



**HAL**  
open science

## **Inattention and the Taxation Bias**

Jérémy Boccanfuso, Antoine Ferey

► **To cite this version:**

Jérémy Boccanfuso, Antoine Ferey. Inattention and the Taxation Bias. Journal of the European Economic Association, 2023, 66 p. <10.1093/jeea/jvad056>. <hal-04603450>

**HAL Id: hal-04603450**

**<https://sciencespo.hal.science/hal-04603450v1>**

Submitted on 6 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

# Inattention and the Taxation Bias\*

Jérémy Boccanfuso<sup>†</sup>      Antoine Ferey<sup>‡</sup>

This version: April 2023

## Abstract

This paper shows that inattention to taxes generates a time-inconsistency problem in the choice of tax policy, leading to higher taxes in equilibrium. These discretionary tax increases are inefficient as they are deviations from the socially optimal commitment policy. We call these deviations a *taxation bias*. Combining sufficient statistics and structural approaches, we quantify the magnitude of this policy distortion for the U.S. redistributive tax-transfer system. We find that the taxation bias ranges between 3 and 8 percentage points, alters tax-transfer progressivity, and has significant welfare effects. Overall, our findings shed new light on the implications of inattention and misperceptions.

Keywords: Inattention; misperceptions; optimal taxation; policy distortions.

JEL classification code: D03; H21.

---

\*We are grateful to Pierre Boyer, Philippe Choné, Olivier Compte, Allan Drazen, Chishio Furukawa, Xavier Gabaix, Ulrich Glogowsky, Emanuel Hansen, Andreas Haufler, Laurence Jacquet, Etienne Lehmann, Benjamin Lockwood, Jean-Baptiste Michau, Xavier Ragot, Anasuya Raj, Daniel Reck, Alessandro Riboni, Dominik Sachs, Emmanuel Saez, Stefanie Stantcheva, Dmitry Taubinsky, Clémence Tricaud, Nicolas Werquin, and Mirko Wiederholt for their helpful and constructive comments in improving the paper. We also thank audiences at conferences (AFSE, EEA, GRASS, IIPF, LAGV, NTA) and seminars (CMR Cologne, CREST Paris, CRED Paris, ECARES Brussels, LMU Munich, UC Berkeley, U Bologna) for useful discussions and feedback. Boccanfuso gratefully acknowledges the financial support of Labex OSE (PSE) and Ferey that of Labex ECODEC (CREST), of the Deutsche Forschungsgemeinschaft through CRC TRR 190 (LMU Munich), and UC Berkeley for its hospitality. Brian Bloch provided editing services. This paper was awarded the *ITAX Best PhD Paper Award* at the 2019 IIPF Conference.

<sup>†</sup>University of Bologna, Piazza Scaravilli 2, 40126 Bologna. E-mail: jeremy.boccanfuso@unibo.it

<sup>‡</sup>LMU Munich, Akademiestr. 1/II, 80799 München. E-mail: antoine.ferey@econ.lmu.de

# 1 Introduction

A growing body of evidence reveals that individuals tend to be inattentive to opaque financial fees, add-on charges, and non-salient taxes. This inattention influences their behavior and thus a wide range of economic outcomes (see Gabaix, 2019, for a review). In the context of taxation, inattention reduces individual responsiveness to taxes (Chetty et al., 2009; Abeler and Jäger, 2015). Inattention is thus usually interpreted as reducing tax-induced distortions, thereby improving economic efficiency and warranting higher optimal tax rates (Goldin, 2015; Farhi and Gabaix, 2020).

However, recent evidence shows that, far from ignoring taxes, individuals rely on particular beliefs, heuristics, or rules of thumb when taxes are opaque (Rees-Jones and Taubinsky, 2020; Morrison and Taubinsky, 2021). This implies that tax perceptions, and thus behaviors, are shaped both by attention to taxes and individual priors. Hence, while inattention reduces short-run responses to taxes by anchoring perceptions on priors, long-run responses depend on how these priors adjust in equilibrium.

Building on this insight, this paper shows that the anchoring effect of inattention creates a time-inconsistency problem in the choice of tax policy (Kydland and Prescott, 1977). When the government cannot credibly commit to its tax policy, the equilibrium policy deviates from the socially optimal commitment policy. We call this deviation a *taxation bias* in analogy to the *inflation bias* (Barro and Gordon, 1983a). We thus show that inattention reduces economic efficiency by inducing inefficient tax increases.

We first derive this result in a stylized representative agent model in which a revenue-maximizing government levies a linear tax (Section 2). A higher tax rate reduces the tax base as it negatively affects individuals' decisions to, e.g., supply labor. Because of information frictions, these decisions depend on a perceived tax rate given by a weighted average between the actual tax rate and a prior, where the weight attached to the former measures attention to taxes. To build the intuition, we assume that the prior coincides with the actual tax rate in equilibrium. It implies that equilibrium perceptions are correct and that the optimal equilibrium policy is not affected by inattention: it follows a standard inverse elasticity rule (Dupuit, 1844; Ramsey, 1927). Nevertheless, we show that the equilibrium policy follows a modified inverse elasticity rule, where the elasticity is scaled by attention to taxes. In the presence of inattention, the equilibrium tax rate is

thus above the revenue-maximizing tax rate, on the wrong side of the Laffer curve.

The intuition behind this result is that inattention anchors tax perceptions on the prior. This creates an incentive for the government to implement policy deviations that individuals are going to partially ignore. However, because the prior adjusts in equilibrium, these policy deviations turn out to be inefficient *ex post*.<sup>1</sup> The anchoring effect of inattention thus makes the government’s choice of tax policy time-inconsistent. If the government could credibly commit to implementing the optimal equilibrium policy, it would commit to it and solve the time-inconsistency. Otherwise, discretionary policymaking leads to inefficient policy deviations. We define the taxation bias as the difference between the tax rates under discretion and commitment. In this stylized model, we derive a simple formula showing that the taxation bias is proportional to inattention.

