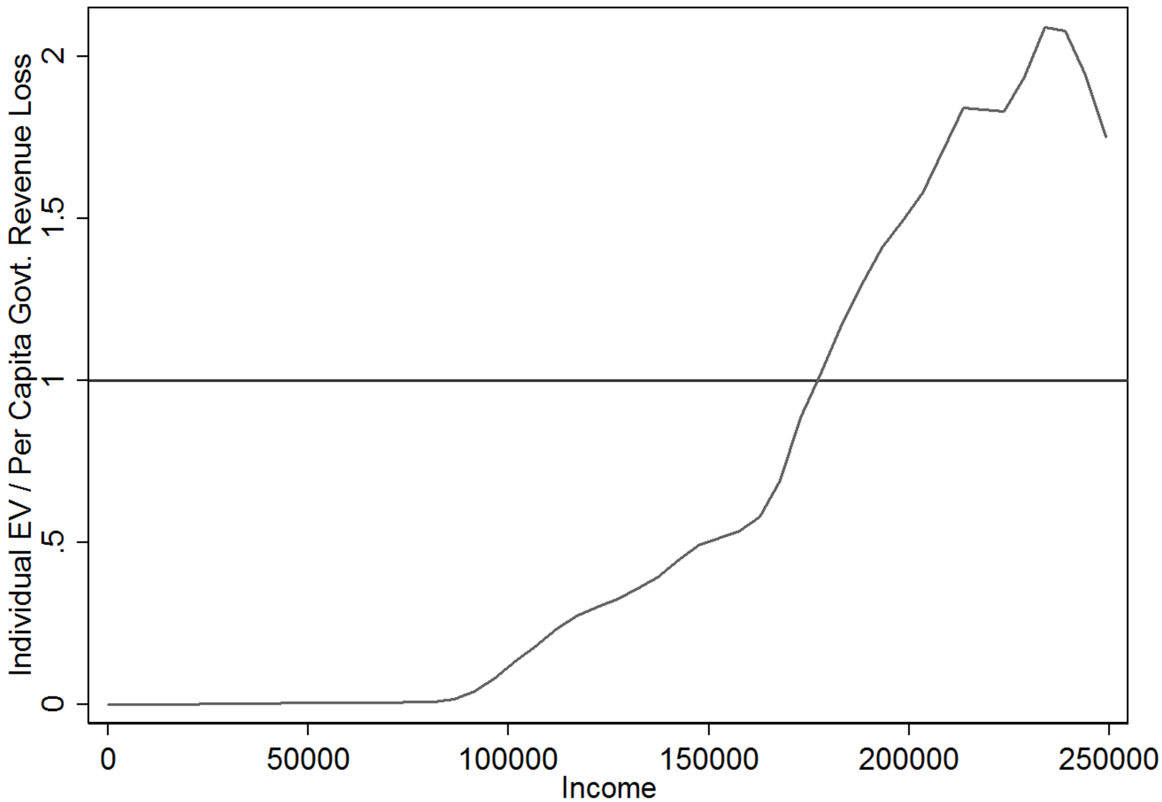
Notes: This figure plots the residual misperception functions estimated in columns 2 and 4 of table 2. These estimates indicate systematic overestimation of the taxes due when true taxes are comparatively small. For sufficiently large tax liabilities, this bias reverses into systematic underestimation of the taxes due.

Figure 5: Individual-Specific Estimates of Heuristic Propensity



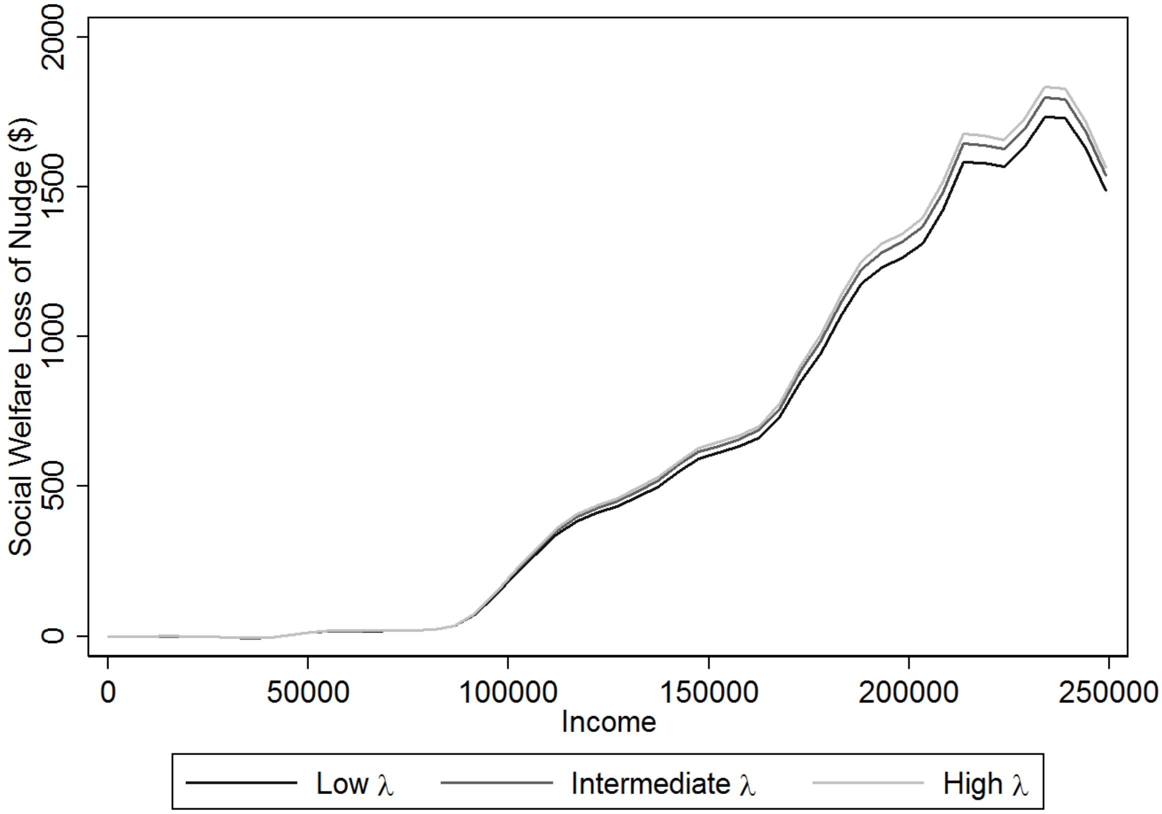
Notes: This table presents a kernel-density estimate of the joint distribution of individual-specific ironing and spotlighting parameters, as estimated in the exercise described in section 2.2. Note that individual-level NLLS regressions failed to converge for 7 respondents. Bandwidth: .2. Kernel: Gaussian.

Figure 6: Consequences of Debiasing Across Income Distribution



Notes: This table presents local polynomial estimates that summarize the regressivity of a fully debiasing nudge. For each taxpayer, we calculate the equivalent variation in income that would lead the biased taxpayer to have the same utility level as achieved by becoming debiased. We plot the ratio of this EV measure to the per capita government revenue loss that occurs with a full debiasing policy. When this ratio takes a value over 1, the individual may be considered a net beneficiary of this policy. As illustrated in this figure, a full debiasing policy is highly attractive to the rich, but not the poor. Bandwidth: 5000. Kernel: Epanechnikov.

Figure 7: Individual Contribution to Social Welfare Across Income Distribution



Notes: This table presents local polynomial estimates of individual contribution to the social welfare loss of full debiasing, conditional on the taxpayer's income. For each taxpayer, we calculate $(U_i^{opt} - U_i^{biased} - \lambda(T_i^{opt} - T_i^{biased}))/\lambda$. This quantifies that individual's contribution to the total social welfare loss that occurs when implementing a full debiasing policy, measured in units of dollars spent on the public good. Social welfare calculations correspond to the models estimated in table 4 for elasticity parameter $k = 3$. Bandwidth: 5000. Kernel: Epanechnikov.

Table 1: Testing for "Flattening" of the Tax Schedule

	All Incomes	Income Quartiles			
	Pooled	1	2	3	4
Estimation Sample: Primary and High-Income Sampling Distributions					
Scale of slope (β)	0.62***	0.53***	0.56***	0.61***	0.76***
	(0.010)	(0.020)	(0.021)	(0.019)	(0.017)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.000
Respondents	4197	1050	1062	1040	1045
Forecasts	58758	14700	14868	14560	14630
Estimation Sample: Primary Sampling Distribution					
Scale of slope (β)	0.82***	0.70***	0.78***	0.80***	0.94***
	(0.013)	(0.029)	(0.030)	(0.026)	(0.023)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.005
Respondents	4197	1050	1062	1040	1045
Forecasts	41970	10500	10620	10400	10450
Estimation Sample: Local Draws					
Scale of slope (β)	0.81***	1.01***	1.07***	0.83***	0.78***
	(0.043)	(0.205)	(0.113)	(0.058)	(0.054)
P-value of H_0: $\beta = 1$	0.000	0.975	0.552	0.003	0.000
Respondents	4197	1050	1062	1040	1045
Forecasts	17937	3143	4074	5293	5427

Notes: Standard errors, clustered by respondent, in parentheses. Presented are coefficients from OLS fixed-effect regressions of the form $\tilde{T}_{i,f} = \beta * T_{i,f} + \nu_i + \epsilon_{i,f}$. The coefficient β can be interpreted as a scaling of the slope induced by the true tax schedule, where estimated values less than 1 indicate a "flattening" of the tax schedule. Two-sided Wald-test p-values, testing the hypothesis that $\beta = 1$, are presented below each regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Parameter Estimates of Heuristic-Perception Model