Turning to a general framework, we derive necessary and sufficient conditions for the existence of a taxation bias (Section 3). We consider a welfare-maximizing government levying a linear income tax to finance a transfer, thereby redistributing resources across individuals with heterogeneous earnings abilities. Individuals’ labor supply depends on their tax perceptions captured through (ability-specific probabilistic) posteriors which can be any function of actual tax policy and their priors. This nests a wide range of perception and attention models, and allows for misperceptions in equilibrium – in which case inattention affects the optimal equilibrium policy and may warrant higher optimal tax rates (Goldin, 2015; Farhi and Gabaix, 2020). In this general framework, we find that discretionary policymaking leads to a taxation bias whenever (i) priors influence posteriors, and (ii) tax policy influences equilibrium priors.

Reviewing empirical evidence (Section 4), we first discuss empirical work on perceptions supporting conditions (i) and (ii), which implies that discretionary policymaking leads to a taxation bias. Second, we document discretionary policymaking in the context of income taxes, focusing on a tax provision (the Alternative Minimum Tax) that enabled U.S. policymakers to regularly implement non-salient discretionary tax increases in past decades. Third, we argue that Cabral and Hoxby (2012) and Finkelstein (2009) provide empirical evidence of a taxation bias, in property taxes and tolls respectively, by showing that reductions in tax salience are associated with inefficient tax increases.<sup>2</sup>

---

<sup>1</sup>We rely on a static model and therefore refer to an equilibrium adjustment. Online Appendix C shows that this static model can be interpreted as the steady state of a dynamic model with learning.

<sup>2</sup>We everywhere abstract from political economy constraints in the choice of tax policy. If tax policies

In light of this evidence, we gauge the magnitude of the taxation bias in the U.S. redistributive tax-transfer system (Section 5). To this end, we derive a sufficient statistics characterization showing that, in the absence of misperceptions, the size of the taxation bias only depends on the actual tax rate, average inattention, and the covariance between inattention and income. In the presence of misperceptions, the relevant sufficient statistics become hard to estimate empirically and we turn to structural models calibrated to the actual U.S. economy. More precisely, we consider three models that differ in the assumptions they place on inattention and misperceptions.<sup>3</sup> The sufficient statistics and structural estimates point to a taxation bias between 3 and 8 percentage points. Since the actual U.S. tax rate is close to 30%, this implies that between a tenth and a fourth of actual tax liabilities can be attributed to a taxation bias.

While inattention reduces welfare through the emergence of a taxation bias, it is also conducive to misperceptions, which reduce tax distortions and increase welfare when individuals underestimate marginal tax rates. In our calibrated structural models, the net welfare effect of inattention is positive. Nevertheless, in the absence of a taxation bias, the welfare gains from tax underestimation would be 20% to 45% higher.

Extending the analysis to nonlinear tax-transfer schedules, we show that the taxation bias also distorts tax-transfer progressivity (Section 6). Indeed, discretionary increases in the marginal tax rate at a given income depend on inattention at that income level. As a result, a negative correlation between inattention and income leads to larger increases in marginal tax rates at low income levels. Quantitatively, we find that this distortion in tax-transfer progressivity more than doubles the welfare impact of the taxation bias.

Last, we discuss the role of information asymmetries, tax complexity and competitive politics in the emergence of a taxation bias (Section 7). Discretionary policy deviations are hardly observable because of information asymmetries, which makes it difficult to hold the government to account and to enforce commitment policies. In this regard, tax simplification or competitive politics may decrease inattention to taxes, but they are unlikely to eliminate the time-inconsistency problem faced by the government. Hence, we are constrained to be inefficiently low and inattention relaxes these constraints, this may lead to efficient increases in tax rates. See Schenk (2011) and Matějka and Tabellini (2021) on political salience.

---

<sup>3</sup>We consider: (1) a model where inattention and misperceptions are exogenous and uniform ; (2) a rational inattention model where inattention and misperceptions endogenously decrease with income ; (3) an ironing model where individuals use their average tax rate as a prior for their marginal tax rate.

argue that existing institutions fail to offer a credible commitment technology and that solutions to the taxation bias would require major institutional changes.

This paper adds to the growing field of behavioral public economics (Bernheim and Taubinsky, 2018) by showing that inattention to taxes creates a time-inconsistency problem in the choice of tax policy, leading to policy distortions. Previous papers analyzing optimal tax policy with behavioral agents implicitly characterize tax policy under commitment, because they treat tax perceptions as a sufficient statistic without modeling the channels through which these perceptions adjust (Goldin, 2015; Gerritsen, 2016; Allcott et al., 2018; Farhi and Gabaix, 2020; Moore and Slemrod, 2021; Craig and Slemrod, 2022). Building on recent empirical evidence (Gabaix, 2019; Rees-Jones and Taubinsky, 2020; Morrison and Taubinsky, 2021), we distinguish between attention to taxes and priors about taxes, in which case the anchoring effect of inattention gives rise to a taxation bias. Focusing on sales taxes, Furukawa (2022) studies in a Bayesian learning model the emergence of a taxation bias through a deadweight loss analysis and analyzes in this setting the value of tax transparency. Taubinsky and Rees-Jones (2018) also show that inattention reduces economic efficiency focusing on a different mechanism, which is that inattention increases heterogeneity in tax perceptions, and they estimate the associated efficiency losses for U.S. sales taxes in an online experiment.

This paper also relates to the fiscal illusion literature whose origins can be traced back to Mill (1848). Its modern roots lie in the public choice literature which argues that an inefficient and self-aggrandizing government will take advantage of a less visible tax system to grow larger than would be socially desirable (Buchanan, 1967; Wagner, 1976; Dollery and Worthington, 1996). We add to this literature by showing that inattention to taxes creates a taxation bias by itself, implying that a government that is neither inefficient nor self-aggrandizing grows beyond the optimum in the presence of inattention. This is a novel and broader microfoundation for fiscal illusion which provides a new rationale for some of the recent findings of Finkelstein (2009) and Cabral and Hoxby (2012). In this regard, the taxation bias is consistent with people disliking and distrusting complex tax systems (Blesse et al., 2019) as they lead to higher inefficient tax increases by fostering inattention.

Following the contribution of Kydland and Prescott (1977), the discretionary nature

of policymaking has mostly been analyzed in the context of fiscal and monetary policies (Barro and Gordon, 1983a; Billi, 2011; Halac and Yared, 2014; Dovis and Kirpalani, 2020), and capital taxation (Eichengreen, 1989; Farhi et al., 2012; Scheuer and Wolitzky, 2016; Saez and Stantcheva, 2018). Our analysis highlights that the anchoring effect of inattention provides a natural justification for the timing assumption that is central to this literature. Hence, the insight that inattention leads to time-inconsistency issues extends to virtually any setting where inattention and information frictions have been documented (Coibion and Gorodnichenko, 2015; Gabaix, 2019) and where the framework developed in this paper could be fruitfully applied.

The rest of the paper is organized as follows. Section 2 introduces the taxation bias in a stylized model. Section 3 extends the analysis to a general framework. Section 4 reviews empirical evidence in support of the taxation bias. Section 5 provides a sufficient statistics characterization as well as empirical estimates for the U.S. tax-transfer system. Section 6 provides an extension to nonlinear tax-transfer schedules. Section 7 discusses the role of information asymmetries, tax complexity and competitive politics. The final section concludes. All proofs are presented in the Appendix.

## 2 Taxation bias in a stylized model

**Setting.** We consider a revenue-maximizing government setting a linear income tax rate  $\tau$  on a mass of homogeneous individuals who make labor supply decisions. Because of information frictions, taxpayers are unable to perfectly observe the tax rate and thus have to form an estimate of this tax rate to decide how much to work. Call this estimate the perceived tax rate,  $\tilde{\tau}$ , and suppose it is determined by a convex combination of a prior,  $\hat{\tau}$ , and the actual tax rate,  $\tau$ ,

$$\tilde{\tau} = \xi\tau + (1 - \xi)\hat{\tau}. \tag{1}$$

The weight  $\xi \in [0, 1]$  is a measure of attention to taxes: when  $\xi = 1$ , individuals are fully attentive and perfectly observe the actual tax rate; when  $\xi = 0$ , individuals are fully inattentive and completely anchor their perceptions on the prior.<sup>4</sup> Earnings choices

---

<sup>4</sup>Gabaix (2019) argues this is a reduced-form framework nesting many biases and attention theories.

negatively depend on the perceived tax such that aggregate earnings,  $Y(1 - \tilde{\tau})$ , increase with the perceived net-of-tax rate. We assume that the elasticity of earnings with respect to the perceived net-of-tax rate,  $e$ , is constant.<sup>5</sup>

**Tax policy.** With information frictions, the government must account for individuals' prior and attention to taxes when setting its tax policy. Formally, it chooses the tax rate which maximizes tax revenue,  $\tau Y(1 - \tilde{\tau})$ , subject to the tax perception model (1). As shown in Appendix B.1, this is a concave problem leading to the policy function,

$$\tau(\xi, \hat{\tau}) = \begin{cases} \frac{1-(1-\xi)\hat{\tau}}{\xi(1+e)} & \text{if } \hat{\tau} \geq 1 - \frac{\xi}{1-\xi}e \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

which is decreasing in attention,  $\xi$ , and prior,  $\hat{\tau}$ . The solution is interior when attention and prior are high enough, otherwise the government chooses a 100% tax.

When taxpayers are fully attentive,  $\xi = 1$ , the government thus chooses a tax policy that follows an inverse elasticity rule:

$$\tau^r = \frac{1}{1+e}. \quad (3)$$

Absent inattention, the perceived tax rate is always equal to the actual tax rate. Hence, the revenue-maximizing tax rate,  $\tau^r$ , is such that a small increase in the tax rate,  $d\tau$ , induces a positive mechanical effect,  $M^r = Y(1 - \tau^r) d\tau$ , that exactly balances out the negative fiscal externality,  $F^r = \tau^r \frac{Y(1-\tau^r)}{1-\tau^r} e d\tau$ , from reductions in earnings.

To understand why inattention leads to deviations from the inverse elasticity rule, suppose taxpayers' prior is that the government implements it,  $\hat{\tau} = \tau^r$ . Then, starting from this policy,  $\tau = \tau^r$ , any increase in the tax rate generates the same positive mechanical effect,  $M = M^r$ , but a smaller negative fiscal externality,  $F = \xi F^r$ . Indeed, reductions in earnings now depend on the change in the *perceived* tax rate, equal to  $d\tilde{\tau} = \xi d\tau$ , since inattentive taxpayers only observe a fraction  $\xi$  of the tax increase. That is, inattention reduces individuals' responses to tax changes by anchoring perceptions on prior beliefs. As a result, the mechanical effect outweighs the fiscal externality and the government systematically implements a higher tax rate.

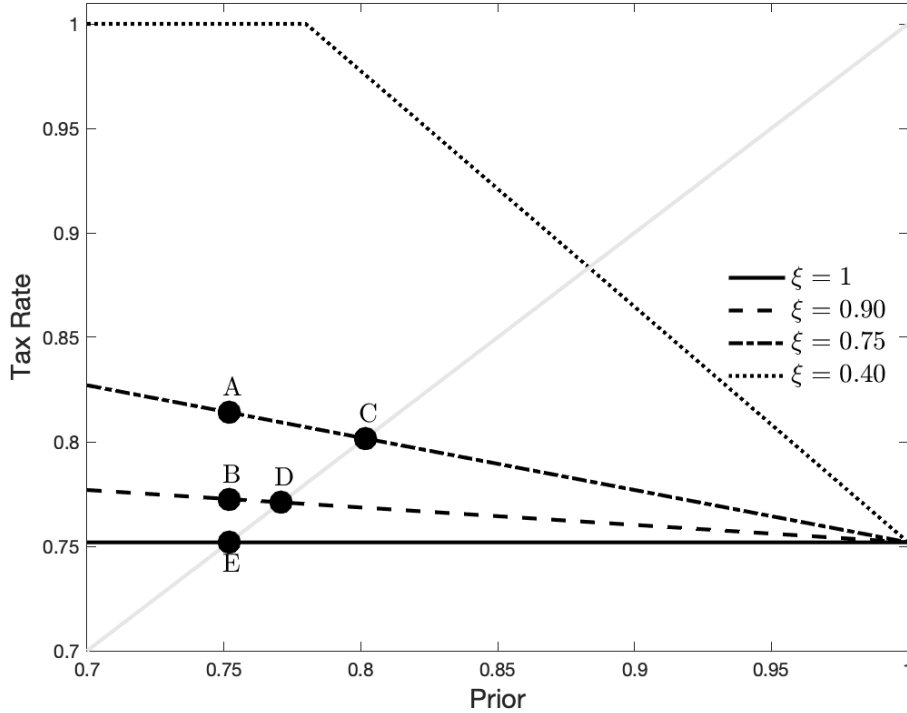
Figure 1 plots tax policy,  $\tau(\xi, \hat{\tau})$ , assuming that  $e = 0.33$ . If the prior is that the government implements the inverse elasticity rule  $\tau^r = 75\%$ , the government chooses a