	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.21***	0.29***	0.47***	0.43***
	(0.037)	(0.052)	(0.048)	(0.095)
γ_S: weight on spotlighting forecast	-0.09*	-0.02	-0.03	-0.02
	(0.050)	(0.057)	(0.062)	(0.076)
Residual misperception function included	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	4197	4197	4197	4197
Forecasts	41970	41970	58758	58758

Notes: Standard errors, clustered by respondent, in parentheses. Presented are non-linear least squares estimates of ironing and spotlighting propensity. Under the ironing heuristic, the taxpayer forecasts by applying his average tax rate at all points. Under the spotlighting heuristic, the taxpayer forecasts by applying his marginal tax rate to the change in income that would occur. See figure 3 for a graphical illustration of the heuristics. The estimated residual misperception function from columns 2 and 4 is plotted in figure 4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Classification of Individuals to Ironing Parameters

		Weight on Spotlighting Heuristic											
		0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Total
Weight on Ironing Heuristic	0%	1155	77	36	31	42	42	42	34	26	39	60	1584
	10%	38	13	8	3	6	3	7	5	13	16	0	112
	20%	24	8	4	4	3	4	2	2	22	0	0	73
	30%	24	11	7	0	1	4	5	30	0	0	0	82
	40%	20	12	5	4	3	1	34	0	0	0	0	79
	50%	20	10	3	4	3	38	0	0	0	0	0	78
	60%	20	11	9	4	40	0	0	0	0	0	0	84
	70%	29	14	6	47	0	0	0	0	0	0	0	96
	80%	20	12	64	0	0	0	0	0	0	0	0	96
	90%	32	81	0	0	0	0	0	0	0	0	0	113
	100%	1155	0	0	0	0	0	0	0	0	0	0	1155
	Total	2537	249	142	97	98	92	90	71	61	55	60	3552

Notes: This table presents the distribution of individual-level classifications of heuristic-use parameters for all respondents with positive tax liability. For each respondent, we compared their 14 tax forecasts to the forecast of the model $\tilde{T_{f,i}} = (1-\gamma_I-\gamma_S)T(z_{f,i}|\theta_i) + \gamma_I\tilde{T}_I(z_{f,i}|z_i^*,\theta_i) + \gamma_S\tilde{T}_S(z_{f,i}|z_i^*,\theta_i) + \epsilon_{f,i}$, corresponding to the model of column 3 of table 2. We calculated this forecast for the grid of values of (γ_I, γ_S) indicated in the table above, and assigned each respondent to the parameter values which minimized the mean squared error of the difference. Analogs of this table applying the restrictions of the other columns of table 2 yield similar results (see appendix tables A3-A5).

Table 4: Welfare Analysis of Debiasing Policy

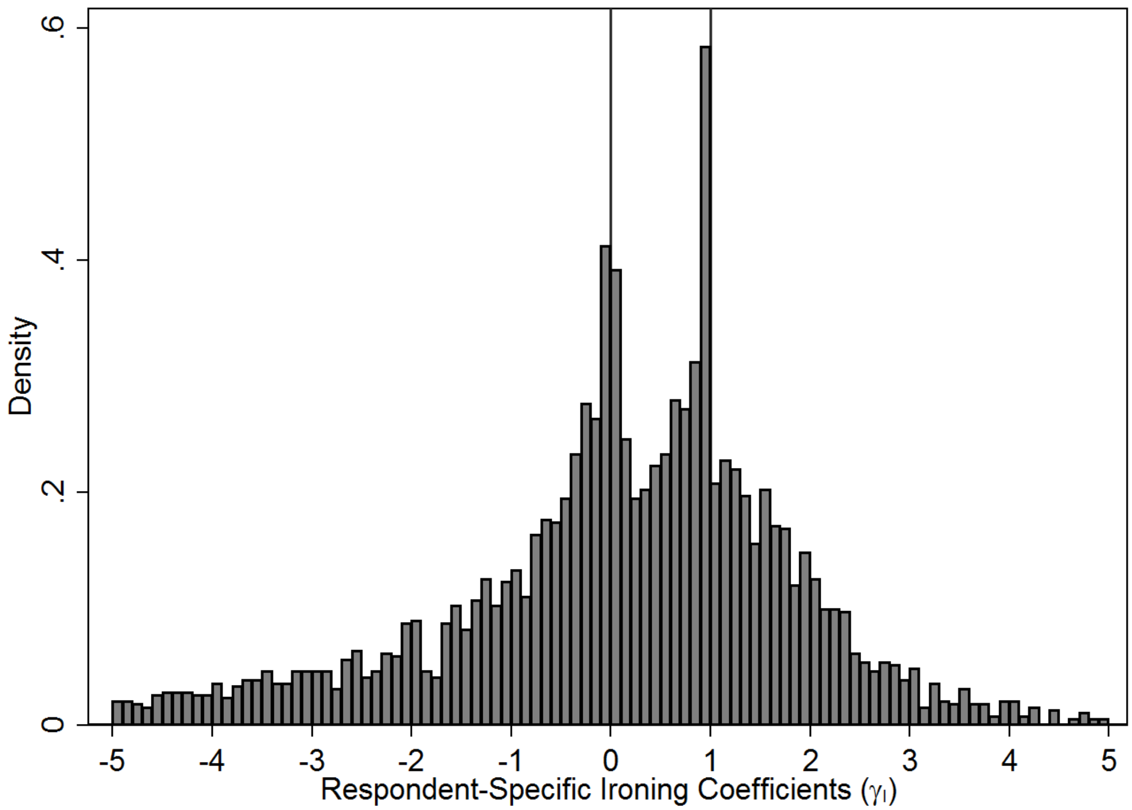
	Elasticity Parameter (k)	Lost Govt. Revenue (%)	*Net Welfare Loss (%)*		
			Low λ	—	High λ
			$\lambda = U'_{10}$	$\lambda = U'_{50}$	$\lambda = U'_{90}$
Elasticity →	1	4.4	4.1	4.3	4.4
	2	2.3	2.2	2.2	2.3
	3	1.6	1.5	1.5	1.6
	4	1.2	1.1	1.2	1.2
	5	1.0	0.9	0.9	1.0

Notes: This table summarizes the simulation exercises described in section 3.4. Each row summarizes the social welfare consequences of fully debiasing our experimental sample, under the assumption that perceptions of tax rates are determined by the estimated model in column 4 of table 2. The first column presents the parameter governing elasticity in our assumed utility model: $U(z) = log(z - T(z) - \frac{(z/w)^{1+k}}{1+k})$. The second column presents the loss of tax revenue that results from full debiasing, expressed as a percentage of baseline tax revenue. The final three columns present estimates of the full welfare effect of debiasing under alternative assumptions on the cost of public funds, expressed as the percentage of tax revenues that a social planner would pay to avoid a full-debiasing nudge.

Appendices (not for publication)