---

<sup>5</sup>This elasticity corresponds to the structural elasticity parameter that reflects individuals' preferences, it is often referred to as the frictionless elasticity (Chetty, 2012; Kostøl and Myhre, 2021).

tax rate of 82% when the attention parameter is  $\xi = 0.75$  (point A), and a tax rate of 78% when the attention parameter is  $\xi = 0.90$  (point B).<sup>6</sup>

Figure 1: Policy functions and equilibrium in the stylized model



NOTE: Black lines represent policy functions,  $\tau(\xi, \hat{\tau})$  for different attention parameters,  $\xi$ . The grey line represents the (rational) equilibrium condition,  $\hat{\tau} = \tau^{\text{eq}}$ . With a frictionless elasticity  $e = 0.33$  (Chetty, 2012), the revenue maximizing tax rate is 75% (point E). Points A and B represent the government's policy deviations when inattentive agents expect the government to implement this revenue maximizing tax rate, and points C and D are the corresponding equilibrium tax rates.

**Equilibrium.** In equilibrium, (i) neither taxpayers nor the government has an incentive to deviate, and (ii) taxpayers' actions and perceptions are mutually consistent with the government choice of tax policy.<sup>7</sup> To build the intuition, we focus in this stylized model on *rational* equilibria in which taxpayers' prior is correct,  $\hat{\tau} \stackrel{\text{eq}}{=} \tau$ . Plugging this equilibrium condition into the policy function (2), the equilibrium tax policy follows a modified inverse elasticity rule, where the elasticity is scaled by attention:

$$\tau^{\text{eq}} = \frac{1}{1 + \xi e}. \quad (4)$$

<sup>6</sup>Gabaix (2019) states (p. 5) that "on average, the attention parameter estimated in the literature is 0.44, roughly halfway between no attention and full attention", while adding that "attention is higher when the incentives to pay attention are stronger", which should be the case for a 75% tax on income.

<sup>7</sup>Whenever relevant, we use the symbol  $\stackrel{\text{eq}}{=}$  to emphasize that an equality holds in equilibrium.

Graphically, the equilibrium is such that the 45-degree line representing the equilibrium condition,  $\hat{\tau} \stackrel{\text{eq}}{=} \tau^{\text{eq}}$ , crosses the policy function (2). In a rational equilibrium, the tax rate is equal to 80% when the attention parameter is  $\xi = 75\%$  (point C), and is equal to 77% when the attention parameter is  $\xi = 90\%$  (point D).

**Taxation bias.** In a rational equilibrium, inattention has no effect on the equilibrium perceived tax rate, which is always equal to the actual tax rate,  $\tilde{\tau} \stackrel{\text{eq}}{=} \tau$ . The equilibrium Laffer curve is thus also unaffected by inattention: the equilibrium tax revenue is simply given by  $\tau Y(1 - \tau)$ . Hence, the equilibrium revenue-maximizing tax rate follows the standard inverse elasticity rule,  $\tau^* = \tau^r$  (point E). As this would be the policy implemented by a social planner who can credibly commit to its tax policy, we refer to  $\tau^*$  as the tax policy under commitment. By definition, it is the optimal equilibrium policy.

However, in the presence of inattention the government systematically deviates from this optimal policy,  $\tau^*$ , and implements policy,  $\tau^{\text{eq}}$ , given by a modified inverse elasticity rule. This is because the optimal policy is time-inconsistent: although the government would prefer to implement this policy ex ante, it will systematically deviate ex post. This time-inconsistency problem arises because of the anchoring effect of inattention and restricts the set of implementable policies in equilibrium.<sup>8</sup> As a result, the equilibrium tax rate is inefficiently high and on the wrong side of the Laffer curve.

We call this phenomenon the *taxation bias* in analogy to the *inflation bias* (Barro and Gordon, 1983a), because in both situations, policy deviations reflects the discretionary nature of policymaking. Hence, inattention resurges the possibility that discretionary policies lead to inefficient outcomes (Kydlan and Prescott, 1977). Defining the taxation bias as the difference between the tax rate under discretion and the tax rate under commitment, we have

$$\frac{\tau^{\text{eq}} - \tau^*}{\tau^*} = (1 - \tau^{\text{eq}}) \frac{1 - \xi}{\xi} . \quad (5)$$

The taxation bias is thus strictly positive when individuals are not perfectly attentive to taxes,  $\xi < 1$ , and it increases with inattention. Other determinants of tax policy, such as the frictionless elasticity, are implicitly captured in the equilibrium tax rate,  $\tau^{\text{eq}}$ .

---

<sup>8</sup>In Barro and Gordon's (1983a) words, "*the equality of policy expectations and realizations is a characteristic of equilibrium – not a prior constraint*" (p. 591).

**Discussion.** Despite its simplicity, this stylized model goes a long way in capturing the essence of the taxation bias.

As we show in the Online Appendix, it can be easily extended to dynamic settings. These extensions reveal that the equilibrium of the static model coincides with the steady state of a dynamic model in which taxpayers learn tax rates over time.<sup>9</sup> This type of dynamic learning provides an intuitive justification for taking taxpayers' prior as given when setting tax policy under discretion: taxpayers' prior depends on the history of tax policies and is thus independent of current tax policy.

Furthermore, we show that long-run sophistication of the government – e.g. that it internalizes taxpayers' learning – still leads to inefficiently high tax rates in the steady state, unless the government is infinitely patient. This indicates that imposing further constraints on the government is unlikely to eliminate the taxation bias, although this could undoubtedly reduce its magnitude. We discuss these issues in Section 7.

### 3 Taxation bias in a general model

This section introduces a broader framework that features a welfarist government, heterogeneous individuals, and a general tax perception model allowing for endogenous attention and arbitrary misperceptions. When tax policy is discretionary, we identify necessary and sufficient conditions for the emergence of a taxation bias, which essentially hinge on the way individuals form their tax perceptions.

#### 3.1 Setup

Consider a population of heterogeneous individuals indexed by their earnings ability,  $w$ , which are private information and distributed from a probability density function,  $f_w(w)$ . The government levies a linear income tax financing a lump-sum transfer. Because of information frictions, individuals are unable to freely observe the tax-transfer system and thus treat the (marginal) tax rate,  $\tau$ , and the transfer,  $R$ , as random variables. We

---

<sup>9</sup>It would therefore be wrong to interpret the static model as one in which the government perpetually tries to fool taxpayers (see the Online Appendix for a discussion). Similarly, the equilibrium condition that the prior adjusts to actual tax policy should not be interpreted as an assumption about the sophistication of taxpayers: taxpayers need not anticipate the government's incentive to deviate.

assume individual preferences are quasi-linear in consumption and additively separable with respect to the disutility from labor supply,  $U(c, y; w) = c - k(y; w)$ , implying that there are no income effects in labor supply decisions. As a result, the perceived tax rate,  $\tilde{\tau}$ , is a sufficient statistic for earnings decisions,  $y(\tilde{\tau}; w)$ .<sup>10</sup> Given the actual tax rate,  $\tau_0$ , and the actual lump-sum transfer,  $R_0$ , we assume that consumption costlessly adjusts such that individuals' budget constraints are binding,  $c(\tilde{\tau}, \tau_0, R_0; w) = (1 - \tau_0)y(\tilde{\tau}; w) + R_0$ .<sup>11</sup> The associated indirect utility is denoted by  $\mathcal{V}(\tilde{\tau}, \tau_0, R_0; w)$ .

We remain as agnostic as possible on the way individuals form their tax perceptions. We assume the existence of a probability density function,  $\hat{q}(\tau; w)$ , capturing priors about the tax rate (or beliefs) and that may depend on earnings ability. We also assume the existence of a probability density function,  $f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)$ , capturing posteriors about the tax rate (or perceived tax rates) and that may depend on the actual tax rate, on priors, and on earnings ability. Importantly, we do not impose any assumption about the relationship between priors and posteriors.

The government sets its tax policy to maximize a social welfare function summing an increasing concave transformation,  $G(\cdot)$ , of taxpayers' indirect utilities,  $\mathcal{V}(\cdot)$ , subject to a resource constraint. More precisely, it chooses a target tax-transfer schedule,  $(\tau_g, R_g)$ , implemented up to an implementation shock,  $\vartheta$ , which is a white noise drawn from an exogenous distribution,  $f_{\vartheta}(\vartheta)$ , that is common knowledge.<sup>12</sup> Given a realization of the implementation shock, the actual tax rate is  $\tau_0 = \tau_g + \vartheta$  and the actual demogrant,  $R_0$ , adjusts such that the government's resource constraint is binding.

With an exogenous expenditure requirement,  $Exp$ , and denoting  $E_{\vartheta}[\cdot]$  the expectation over the implementation shock,  $\vartheta$ , the government's problem is

$$\max_{\tau_g, R_g} E_{\vartheta} \left[ \iint G\left(\mathcal{V}(\tau, \tau_0, R_0; \kappa, w)\right) f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w) f_w(w) d\tau dw \right] \quad (6)$$

$$\text{s.t.} \quad \iint \tau_0 y(\tau; w) f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w) f_w(w) d\tau dw = R_0 + Exp. \quad (7)$$

---

<sup>10</sup>We show in Online Appendix G how the main results extend with income effects, although earnings decisions then depend on both the perceived tax rate,  $\tilde{\tau}$ , and the perceived transfer,  $\tilde{R}$ .

<sup>11</sup>See Reck (2016) or Farhi and Gabaix (2020) for discussions of costly budget adjustment rules with misperceptions. They have intuitively little impact on the taxation bias as it mostly relates to inattention.

<sup>12</sup>Implementation shocks are introduced as an underlying source of uncertainty to ensure that, e.g., rationally inattentive agents have an incentive to learn in equilibrium as in Matějka and Tabellini (2021). By creating a source of randomness, they allow us to close the model when perceptions are stochastic.

### 3.2 Tax policy under discretion

When setting its policy, the government cannot credibly commit to a predefined policy because of the time-inconsistency problem discussed in Section 2. As a result, the government cannot directly change individuals' priors about the tax rate (e.g., with a public announcement) and must take these priors as given when choosing its target policy.

While the problem under consideration is fundamentally simultaneous, it can be equivalently described by the following sequence of events:

0. Taxpayers are endowed with priors  $\hat{q}(\tau; w)$ .
1. The discretionary government sets its target policy  $(\tau_g, R_g)$  solving problem (6)-(7).
2. The implementation shock  $\vartheta$  is drawn and the actual tax rate is  $\tau_0 = \tau_g + \vartheta$ .
3. Taxpayers' posteriors  $f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)$  adjust and they make their earnings decisions.
4. The government levies taxes and gives the actual transfer  $R_0$  satisfying (7).

The government understands that individuals are partially attentive to actual policy and adjust their decisions in reaction to its policy choice. It thus 'plays first' in the above-described sequence of events. However, because the government treats priors,  $\hat{q}(\tau; w)$ , as predetermined state variables, it does not have a strategic advantage from 'playing first' which reflects the simultaneous nature of the problem. Importantly, the government is as rational and informed as in a standard optimal tax model and the novelty relates to the presence of information frictions in individuals' tax perceptions.