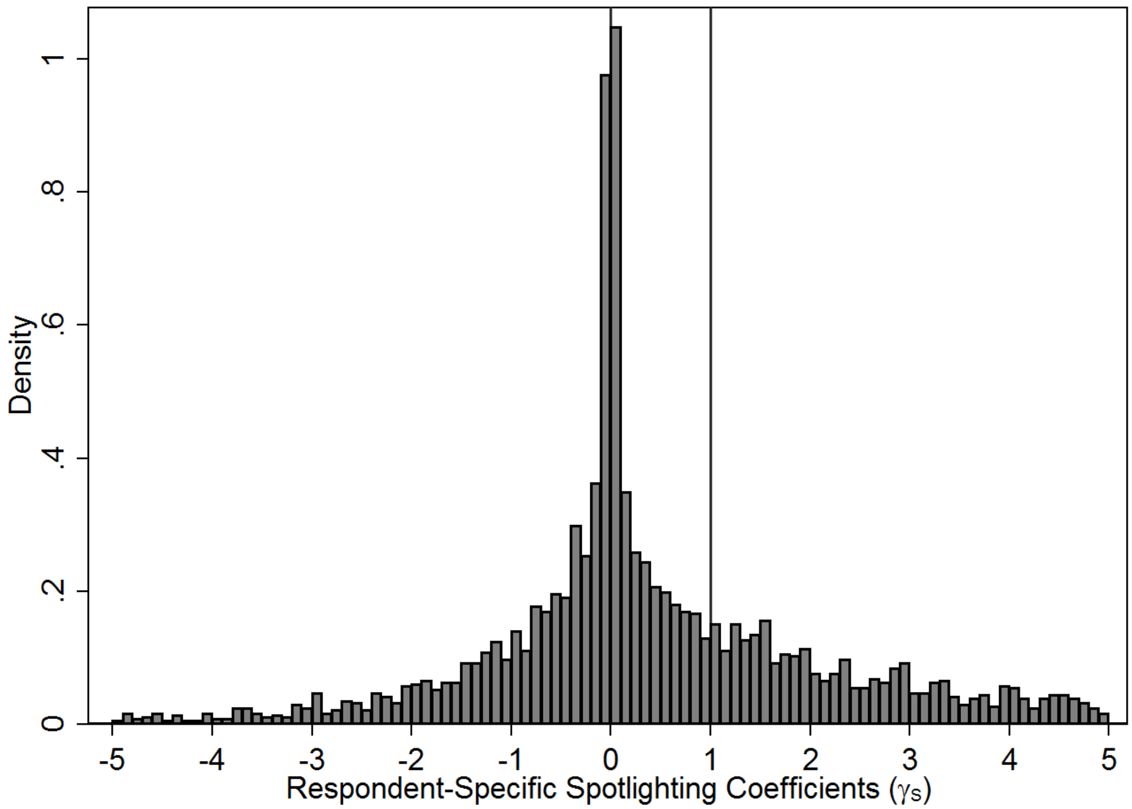
A Supplemental Figure and Tables

Figure A1: Marginal Distribution of Ironing Parameters in Figure 5



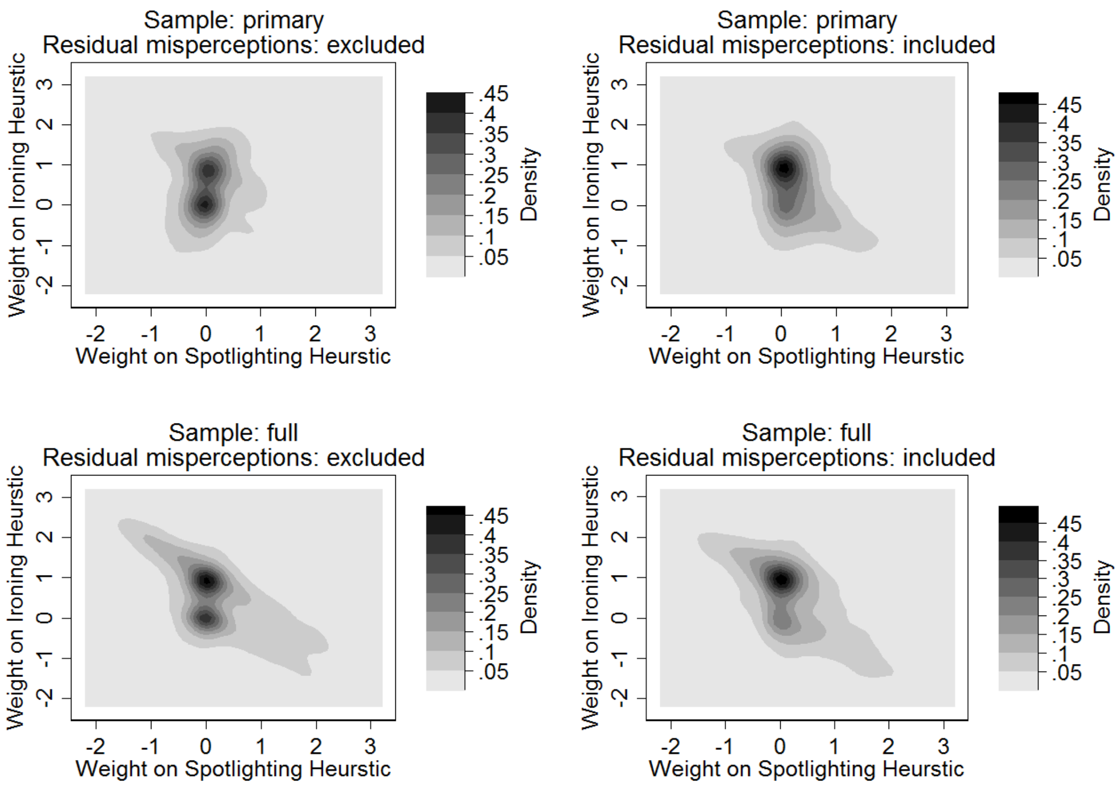
Notes: This figure plots a histogram of the marginal distribution of ironing coefficients corresponding to the joint distribution plotted in 5.

Figure A2: Marginal Distribution of Spotlighting Parameters in Figure 5



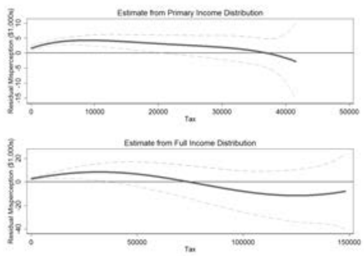
Notes: This figure plots a histogram of the marginal distribution of spotlighting coefficients corresponding to the joint distribution plotted in 5.

Figure A3: Alternative Versions of Figure 5

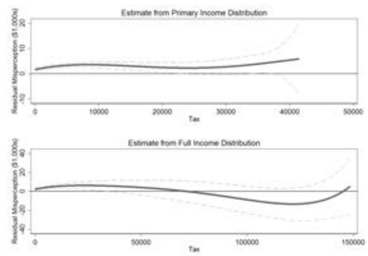


Notes: This figure plots alternative constructions of Figure 5, made to match the restrictions applied in each of the four columns of table 2. Note that individual-level NLLS regressions failed to converge for 13 respondents when using the primary sample, and for 7 respondents when using the full sample.

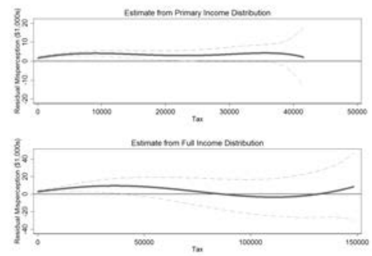
Figure A4: Estimates of Residual Tax Misperception: Alternative Sample Restrictions



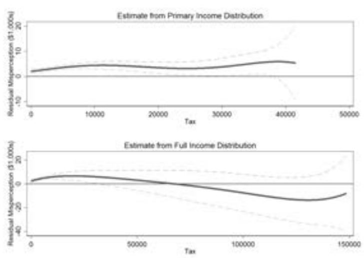
(a) Unemployed



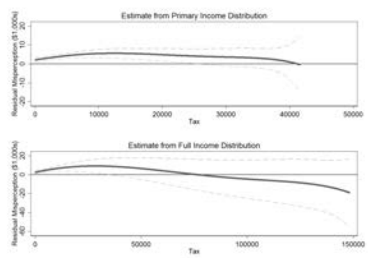
(b) Did own return



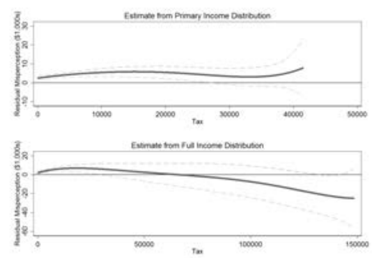
(c) Used tax prep software



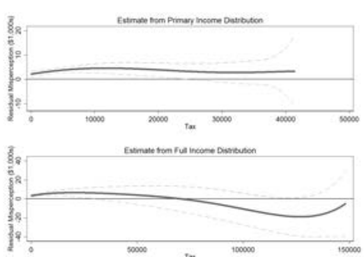
(d) Employed



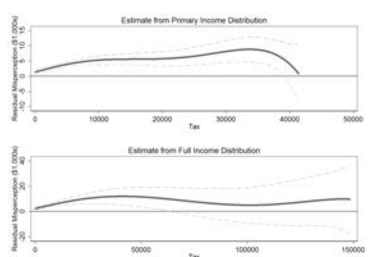
(e) Didn't do own return



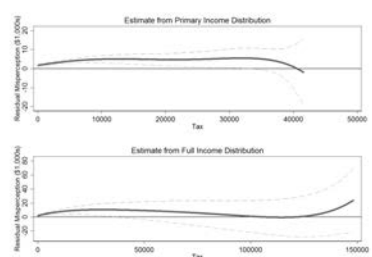
(f) Didn't use tax prep software



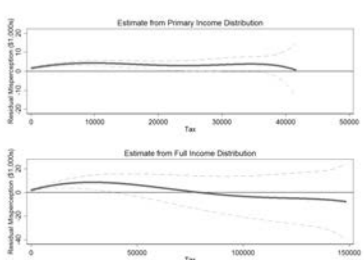
(g) Before tax day



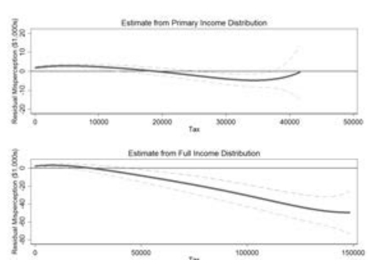
(h) Financially literate



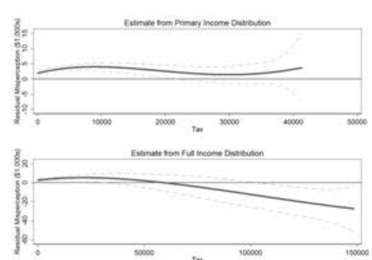
(i) Age < 51



(j) After tax day



(k) Financially illiterate



(l) Age $\geq$ 51

Notes: This figure plots the residual misperception functions estimated in tables A6-A11.

Table A1: Demographics of Sample Compared to Census Data

	In-sample distribution	Census distribution
Gender		
Male	49%	49%
Female	51%	51%
Age		
18-44	39%	48%
45-64	44%	35%
65+	17%	17%
Income		
Under $15,000	16%	12%
$15,000 to $24,999	12%	10%
$25,000 to $34,999	11%	10%
$35,000 to $49,999	15%	13%
$50,000 to $74,999	19%	17%
$75,000 to $99,999	13%	12%
$100,000 to $149,999	10%	14%
$150,000 to $200,999	3%	6%
$200,000 +	1%	6%

Notes: This table presents tabulations of the gender, age, and income distributions reported in our sample for analysis, compared against the distributions reported in the census. Age distributions condition on being 18+.
Source: http://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf and
https://www.census.gov/data/tables/2016/demo/income-poverty/p60-256.html.