**Proposition 1.** *Given priors,  $\hat{q}(\tau; w)$ , the discretionary policy,  $(\tau_g, R_g)$ , follows from*

$$\begin{aligned}
 (\tau_g) : E_{\vartheta} \left[ \int \left\{ \underbrace{\int \left[ y - \frac{G'(\mathcal{V})}{p} y \right] f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w) d\tau}_{\text{mechanical effects}} \right. \right. \\
 \left. \left. + \underbrace{\int \left[ \tau_0 y + \frac{G(\mathcal{V})}{p} \right] \frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}} d\tau}_{\text{direct behavioral effects}} \right\} f_w(w) dw \right] = 0
 \end{aligned} \tag{8}$$

$$(R_g) : E_{\vartheta} \left[ \iint \left[ \frac{G'(\mathcal{V})}{p} - 1 \right] f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w) f_w(w) d\tau dw \right] = 0 \tag{9}$$

together with the resource constraint (7), where  $p$  is the marginal value of public funds.

Condition (8) characterizes the target tax rate,  $\tau_g$ . The first line relates to mechanical fiscal and welfare effects of changes in the tax rate, holding the distribution of perceived tax rates and thus earnings constant. More precisely, a  $d\tau_g$  increase in the tax rate increases tax revenue by  $y d\tau_g$  additional dollars, and reduces consumption and thus welfare by  $\frac{G'(\mathcal{V})}{p} y d\tau_g$  dollars.

The second line relates to direct behavioral effects associated with earnings responses. These responses are captured by changes in the distribution of perceived tax rates holding priors constant,  $\frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}}$ , which depend on attention to tax changes. In the stylized model of Section 2, this is saying that a  $d\tau_g$  increase in the tax rate increases the perceived tax rate by  $d\tilde{\tau} = \xi d\tau_g$ . This shifts the earnings of an individual with ability  $w$  from  $y(\tilde{\tau}; w)$  to  $y(\tilde{\tau} + d\tilde{\tau}; w)$ , which corresponds to a  $\frac{dy(\tilde{\tau}; w)}{d\tilde{\tau}} d\tilde{\tau}$  change in earnings that affects tax revenue,  $\tau_0 y$ , and potentially welfare,  $\frac{G(\mathcal{V})}{p}$ .<sup>13</sup> Indeed, in the presence of misperceptions, choices are not utility-maximizing and welfare changes do not envelope out, introducing a corrective motive for taxation as in, e.g., Gerritsen (2016) and Farhi and Gabaix (2020).

Condition (9) shows that a dollar increase in transfer increases welfare by  $\frac{G'(\mathcal{V})}{p}$  dollars but costs a dollar to the government. Social marginal welfare weights,  $\frac{G'(\mathcal{V})}{p}$ , thus average to one at the optimum: the government is indifferent between having an additional dollar or redistributing an additional dollar (Saez, 2001). This characterizes the marginal value of public funds,  $p$ , and the transfer follows from the resource constraint (7).

**Equilibrium.** Given the structure of the problem, the only free variables are the target tax rate,  $\tau_g$ , posterior distributions,  $f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)$ , and prior distributions,  $\hat{q}(\tau; w)$ . We formalize the equilibrium concept in Definition 1.

**Definition 1** (equilibrium). *Given a distribution of implementation shocks,  $f_{\Theta}(\vartheta)$ , an equilibrium is a set of target tax rate chosen by the government and attention strategies chosen by taxpayers such that*

- (a) *Given priors  $\hat{q}(\tau; w)$ , the target tax rate  $\tau_g$  solves the government's problem.*
- (b) *Given priors  $\hat{q}(\tau; w)$ , posteriors  $f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)$  follow from attention strategies.*
- (c) *Taxpayers' priors adjust to the target policy of the government,  $\hat{q}(\tau; w) \stackrel{\text{eq}}{=} \Gamma(\tau, \tau_g; w)$ .*

This definition of equilibrium is general enough to encompass many different models of inattention and tax perceptions. Conditions (a) and (b) ensure that neither the government nor taxpayers have an incentive to deviate, while condition (c) ensures that taxpayers' priors and the government's tax policy are mutually consistent in equilibrium.

<sup>13</sup>Normalizing implementation shocks to zero such that  $\tau_0 = \tau_g$ , the stylized model of Section 2 implies  $f_{\tilde{\tau}}(\tau|\tau_g, \hat{q}; w) = \delta(\tau - \tilde{\tau})$  with  $\delta(\cdot)$  the Dirac distribution and  $\tilde{\tau} = \xi\tau_g - (1 - \xi)\hat{\tau}$ . Integrating by parts,

$$\int y(\tau; w) \frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}} d\tau = \int y(\tau; w) \frac{d\delta(\tau - \tilde{\tau})}{d\tau_g} \Big|_{\tilde{\tau}} d\tau = -\xi \int y(\tau; w) \frac{d\delta(\tau - \tilde{\tau})}{d\tau} d\tau = \xi \int \frac{dy(\tau; w)}{d\tau} \delta(\tau - \tilde{\tau}) d\tau = \xi \frac{dy(\tilde{\tau}; w)}{d\tilde{\tau}}.$$

Indeed, the condition  $\hat{q}(\tau; w) \stackrel{\text{eq}}{=} \Gamma(\tau, \tau_g; w)$  implies that priors generally depend on the target tax rate,  $\tau_g$ , in equilibrium. For instance, in the stylized model of Section 2, the prior exactly coincides with the target tax rate in equilibrium, meaning that  $\Gamma(\tau, \tau_g; w) = \delta(\tau - \tau_g)$  where  $\delta(\cdot)$  is the Dirac distribution. Alternatively, misperceptions arise when  $\Gamma(\tau, \tau_g; w)$  is not centered around the target tax rate in equilibrium. We remain agnostic about  $\Gamma(\tau, \tau_g; w)$  at this stage and consider different specifications in Section 5.

Given Definition 1, the equilibrium discretionary tax policy,  $(\tau_g^{\text{eq}}, R_g^{\text{eq}})$ , follows from Proposition 1 with priors evaluated at their equilibrium values,  $\hat{q}(\tau; w) \stackrel{\text{eq}}{=} \Gamma(\tau, \tau_g^{\text{eq}}; w)$ .

### 3.3 Commitment and the taxation bias

The equilibrium discretionary tax policy is in general suboptimal. To formalize this point, we characterize the policy that maximizes welfare in equilibrium. This is the policy that would be chosen by a government who could credibly commit to implementing a given policy, we thus refer to it as the policy under commitment.

**Commitment policy.** When a government can credibly commit to a predefined policy, tax policy solves the government's problem (6)-(7) subject to the additional condition that taxpayers' priors and the government's choice of tax policy are mutually consistent in equilibrium, i.e., condition (c) in Definition 1.

**Proposition 2.** *Given equilibrium priors,  $\hat{q}(\tau; w) \stackrel{\text{eq}}{=} \Gamma(\tau, \tau_g; w)$ , the commitment tax policy,  $(\tau_g^*, R_g^*)$ , follows from*

$$(\tau_g^*): E_{\vartheta} \left[ \underbrace{\int \left\{ \int \left[ y - \frac{G'(\mathcal{V})}{p} y \right] f_{\tilde{\tau}}(\tau | \tau_0, \hat{q}; w) d\tau}_{\text{mechanical effects}} \right.}_{\text{direct \& equilibrium behavioral effects}} + \int \left[ \tau_0 y + \frac{G(\mathcal{V})}{p} \right] \left[ \frac{df_{\tilde{\tau}}(\tau | \tau_0, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}} + \frac{df_{\tilde{\tau}}(\tau | \tau_0, \hat{q}; w)}{d\hat{q}} \frac{d\Gamma(\tau, \tau_g; w)}{d\tau_g} \right] d\tau \Big]_{\hat{q} \stackrel{\text{eq}}{=} \Gamma} f_w(w) dw \Big] = 0 \quad (10)$$

$$(R_g^*): E_{\vartheta} \left[ \iint \left\{ \left[ \frac{G'(\mathcal{V})}{p} - 1 \right] f_{\tilde{\tau}}(\tau | \tau_0, \hat{q}; w) f_w(w) d\tau \right\}_{\hat{q} \stackrel{\text{eq}}{=} \Gamma} dw \right] = 0 \quad (11)$$

together with the resource constraint (7) and  $p$  is the marginal value of public funds. This is the equilibrium welfare-maximizing tax policy.

Proposition 2 characterizes the commitment tax policy. The main difference with Proposition 1 characterizing the discretionary tax policy is that, in the second line of (10),

a change in the tax rate changes taxpayers' perceived tax rates not only through direct effects,  $\left. \frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\tau_g} \right|_{\hat{q}}$ , but also through equilibrium effects,  $\frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\hat{q}} \frac{d\Gamma(\tau, \tau_g; w)}{d\tau_g}$ , related to the equilibrium adjustments of taxpayers' priors. In the stylized model of Section 2, this is saying that a  $d\tau_g$  increase in the tax rate increases the equilibrium perceived tax rate by  $d\tilde{\tau} = \xi d\tau_g + (1 - \xi)d\tau_g = d\tau_g$ . As equilibrium adjustments increase responses to tax reforms, the commitment tax rate is lower than the discretionary tax rate.

Implicitly, this is the policy analyzed in optimal tax papers with behavioral agents. Indeed, the appeal of the sufficient statistics approach is to characterize optimal policy in terms of sufficient statistics evaluated at the optimum, without modeling where sufficient statistics come from and how they adjust. For instance, Gerritsen (2016) and Farhi and Gabaix (2020) characterize welfare-maximizing taxes accounting for potential misperceptions (or other behavioral biases) through a general *behavioral wedge*. Since they characterize welfare-maximizing policies accounting for all adjustments in sufficient statistics, they characterize commitment policies.

**Taxation bias.** The taxation bias is defined as the deviation from the equilibrium welfare-maximizing tax rate,  $\tau_g^*$ .

**Definition 2** (taxation bias). *The taxation bias is the difference between the equilibrium tax rates under discretion,  $\tau_g^{\text{eq}}$ , and under commitment,  $\tau_g^*$ .*

Inattention creates a time-inconsistency problem in policymaking linked to the equilibrium adjustments of perceptions. As a result, when the government cannot credibly commit to a predefined tax policy, equilibrium responses to tax reforms lead to a taxation bias: the equilibrium policy does not coincide with the optimal equilibrium policy.

**Proposition 3.** *When both equilibria exist and are unique, discretionary policymaking leads to a taxation bias if and only if*

$$E_{\vartheta} \left[ \underbrace{\int \left\{ \int \left[ \tau_0 y + \frac{G(\mathcal{V})}{p} \right] \overbrace{\frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\hat{q}} \frac{d\Gamma(\tau, \tau_g; w)}{d\tau_g}}^{(i)} d\tau \right\}}_{\text{equilibrium behavioral effects}} \right]_{\hat{q} \stackrel{\text{eq}}{\Gamma}} f_w(w) dw \neq 0 \quad (12)$$

Proposition 3 shows that discretionary policymaking leads to a taxation bias when equilibrium responses to tax reforms are non-zero. Indeed, the left-hand side of (12) represents expected equilibrium behavioral effects upon a change in the tax rate. It

further decomposes equilibrium responses into two channels, implying that a taxation bias emerges when (i) posterior perceptions are affected by prior beliefs, and (ii) prior beliefs are affected by the government choice of tax policy in equilibrium.

## 4 Review of empirical evidence

This section reviews empirical evidence supporting the existence of a taxation bias. First, we present evidence for conditions (i) and (ii), implying that discretionary policymaking leads to a taxation bias (Proposition 3). Second, we document evidence of discretionary policymaking in income tax policy, suggesting the existence of a taxation bias in the U.S. redistributive tax-transfer system. Third, we argue that Cabral and Hoxby (2012) and Finkelstein (2009) present direct evidence of a taxation bias in property taxes and tolls.

### 4.1 Perceptions and prior beliefs

There is substantial evidence reviewed in Bernheim and Taubinsky (2018) and in Malmendier (2021) that perceptions are often anchored and often depend on past experience, implying that (i) perceptions are affected by prior beliefs.<sup>14</sup> Since prior beliefs are generally hard to elicit, it is difficult to provide direct evidence that (ii) prior beliefs are affected by tax policies. It is however possible to provide joint evidence for both (i) and (ii) by identifying variations in tax perceptions correlated with past tax policies but independent of actual tax policy. For instance, if taxpayers who experienced higher taxes in the past report higher perceived taxes than otherwise identical taxpayers facing the same tax schedule today, it suggests that their perceptions are influenced by past tax policy through their beliefs.