Table A2: Parameter Estimates of Heuristic-Perception Model: Alt. Degrees of Polynomial

	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.28***	0.41***	0.29***	0.43***
	(0.052)	(0.094)	(0.052)	(0.095)
γ_S: weight on spotlighting forecast	0.01	0.01	-0.00	-0.01
	(0.055)	(0.075)	(0.056)	(0.076)
Degree of $r(t)$ polynomial	1	1	2	2
γ_I: weight on ironing forecast	0.29***	0.43***	0.29***	0.43***
	(0.052)	(0.095)	(0.052)	(0.095)
γ_S: weight on spotlighting forecast	-0.01	-0.02	-0.02	-0.02
	(0.057)	(0.076)	(0.057)	(0.076)
Degree of $r(t)$ polynomial	3	3	4	4
γ_I: weight on ironing forecast	0.29***	0.43***	0.30***	0.43***
	(0.052)	(0.095)	(0.052)	(0.095)
γ_S: weight on spotlighting forecast	-0.02	-0.02	-0.02	-0.02
	(0.057)	(0.076)	(0.057)	(0.076)
Degree of $r(t)$ polynomial	5	5	6	6
γ_I: weight on ironing forecast	0.30***	0.43***	0.30***	0.43***
	(0.052)	(0.095)	(0.052)	(0.095)
γ_S: weight on spotlighting forecast	-0.02	-0.03	-0.02	-0.02
	(0.057)	(0.076)	(0.057)	(0.076)
Degree of $r(t)$ polynomial	7	7	8	8
γ_I: weight on ironing forecast	0.30***	0.43***	0.30***	0.43***
	(0.052)	(0.095)	(0.052)	(0.095)
γ_S: weight on spotlighting forecast	-0.02	-0.03	-0.02	-0.03
	(0.057)	(0.076)	(0.057)	(0.076)
Degree of $r(t)$ polynomial	9	9	10	10
High-income forecasts included	No	Yes	No	Yes
Respondents	4197	4197	4197	4197
Forecasts	41970	58758	41970	58758

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the estimates from columns 2 and 4 of table 2, while varying the degree of the polynomial used to approximate residual misperception. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Classification of Individuals to Ironing Parameters (Table 2 column 1 analog)

Weight on Ironing Heuristic	**Weight on Spotlighting Heuristic** 0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Total
0%	1407	55	35	23	20	22	16	19	17	31	41	1686
10%	58	10	7	1	2	1	3	4	1	21	0	108
20%	36	11	2	4	3	3	0	3	12	0	0	74
30%	39	6	2	1	3	2	3	19	0	0	0	75
40%	38	12	4	1	1	4	21	0	0	0	0	81
50%	20	8	4	5	2	28	0	0	0	0	0	67
60%	27	16	5	7	24	0	0	0	0	0	0	79
70%	36	13	13	35	0	0	0	0	0	0	0	97
80%	29	18	59	0	0	0	0	0	0	0	0	106
90%	34	91	0	0	0	0	0	0	0	0	0	125
100%	1054	0	0	0	0	0	0	0	0	0	0	1054
Total	2778	240	131	77	55	60	43	45	30	52	41	3552

Notes: This table presents the distribution of individual-level classifications of heuristic-use parameters for all respondents with positive tax liability. For each respondent, we compared their 10 primary-sample tax forecasts to the forecast of the model $\tilde{T_{f,i}} = (1 - \gamma_I - \gamma_S)T(z_{f,i}|\theta_i) + \gamma_I \tilde{T}_I(z_{f,i}|z_i^*, \theta_i) + \gamma_S \tilde{T}_S(z_{f,i}|z_i^*, \theta_i) + \epsilon_{f,i}$. We calculated this forecast for the grid of values of (γ_I, γ_S) indicated in the table above, and assigned each respondent to the parameter values which minimized the mean squared error of the difference.

Table A4: Classification of Individuals to Ironing Parameters (Table 2 column 2 analog)

	Weight on Spotlighting Heuristic											
Weight on Ironing Heuristic	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Total
0%	1101	84	72	70	77	56	42	26	14	24	28	1594
10%	28	13	13	11	4	8	6	8	5	10	0	106
20%	26	11	5	13	7	6	10	7	12	0	0	97
30%	30	8	11	4	9	2	6	18	0	0	0	88
40%	24	19	9	8	7	8	15	0	0	0	0	90
50%	25	16	13	9	11	29	0	0	0	0	0	103
60%	34	19	17	10	28	0	0	0	0	0	0	108
70%	36	24	19	45	0	0	0	0	0	0	0	124
80%	39	21	68	0	0	0	0	0	0	0	0	128
90%	49	88	0	0	0	0	0	0	0	0	0	137
100%	977	0	0	0	0	0	0	0	0	0	0	977
Total	2369	303	227	170	143	109	79	59	31	34	28	3552

Notes: This table presents the distribution of individual-level classifications of heuristic-use parameters for all respondents with positive tax liability. For each respondent, we compared their 10 primary-sample tax forecasts to the forecast of the model $\tilde{T_{f,i}} = (1 - \gamma_I - \gamma_S)(T(z_{f,i}|\theta_i) + \hat{r}(T(z_{f,i}|\theta_i))) + \gamma_I \tilde{T}_I(z_{f,i}|z_i^*, \theta_i) + \gamma_S \tilde{T}_S(z_{f,i}|z_i^*, \theta_i) + \epsilon_{f,i}$, where $\hat{r}$ represents the fitted residual misperception function estimated in column 2 of table 2. We calculated this forecast for the grid of values of (γ_I, γ_S) indicated in the table above, and assigned each respondent to the parameter values which minimized the mean squared error of the difference.

Table A5: Classification of Individuals to Ironing Parameters (Table 2 column 4 analog)

	Weight on Spotlighting Heuristic											
Weight on Ironing Heuristic	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	Total
0%	1118	97	47	33	49	27	35	20	23	15	25	1489
10%	29	16	11	3	9	6	5	5	11	10	0	105
20%	36	17	7	8	3	3	7	4	8	0	0	93
30%	32	9	6	3	1	2	3	20	0	0	0	76
40%	37	11	9	1	0	2	23	0	0	0	0	83
50%	26	12	10	9	7	17	0	0	0	0	0	81
60%	34	20	17	6	24	0	0	0	0	0	0	101
70%	39	27	18	25	0	0	0	0	0	0	0	109
80%	50	30	56	0	0	0	0	0	0	0	0	136
90%	56	77	0	0	0	0	0	0	0	0	0	133
100%	1146	0	0	0	0	0	0	0	0	0	0	1146
Total	2603	316	181	88	93	57	73	49	42	25	25	3552

Notes: This table presents the distribution of individual-level classifications of heuristic-use parameters for all respondents with positive tax liability. For each respondent, we compared their 14 tax forecasts to the forecast of the model $\tilde{T_{f,i}} = (1-\gamma_I-\gamma_S)(T(z_{f,i}|\theta_i)+\hat{r}(T(z_{f,i}|\theta_i)))+\gamma_I\tilde{T}_I(z_{f,i}|z_i^*,\theta_i)+\gamma_S\tilde{T}_S(z_{f,i}|z_i^*,\theta_i)+\epsilon_{f,i}$, where $\hat{r}$ represents the fitted residual misperception function estimated in column 4 of table 2. We calculated this forecast for the grid of values of (γ_I, γ_S) indicated in the table above, and assigned each respondent to the parameter values which minimized the mean squared error of the difference.

Table A6: Parameter Estimates of Heuristic-Perception Model: Robustness to Employment Status

Restricted to unemployed respondents				
γ_I: weight on ironing forecast	0.25***	0.31***	0.56***	0.53***
	(0.064)	(0.090)	(0.073)	(0.136)
γ_S: weight on spotlighting forecast	-0.11	-0.06	-0.13	-0.13
	(0.089)	(0.098)	(0.093)	(0.110)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	1685	1685	1685	1685
Forecasts	16850	16850	23590	23590

Restricted to employed respondents				
γ_I: weight on ironing forecast	0.18***	0.28***	0.41***	0.35***
	(0.045)	(0.065)	(0.064)	(0.130)
γ_S: weight on spotlighting forecast	-0.10*	0.01	0.03	0.06
	(0.056)	(0.068)	(0.083)	(0.105)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2512	2512	2512	2512
Forecasts	25120	25120	35168	35168

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the analysis of table 2, restricting the sample to employed or unemployed respondents. Estimated residual misperception functions are plotted in appendix table A4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Parameter Estimates of Heuristic-Perception Model: Robustness to Who Completed Tax Return

Restricted to respondents who completed their own tax return				
γ_I: weight on ironing forecast	0.12**	0.20***	0.36***	0.31**
	(0.050)	(0.066)	(0.061)	(0.123)
γ_S: weight on spotlighting forecast	-0.04	0.03	0.05	0.07
	(0.065)	(0.073)	(0.077)	(0.098)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2301	2301	2301	2301
Forecasts	23010	23010	32214	32214

Restricted to respondents who did not complete their own tax return				
	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.31***	0.39***	0.59***	0.58***
	(0.054)	(0.081)	(0.077)	(0.144)
γ_S: weight on spotlighting forecast	-0.15*	-0.05	-0.13	-0.14
	(0.077)	(0.089)	(0.102)	(0.120)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	1896	1896	1896	1896
Forecasts	18960	18960	26544	26544

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the analysis of table 2, restricting the sample to respondents who did, or did not, complete their own tax return. Estimated residual misperception functions are plotted in appendix table A4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Parameter Estimates of Heuristic-Perception Model: Robustness to Use of Tax Preparation Software

Restricted to respondents who used tax preparation software				
γ_I: weight on ironing forecast	0.19***	0.28***	0.38***	0.42***
	(0.052)	(0.075)	(0.064)	(0.135)
γ_S: weight on spotlighting forecast	-0.07	0.00	0.07	0.04
	(0.064)	(0.075)	(0.081)	(0.105)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2242	2242	2242	2242
Forecasts	22420	22420	31388	31388
Restricted to respondents who did not use tax preparation software				
	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.20***	0.31***	0.52***	0.43***
	(0.055)	(0.074)	(0.079)	(0.144)
γ_S: weight on spotlighting forecast	-0.08	0.02	-0.10	-0.07
	(0.075)	(0.085)	(0.104)	(0.121)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	1590	1590	1590	1590
Forecasts	15900	15900	22260	22260

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the analysis of table 2, restricting the sample to respondents who did, or did not, use tax preparation software. 365 respondents who did not know if tax preparation software was used are excluded. Estimated residual misperception functions are plotted in appendix table A4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Parameter Estimates of Heuristic-Perception Model: Differences by Survey Date

Restricted to responses on or before tax day				
	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.23***	0.30***	0.42***	0.34**
	(0.054)	(0.075)	(0.073)	(0.133)
γ_S: weight on spotlighting forecast	-0.06	0.03	0.07	0.11
	(0.074)	(0.086)	(0.094)	(0.112)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2080	2080	2080	2080
Forecasts	20800	20800	29120	29120

Restricted to responses after tax day				
	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.19***	0.29***	0.51***	0.52***
	(0.051)	(0.071)	(0.064)	(0.135)
γ_S: weight on spotlighting forecast	-0.12*	-0.05	-0.12	-0.14
	(0.067)	(0.076)	(0.082)	(0.104)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2117	2117	2117	2117
Forecasts	21170	21170	29638	29638

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the analysis of table 2, restricting the sample by the time period when the respondent completed the survey. Estimated residual misperception functions are plotted in appendix table A4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Parameter Estimates of Heuristic-Perception Model: Differences by Financial Literacy

Restricted to financially literate respondents				
	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.02	0.26***	0.21***	0.35***
	(0.040)	(0.059)	(0.060)	(0.108)
γ_S: weight on spotlighting forecast	-0.11**	-0.02	0.10	0.02
	(0.049)	(0.061)	(0.082)	(0.095)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	1899	1899	1899	1899
Forecasts	18990	18990	26586	26586

Restricted to financially illiterate respondents				
	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.39***	0.23***	0.78***	0.36**
	(0.062)	(0.087)	(0.073)	(0.150)
γ_S: weight on spotlighting forecast	-0.18**	-0.06	-0.29***	-0.09
	(0.083)	(0.091)	(0.091)	(0.114)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2298	2298	2298	2298
Forecasts	22980	22980	32172	32172

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the analysis of table 2, restricting the sample by respondents' financial literacy classification. Respondents are coded if they answered all of the Lusardi and Mitchell "Big Three" financial literacy questions correctly; otherwise, they are classified as financially illiterate. Estimated residual misperception functions are plotted in appendix table A4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Parameter Estimates of Heuristic-Perception Model: Differences by Age

Restricted to age < 51				
	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.24***	0.40***	0.47***	0.55***
	(0.053)	(0.077)	(0.070)	(0.140)
γ_S: weight on spotlighting forecast	-0.13*	-0.08	-0.02	-0.08
	(0.072)	(0.081)	(0.089)	(0.108)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2078	2078	2078	2078
Forecasts	20780	20780	29092	29092

Restricted to age ≥ 51				
	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.18***	0.21***	0.47***	0.34***
	(0.051)	(0.070)	(0.067)	(0.129)
γ_S: weight on spotlighting forecast	-0.06	0.03	-0.05	0.01
	(0.069)	(0.081)	(0.087)	(0.108)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2119	2119	2119	2119
Forecasts	21190	21190	29666	29666

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the analysis of table 2, restricting the sample to those above and below the median age of 51. Estimated residual misperception functions are plotted in appendix table A4. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Welfare Analysis of Debiasing Policy: Simulation with Alternative CRRA Utility

Panel A: CRRA parameter = 0.5					
	Elasticity Parameter (k)	Lost Govt. Revenue (%)	*Net Welfare Loss (%)*		
			Low λ	—	High λ
			$\lambda = U'_{10}$	$\lambda = U'_{50}$	$\lambda = U'_{90}$
Elasticity →	1	4.4	4.1	4.2	4.3
	2	2.3	2.2	2.2	2.3
	3	1.6	1.5	1.5	1.6
	4	1.2	1.1	1.1	1.2
	5	1.0	0.9	0.9	0.9

Panel B: CRRA parameter = 2					
	Elasticity Parameter (k)	Lost Govt. Revenue (%)	*Net Welfare Loss (%)*		
			Low λ	—	High λ
			$\lambda = U'_{10}$	$\lambda = U'_{50}$	$\lambda = U'_{90}$
Elasticity →	1	4.4	4.1	4.4	4.4
	2	2.3	2.2	2.3	2.3
	3	1.6	1.5	1.6	1.6
	4	1.2	1.1	1.2	1.2
	5	1.0	0.9	1.0	1.0

Notes: This table reproduces the analysis of table 4, but generalizes the ln utility component to CRRA preferences of differing parameters governing redistributive preferences.

Table A13: Welfare Analysis of Debiasing Policy: Simulation with Locally Estimated Misperceptions

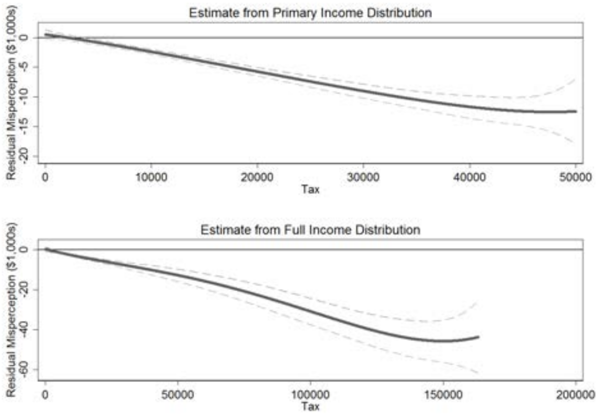
Panel A: Applying Full-Sample Estimate					
	Elasticity Parameter (k)	Lost Govt. Revenue (%)	*Net Welfare Loss (%)* Low λ $\lambda = U'_{10}$	— $\lambda = U'_{50}$	High λ $\lambda = U'_{90}$
Elasticity →	1	8.2	7.1	7.8	8.0
	2	4.3	3.8	4.1	4.2
	3	2.9	2.6	2.8	2.9
	4	2.2	1.9	2.1	2.2
	5	1.8	1.6	1.7	1.8

Panel B: Applying Quartile-Specific Estimates					
	Elasticity Parameter (k)	Lost Govt. Revenue (%)	*Net Welfare Loss (%)* Low λ $\lambda = U'_{10}$	— $\lambda = U'_{50}$	High λ $\lambda = U'_{90}$
Elasticity →	1	8.0	7.0	7.6	7.9
	2	4.2	3.7	4.0	4.2
	3	2.8	2.5	2.7	2.8
	4	2.2	1.9	2.1	2.2
	5	1.8	1.6	1.7	1.7

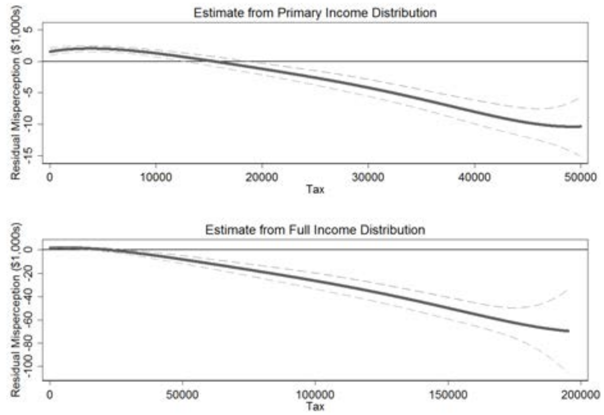
Notes: This table reproduces the analysis of table 4, but assumes that taxpayers apply the locally estimated flattening parameter (bottom panel of table 1) rather than the heuristic forecasting model. The top panel applies the estimate of column 1, and the bottom panel applies the quartile-specific estimates of columns 2-5.

B Influence of the Inclusion of Additional Taxes

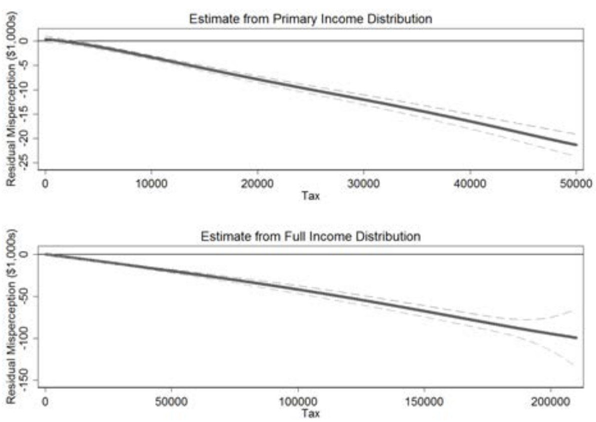
Figure B1: Estimates of Residual Tax Misperception: Additional Taxes Included



(a) FICA included



(b) State tax included



(c) FICA and state tax included

Notes: This figure plots the residual misperception functions estimated in tables B1-B3.

Table B1: Parameter Estimates of Heuristic-Perception Model: Assuming Respondents Included FICA Tax in Forecasts

	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.88***	0.32***	1.06***	0.55***
	(0.046)	(0.065)	(0.049)	(0.098)
γ_S: weight on spotlighting forecast	-0.24***	-0.03	-0.39***	-0.12
	(0.063)	(0.072)	(0.065)	(0.081)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	4197	4197	4197	4197
Forecasts	41970	41970	58758	58758

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the estimates of table 2, but calculates the true tax, ATR, and MTR to include FICA taxes. Estimated residual misperception functions are plotted in appendix figure B1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Parameter Estimates of Heuristic-Perception Model: Assuming Respondents Included State Tax in Forecasts

	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.47***	0.23***	0.79***	0.37***
	(0.032)	(0.045)	(0.040)	(0.079)
γ_S: weight on spotlighting forecast	-0.16***	-0.02	-0.29***	-0.06
	(0.044)	(0.050)	(0.053)	(0.065)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	4163	4163	4163	4163
Forecasts	41630	41630	58282	58282

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the estimates of table 2, but calculates the true tax, ATR, and MTR to include State taxes. Estimated residual misperception functions are plotted in appendix figure B1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Parameter Estimates of Heuristic-Perception Model: Assuming Respondents Included State and FICA Tax in Forecasts

	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	1.09***	0.25***	1.31***	0.47***
	(0.039)	(0.054)	(0.040)	(0.080)
γ_S: weight on spotlighting forecast	-0.36***	-0.04	-0.60***	-0.14**
	(0.054)	(0.061)	(0.054)	(0.068)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	4163	4163	4163	4163
Forecasts	41630	41630	58282	58282

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the estimates of table 2, but calculates the true tax, ATR, and MTR to include both State and FICA taxes. Estimated residual misperception functions are plotted in appendix figure B1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B4: Welfare Analysis of Debiasing Policy

Panel A: State + FICA Taxes Perfectly Perceived Before and After Nudge

	Elasticity Parameter (k)	Lost Govt. Revenue (%)	*Net Welfare Loss (%)* Low λ $\lambda = U'_{10}$	— $\lambda = U'_{50}$	High λ $\lambda = U'_{90}$
Elasticity →	1	4.9	4.6	4.8	4.9
	2	2.6	2.4	2.5	2.6
	3	1.8	1.7	1.7	1.7
	4	1.3	1.3	1.3	1.3
	5	1.1	1.0	1.1	1.1

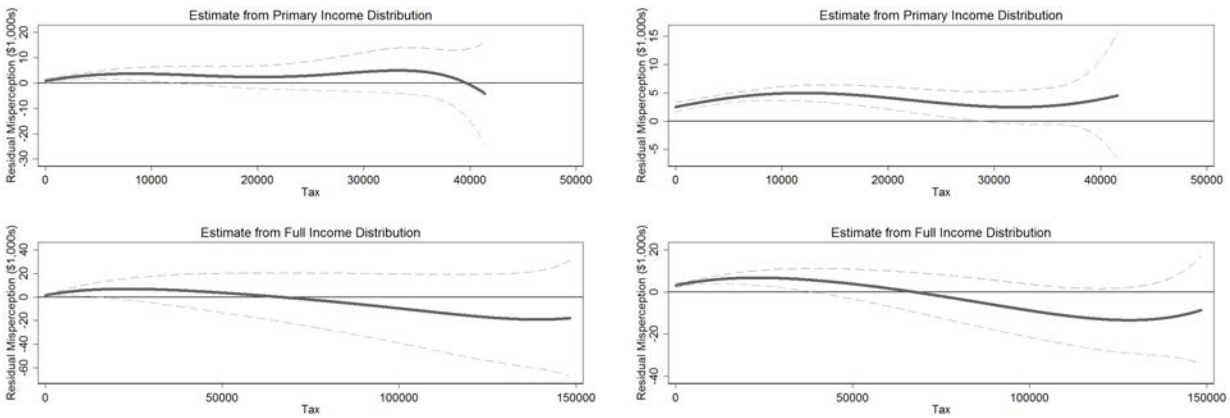
Panel B: State + FICA Taxes Ignored Before Nudge, Perfectly Perceived After

	Elasticity Parameter (k)	Lost Govt. Revenue (%)	*Net Welfare Loss (%)* Low λ $\lambda = U'_{10}$	— $\lambda = U'_{50}$	High λ $\lambda = U'_{90}$
Elasticity →	1	22.8	9.7	17.1	21.5
	2	12.8	5.9	9.8	12.1
	3	8.9	4.2	6.9	8.5
	4	6.8	3.2	5.3	6.5
	5	5.5	2.6	4.3	5.3

Notes: This table reproduces the analysis of table 4 under two alternative treatments of State and FICA taxes. In both panels, we include both State taxes and FICA taxes when calculating the actual tax burden, although we model these as externally imposed taxes, with revenue not controlled by our social planner. In the simulations reported in the top panel, we assume that taxpayers correctly perceive State and FICA taxes both before and after the nudge: prior to the nudge, tax misperceptions are limited to the federal income tax. In the simulations reported in the bottom panel, we assume that taxpayers completely ignore State and FICA taxes before the nudge, but correctly account for them after the nudge.

C Influence of Similarity to Fred

Figure C1: Estimates of Residual Tax Misperception: Robustness to Similarity with Hypothetical Filer



(a) Simple

(b) Complex

Notes: This figure plots the residual misperception functions estimated in columns 2 and 4 of table C3.

Table C1: Testing for "Flattening" of the Tax Schedule: Robustness to Similarity with Hypothetical Filer (Similar Filers)

	All Incomes	Income Quartiles			
	Pooled	1	2	3	4
Estimation Sample: Primary and High-Income Sampling Distributions					
Scale of slope (β)	0.55***	0.50***	0.53***	0.57***	0.73***
	(0.018)	(0.028)	(0.032)	(0.039)	(0.057)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.000
Respondents	1357	465	436	313	143
Forecasts	18998	6510	6104	4382	2002
Estimation Sample: Primary Sampling Distribution					
Scale of slope (β)	0.77***	0.70***	0.78***	0.74***	0.94***
	(0.027)	(0.047)	(0.051)	(0.054)	(0.073)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.395
Respondents	1357	465	436	313	143
Forecasts	13570	4650	4360	3130	1430
Estimation Sample: Local Draws					
Scale of slope (β)	0.78***	0.91***	1.10***	0.78***	0.72***
	(0.077)	(0.161)	(0.208)	(0.084)	(0.133)
P-value of H_0: $\beta = 1$	0.005	0.587	0.628	0.009	0.037
Respondents	1357	465	436	313	143
Forecasts	5489	1408	1725	1594	762

Notes: Standard errors, clustered by respondent, in parentheses. Presented are coefficients from OLS fixed-effect regressions of the form $\tilde{T}_{i,f} = \beta * T_{i,f} + \nu_i + \epsilon_{i,f}$. The coefficient β can be interpreted as a scaling of the slope induced by the true tax schedule, where estimated values less than 1 indicate a "flattening" of the tax schedule. Two-sided Wald-test p-values, testing the hypothesis that $\beta = 1$, are presented below each regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C2: Testing for "Flattening" of the Tax Schedule: Robustness to Similarity with Hypothetical Filer (Dissimilar Filers)

	All Incomes	Income Quartiles			
	Pooled	1	2	3	4
Estimation Sample: Primary and High-Income Sampling Distributions					
Scale of slope (β)	0.65***	0.56***	0.57***	0.62***	0.76***
	(0.011)	(0.027)	(0.027)	(0.021)	(0.018)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.000
Respondents	2840	585	626	727	902
Forecasts	39760	8190	8764	10178	12628
Estimation Sample: Primary Sampling Distribution					
Scale of slope (β)	0.84***	0.71***	0.79***	0.83***	0.94***
	(0.015)	(0.036)	(0.036)	(0.029)	(0.024)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.007
Respondents	2840	585	626	727	902
Forecasts	28400	5850	6260	7270	9020
Estimation Sample: Local Draws					
Scale of slope (β)	0.81***	1.17**	1.05***	0.86***	0.79***
	(0.049)	(0.476)	(0.127)	(0.075)	(0.059)
P-value of H_0: $\beta = 1$	0.000	0.724	0.723	0.055	0.000
Respondents	2840	585	626	727	902
Forecasts	12448	1735	2349	3699	4665

Notes: Standard errors, clustered by respondent, in parentheses. Presented are coefficients from OLS fixed-effect regressions of the form $\tilde{T}_{i,f} = \beta * T_{i,f} + \nu_i + \epsilon_{i,f}$. The coefficient β can be interpreted as a scaling of the slope induced by the true tax schedule, where estimated values less than 1 indicate a "flattening" of the tax schedule. Two-sided Wald-test p-values, testing the hypothesis that $\beta = 1$, are presented below each regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C3: Parameter Estimates of Heuristic-Perception Model: Robustness to Similarity with Hypothetical Filer

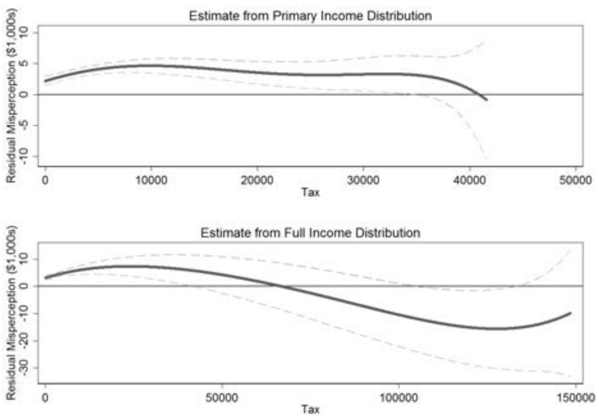
Restricted to respondents with simple tax returns				
γ_I: weight on ironing forecast	0.29***	0.46***	0.63***	0.57***
	(0.098)	(0.145)	(0.104)	(0.210)
γ_S: weight on spotlighting forecast	-0.18	-0.21	-0.18	-0.18
	(0.147)	(0.153)	(0.133)	(0.156)
Residual misperception function included	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	1357	1357	1357	1357
Forecasts	13570	13570	18998	18998

Restricted to respondents with complex tax returns				
γ: weight on ironing forecast	0.18***	0.23***	0.41***	0.34***
	(0.040)	(0.059)	(0.054)	(0.110)
γ§: weight on spotlighting forecast	-0.08	0.05	0.01	0.04
	(0.048)	(0.058)	(0.069)	(0.088)
Residual misperception function included	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	2840	2840	2840	2840
Forecasts	28400	28400	39760	39760

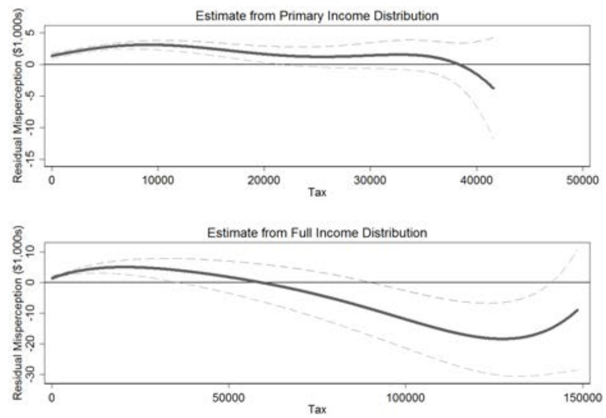
Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the analysis of table 2, restricting the sample to respondents who do (or do not) have the same simple tax filing behavior that was specified in our scenarios. Estimated residual misperception functions are plotted in appendix table C1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Robustness to Data Cleaning Decisions

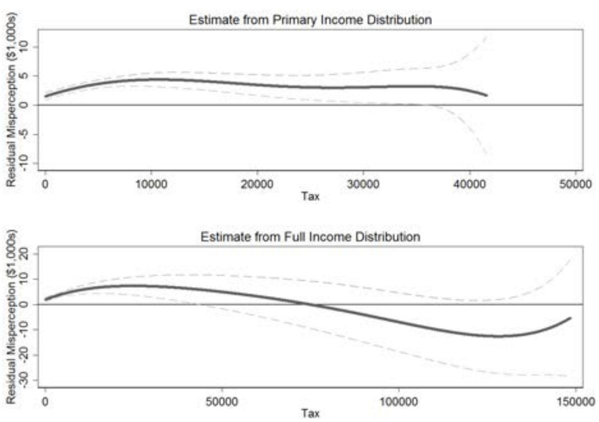
Figure D1: Estimates of Residual Tax Misperception: Alternative Outlier Control



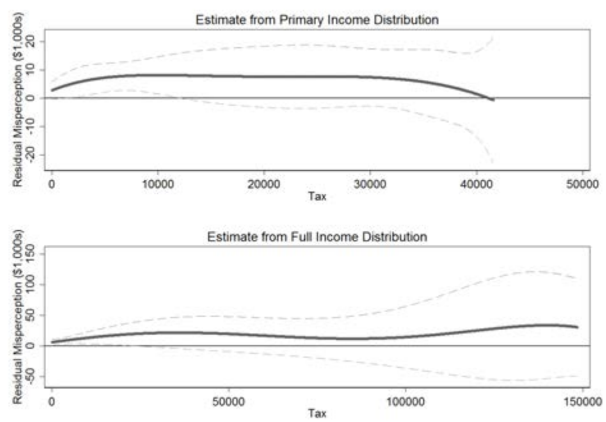
(a) Reincuding those who failed attention check



(b) Windsorized at 5th and 9th pctile



(c) Winsorized at 0% and 100% tax rates



(d) No winsorization

Notes: This figure plots the residual misperception functions estimated in tables D2 and D6.

Table D1: Testing for "Flattening" of the Tax Schedule: Including Respondents Who Failed Attention Check

	All Incomes	Income Quartiles			
	Pooled	1	2	3	4
Estimation Sample: Primary and High-Income Sampling Distributions					
Scale of slope (β)	0.60***	0.51***	0.53***	0.59***	0.74***
	(0.009)	(0.019)	(0.020)	(0.018)	(0.017)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.000
Respondents	4633	1159	1167	1149	1158
Forecasts	64862	16226	16338	16086	16212
Estimation Sample: Primary Sampling Distribution					
Scale of slope (β)	0.81***	0.69***	0.76***	0.79***	0.93***
	(0.013)	(0.029)	(0.030)	(0.026)	(0.023)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.003
Respondents	4633	1159	1167	1149	1158
Forecasts	46330	11590	11670	11490	11580
Estimation Sample: Local Draws					
Scale of slope (β)	0.78***	0.99***	1.09***	0.82***	0.76***
	(0.046)	(0.221)	(0.110)	(0.056)	(0.058)
P-value of H_0: $\beta = 1$	0.000	0.977	0.424	0.002	0.000
Respondents	4633	1159	1167	1149	1158
Forecasts	19672	3436	4419	5788	6029

Notes: Standard errors, clustered by respondent, in parentheses. Presented are coefficients from OLS fixed-effect regressions of the form $\tilde{T}_{i,f} = \beta * T_{i,f} + \nu_i + \epsilon_{i,f}$. The coefficient β can be interpreted as a scaling of the slope induced by the true tax schedule, where estimated values less than 1 indicate a "flattening" of the tax schedule. Two-sided Wald-test p-values, testing the hypothesis that $\beta = 1$, are presented below each regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D2: Parameter Estimates of Heuristic-Perception Model: Including Respondents Who Failed Attention Check

	(1)	(2)	(3)	(4)
γ_I: weight on ironing forecast	0.21***	0.29***	0.47***	0.40***
	(0.040)	(0.056)	(0.047)	(0.094)
γ_S: weight on spotlighting forecast	-0.09*	0.00	-0.00	0.02
	(0.051)	(0.058)	(0.059)	(0.074)
Residual misperception function included?	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	4633	4633	4633	4633
Forecasts	46330	46330	64862	64862

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the estimates of table 2 after including subjects who failed the attention check in the estimation sample. Estimated residual misperception functions are plotted in appendix table D1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D3: Testing for "Flattening" of the Tax Schedule: Winsorized at 5th and 95th percentile forecast values

	All Incomes	Income Quartiles			
	Pooled	1	2	3	4
Estimation Sample: Primary and High-Income Sampling Distributions					
Scale of slope (β)	0.58***	0.49***	0.51***	0.58***	0.72***
	(0.008)	(0.015)	(0.015)	(0.015)	(0.014)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.000
Respondents	4197	1050	1062	1040	1045
Forecasts	58758	14700	14868	14560	14630
Estimation Sample: Primary Sampling Distribution					
Scale of slope (β)	0.77***	0.64***	0.70***	0.76***	0.91***
	(0.010)	(0.020)	(0.021)	(0.020)	(0.018)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.000
Respondents	4197	1050	1062	1040	1045
Forecasts	41970	10500	10620	10400	10450
Estimation Sample: Local Draws					
Scale of slope (β)	0.76***	0.81***	0.95***	0.81***	0.74***
	(0.036)	(0.092)	(0.070)	(0.034)	(0.047)
P-value of H_0: $\beta = 1$	0.000	0.036	0.442	0.000	0.000
Respondents	4197	1050	1062	1040	1045
Forecasts	17937	3143	4074	5293	5427

Notes: Standard errors, clustered by respondent, in parentheses. Presented are coefficients from OLS fixed-effect regressions of the form $\tilde{T}_{i,f} = \beta * T_{i,f} + \nu_i + \epsilon_{i,f}$. The coefficient β can be interpreted as a scaling of the slope induced by the true tax schedule, where estimated values less than 1 indicate a "flattening" of the tax schedule. Two-sided Wald-test p-values, testing the hypothesis that $\beta = 1$, are presented below each regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D4: Testing for "Flattening" of the Tax Schedule: Winsorized to 0-100% tax rates

	All Incomes	Income Quartiles			
	Pooled	1	2	3	4
Estimation Sample: Primary and High-Income Sampling Distributions					
Scale of slope (β)	0.62***	0.54***	0.55***	0.61***	0.75***
	(0.009)	(0.020)	(0.020)	(0.019)	(0.017)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.000
Respondents	4197	1050	1062	1040	1045
Forecasts	58758	14700	14868	14560	14630
Estimation Sample: Primary Sampling Distribution					
Scale of slope (β)	0.83***	0.71***	0.79***	0.81***	0.95***
	(0.013)	(0.029)	(0.030)	(0.026)	(0.023)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.020
Respondents	4197	1050	1062	1040	1045
Forecasts	41970	10500	10620	10400	10450
Estimation Sample: Local Draws					
Scale of slope (β)	0.81***	1.03***	1.09***	0.84***	0.79***
	(0.045)	(0.195)	(0.111)	(0.055)	(0.057)
P-value of H_0: $\beta = 1$	0.000	0.869	0.439	0.003	0.000
Respondents	4197	1050	1062	1040	1045
Forecasts	17937	3143	4074	5293	5427

Notes: Standard errors, clustered by respondent, in parentheses. Presented are coefficients from OLS fixed-effect regressions of the form $\tilde{T}_{i,f} = \beta * T_{i,f} + \nu_i + \epsilon_{i,f}$. The coefficient β can be interpreted as a scaling of the slope induced by the true tax schedule, where estimated values less than 1 indicate a "flattening" of the tax schedule. Two-sided Wald-test p-values, testing the hypothesis that $\beta = 1$, are presented below each regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D5: Testing for "Flattening" of the Tax Schedule: No Winsorization

	All Incomes	Income Quartiles			
	Pooled	1	2	3	4
Estimation Sample: Primary and High-Income Sampling Distributions					
Scale of slope (β)	0.66***	0.54***	0.58***	0.68***	0.82***
	(0.027)	(0.031)	(0.033)	(0.087)	(0.042)
P-value of H_0: $\beta = 1$	0.000	0.000	0.000	0.000	0.000
Respondents	4197	1050	1062	1040	1045
Forecasts	58758	14700	14868	14560	14630
Estimation Sample: Primary Sampling Distribution					
	(1)	(2)	(3)	(4)	(5)
Scale of slope (β)	0.88***	0.73***	1.08***	0.75***	0.94***
	(0.058)	(0.125)	(0.215)	(0.047)	(0.026)
P-value of H_0: $\beta = 1$	0.043	0.031	0.701	0.000	0.019
Respondents	4197	1050	1062	1040	1045
Forecasts	41970	10500	10620	10400	10450
Estimation Sample: Local Draws					
Scale of slope (β)	0.44	1.48***	1.29**	-0.28	0.60**
	(0.317)	(0.476)	(0.561)	(1.197)	(0.251)
P-value of H_0: $\beta = 1$	0.079	0.317	0.602	0.286	0.108
Respondents	4197	1050	1062	1040	1045
Forecasts	17937	3143	4074	5293	5427

Notes: Standard errors, clustered by respondent, in parentheses. Presented are coefficients from OLS fixed-effect regressions of the form $\tilde{T}_{i,f} = \beta * T_{i,f} + \nu_i + \epsilon_{i,f}$. The coefficient β can be interpreted as a scaling of the slope induced by the true tax schedule, where estimated values less than 1 indicate a "flattening" of the tax schedule. Two-sided Wald-test p-values, testing the hypothesis that $\beta = 1$, are presented below each regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D6: Parameter Estimates of Heuristic-Perception Model: Alt. Winsorizations

Winsorized at the 5th and 95th percentile forecast				
γ_I: weight on ironing forecast	0.30***	0.32***	0.56***	0.47***
	(0.027)	(0.041)	(0.036)	(0.072)
γ_S: weight on spotlighting forecast	-0.06*	0.00	-0.07	-0.04
	(0.033)	(0.039)	(0.047)	(0.057)
Winsorized at 0% and 100% of income				
γ_I: weight on ironing forecast	0.20***	0.29***	0.47***	0.44***
	(0.037)	(0.052)	(0.048)	(0.092)
γ_S: weight on spotlighting forecast	-0.08	-0.01	-0.03	-0.03
	(0.050)	(0.057)	(0.062)	(0.075)
No Winsorization				
γ_I: weight on ironing forecast	0.06	0.27	0.33***	0.58**
	(0.119)	(0.245)	(0.083)	(0.233)
γ_S: weight on spotlighting forecast	-0.15	-0.02	0.06	-0.07
	(0.146)	(0.214)	(0.098)	(0.145)
Residual misperception function included	No	Yes	No	Yes
High-income forecasts included	No	No	Yes	Yes
Respondents	4197	4197	4197	4197
Forecasts	41970	41970	58758	58758

Notes: Standard errors, clustered by respondent, in parentheses. This table reproduces the analysis of table 2 under alternative treatments of outliers. Estimated residual misperception functions are plotted in appendix table D1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Proofs

Proof of Proposition 1 Assume that $G(u) = u$; which is without loss of generality since monotonic transformations of utility functions preserve behaviors.

Part 1. Let $B_w(z')$ denote the optimal choice of z by an individual facing tax schedule $\tilde{T}(z|z')$. Note the following:

1. $B_w(z')$ is uniquely defined since $\tilde{T}(z|z')$ is convex. The convexity of $\tilde{T}(z|\cdot)$ follows by assumption A combined with the assumption that $T(z)$ is convex.

2. $B_w(z')$ is guaranteed to be continuous in z' because $\tilde{T}(z|z')$ is continuous in z' and $(c-\psi(z/w))-\tilde{T}(z|z')$ is strictly concave in z.

3. $B_w(z')$ is decreasing in z'. This is because assumption A and the assumption that $T(z)$ is convex guarantee that the perceived marginal tax rates are increasing in z'.

4. $B_w(0) > 0$, since the assumption that $\psi(0) = 0$ guarantees that the optimal choice of z is interior for any perceived tax schedule. Also, $B_w(z') < z'$ for large enough z' by the assumption that

$\lim_{z\to\infty}\psi'(z)=\infty$.

The above four facts show that $B_w(z')$ is a continuous and decreasing function with $B_w(0)>0$ and $B_w(z)<z$ for large enough z. A basic application of the intermediate value theorem establishes that this function must intersect the 45-degree line exactly once.

Part 2. Continuity follows immediately since $c-\psi(z/w)-\tilde{T}(z|z')$ is continuous w and is strictly concave in z. Next, note that for any $z_1>z_2$ and z', $\left(c-\psi(z_1/w)-\tilde{T}(z_1|z')\right)-\left(c-\psi(z_2/w)-\tilde{T}(z_2|z')\right)$ is strictly increasing in w because ψ is convex. The statement in the proposition is thus a standard comparative static on fixed points (see, e.g., Villas-Boas, 2016). Moreover, z^* will be strictly increasing in w if it is interior.

Part 3. Continuity follows immediately since $\tilde{T}$ is continuous in γ_I and γ_r, and since $c-\psi(z/w)-\tilde{T}(z|z')$ is strictly concave in z. Next, it follows immediately by assumption A that for $z_1>z_2$, $T(z_1|z')-T(z_2|z')$ is decreasing in γ_I and γ_r. Thus, for any $z_1>z_2$ and z', $\left(c-\psi(z_1/w)-\tilde{T}(z_1|z')\right)-\left(c-\psi(z_2/w)-\tilde{T}(z_2|z')\right)$ is strictly increasing in γ_I and γ_r. The result then follows as in Part 2. ■

An observation: It is useful to note that convexity of T plays two important roles in the proof of Proposition 1. First, it ensures that the individual's optimization problem is convex, and thus that B_w is single-valued. In particular, this then ensures that B_w has a closed graph, a property that would not hold for all possible T. Second, convexity of T ensures that B_w is a decreasing function. If T were concave, however, B_w would be an increasing function; and more generally, B_w could be increasing in some regions and decreasing in others for some tax schedules T. Existence and uniqueness are thus not guaranteed for all possible T. To ensure existence, the ME concept would need to be extended to allow for "mixed strategies."

Proof of Proposition 2 We denote by $\frac{d_+z^*}{d\gamma_I}:=\lim_{\epsilon\to 0}\frac{z^*(\gamma_I+|\epsilon|)-z^*(\gamma_I)}{|\epsilon|}$ the right derivative. We prove proposition 2 without assuming that T is not everywhere differentiable—that is, we allow T to have kink points. We only make the assumption that the distribution w of types is smooth, and thus that not all consumers are at kink points.

We show that $\frac{d}{d\gamma}_I W|_{\gamma_I=0,\gamma_r=0}>0$, from which the result follows. Let $\tilde{\tau}_w(z)<T'(z^*(w))$ denote the individual's perceived marginal tax rate. By definition, $1-\psi'/w=\tilde{\tau}_w$.

$$\begin{aligned} &= \int G'\cdot(1-\psi'/w-T')\frac{d_+z^*}{d\gamma_I}dF+\lambda\int\frac{d_+z^*}{d\gamma_I}T'dF \\ &= \int G'\cdot(\tilde{\tau}_w-T')dF+\lambda\int\frac{d_+z^*}{d\gamma_I}T'dF \end{aligned}$$

Now when $\gamma_I=\gamma_r=0$, $\tilde{\tau}_w=T'$ for all w; that is, consumers perceive tax rates correctly by definition. Thus $\frac{d}{d\gamma}_I W|_{\gamma_I=0,\gamma_r=0}=\lambda\int\frac{d_+z^*}{d\gamma_I}T'dF$. The result follows by part 3 of proposition 1, which shows that earnings are increasing in γ_I, and are strictly increasing for consumers who are not at kink points, and thus $\frac{d_+z^*}{d\gamma_I}>0$ for a positive measure of consumers.■

Proof of Proposition 3. More generally, we prove the results by replacing $\frac{dz^*}{dn}$ with $\frac{d_+z^*}{dn}$ in part 1, and replacing it with $\frac{d_-z^*}{dn}$ in part 2. To keep the proof concise we abuse notation slightly and use $\frac{dz^*}{dn}$ to denote the appropriate left or right derivative.

Analogously to the proof of Proposition 2, we can write

$$\begin{aligned}\frac{d}{dn}W \quad /\lambda = & \int g(z)\cdot(1-\psi'/w-T')\frac{dz^*}{dn}dF+\int\frac{dz^*}{dn}T'dF \\ = & \int\frac{dz^*}{dn}T'\left(1-g(z^*)\right)dF+\int g(z)\tilde{\tau}_w\frac{dz^*}{dn}dF\end{aligned} \tag{6}$$

Since $\tilde{\tau}_w \geq 0$, the term $\int g(z)\tilde{\tau}_w\frac{dz^*}{dn}$ in equation (6) positive when $\frac{dz^*}{dn}$ is positive and negative when $\frac{dz^*}{dn}$ is negative. Thus the term in equation (6) is less than (greater than) $\int\frac{dz^*}{dn}T'\left(1-g(z^*)\right)dF$ when $\frac{dz^*}{dn}$ is positive (negative). Moreover, note that

$$\int\frac{dz^*}{dn}T'\left(1-g(z^*)\right)dF \quad = - \quad Cov\left[\frac{dz^*}{dn}T', g(z^*)\right]+E\left[\frac{dz^*}{dn}T'\right]E\left[1-g(z^*)\right]$$

Now $Cov\left[\frac{dz^*}{dn}T', g(z^*)\right]$ is non-negative when $\frac{dz^*}{dn}$ is decreasing in w, and is non-positive when $\frac{dz^*}{dn}$ is increasing in w. Similarly, $E\left[\frac{dz^*}{dn}T'\right]E\left[1-g(z^*)\right]$ is non-positive (non-negative) when $E[g(z^*)]\leq 1$ and $\frac{dz^*}{dn}$ is non-positive (non-negative) for all w. This proves that when $\left|\frac{dz^*}{dn}\right|$ is increasing in w, $\int\frac{dz^*}{dn}T'\left(1-g(z^*)\right)dF$ negative (positive) if $\frac{dz^*}{dn}$ is negative (positive) for all w.■