This is exactly what Goldfayn-Frank and Wohlfart (2020) document in the context of inflation expectations. They show that East Germans, who previously experienced higher inflation, continue to have higher inflation expectations than West Germans, decades after the reunification. In the context of taxation, Fochmann and Weimann (2013) find that individuals rely on their actual taxes to form their perceptions in a lab experiment,

---

<sup>14</sup>The effect of priors on posteriors is also visible from the sluggish adjustment of posteriors (Coibion and Gorodnichenko, 2012, 2015) consistent with the many information provision experiments showing that inattentive individuals *partially* adjust their expectations to new information (Haaland et al., 2020).

although the experimental tax schedule differs from the actual tax schedule. Similar anchoring effects are found in empirical papers studying tax reforms and reporting bunching at old and no longer relevant thresholds. While these bunching behaviors may be driven by frictions in labor supply adjustments, Aghion et al. (2017) argue that in the setting they study they are likely driven by information frictions and sluggish learning.

Overall, this strongly suggests that conditions (i) and (ii) hold in practice, implying that discretionary policymaking will lead to a taxation bias (Proposition 3).

## 4.2 Discretionary tax increases

There is a large literature discussing discretionary policymaking in monetary policy (e.g., Barro and Gordon, 1983a), fiscal policy (e.g., Taylor, 2000), and wealth taxation (e.g., Eichengreen, 1989). Looking at income taxes, tax reforms are often discretionary and there is evidence that policymakers regularly mask discretionary tax increases through low-salience provisions. Schenk (2011) reviews some of this evidence for the U.S., including the prominent example of the Alternative Minimum Tax (AMT).

The AMT was introduced during the 1970s to ensure that all high-income taxpayers, including those with significant tax deductions, would pay a minimum amount of federal income taxes. Initially a class tax affecting 200,000 taxpayers in 1982, the AMT gradually became a mass tax affecting 5.1 million taxpayers in 2017 and raising 2.4% of federal income tax revenue (Tax Policy Center, 2023). A major driver of this growth was the lack of automatic indexation of the AMT on inflation, leaving any adjustment to the discretion of the Congress, and contrasting with the automatic indexation of the statutory federal income tax schedule. As discretionary adjustments did not compensate the automatic indexation, more and more taxpayers became subject to the AMT (Harvey and Tempalski, 1997; Feenberg and Poterba, 2004; Burman et al., 2005).<sup>15</sup> This highlights a tension between the salient commitment rule of automatic indexation in place since 1981 for the statutory schedule, and the non-salient discretionary tax increases induced by the non-systematic adjustment of the AMT after 1982. Eventually, the AMT became automatically indexed on inflation in 2013 and, following a 2018 reform, it now affects

---

<sup>15</sup>Another driver of this growth are the 2001 and 2003 tax cuts reducing statutory tax rates without changing the AMT rates. Interestingly, it was well anticipated that these largely advertised tax cuts would be effectively limited by the presence of the less publicized AMT (Gale and Potter, 2002).

about 200,000 taxpayers, as it did initially.

This evidence of non-salient discretionary tax increases suggests that all theoretical conditions for the existence of a taxation bias in the U.S. redistributive tax-transfer system are met. However, we cannot rule out the possibility that policymakers exploit cognitive biases in a welfare-enhancing way (Schenk, 2011; Goldin, 2015; Moore and Slemrod, 2021; Craig and Slemrod, 2022) since the socially optimal commitment policy depends on equilibrium tax perceptions (Proposition 2). To the best of our knowledge, more direct evidence of inefficient tax increases due to inattention only exists for property taxes and tolls, which we now turn to.

### 4.3 Toward evidence of a taxation bias

Providing direct evidence of discretionary tax increases that are inefficient is challenging, because it requires to identify the counterfactual policy under commitment. A key insight from our theory is that discretionary tax policy depends on direct responses to tax reforms (Proposition 1), whereas commitment tax policy depends on both direct and equilibrium responses (Proposition 2). Hence, it only suffices to show that tax policy is based on direct short-run responses to tax reforms rather than on total long-run responses. An ideal experiment would thus be an inattention shock which (a) affects short-run responses without affecting long-run responses, but (b) still leads to a change in tax policy.

A shock to tax salience would satisfy the first requirement, provided that taxpayers have correct perceptions of the tax both before and after the shock. Indeed, when equilibrium perceptions are accurate, a change in inattention typically affects short-run responses and thus tax policy under discretion, but does not affect long-run responses and thus tax policy under commitment. Hence, a shock to tax salience which triggers a change in tax policy in a setting where tax perceptions are correct would be strong evidence for the existence of a taxation bias.

Exploiting geographical variations across U.S. states in the use of escrow (a way of remitting property taxes that makes them less salient), Cabral and Hoxby (2012) show that property tax rates are higher when the tax is less salient. At the same time, they provide survey evidence that taxpayers have on average correct perceptions of their property tax liabilities, regardless of their use of escrow.<sup>16</sup> Cabral and Hoxby (2012) thus

---

<sup>16</sup>They also show that taxpayers who use escrow have more dispersed perceptions, but an increase in

provide evidence for the existence of a taxation bias in property taxes.

Looking at the adoption of electronic toll collection (ETC), Finkelstein (2009) shows that the adoption of ETC reduces the direct short-run elasticity of demand and leads to a 20 to 40 percent increase in tolls. Eliciting toll perceptions seven to nine years after the adoption of ETC, she also reports that drivers tend to overestimate the tolls paid with ETC more than those paid in cash. An increase in tax overestimation should however lead to a *decrease* in the tax rate under commitment (Proposition 2).<sup>17</sup> As a result, the findings of Finkelstein (2009) that a decrease in salience leads to higher tax rates and higher tax overestimation is again evidence for the existence of a taxation bias in tolls.

Overall, empirical evidence thus suggests the existence of a taxation bias in different contexts. Going forward, we focus on a particular application and estimate the magnitude and the welfare implications of the taxation bias for the U.S. tax-transfer system.

## 5 Application to the U.S. tax-transfer system

To estimate the magnitude of the taxation bias, we first derive a sufficient statistics formula that we then apply to the U.S. redistributive tax-transfer system. Key determinants of the taxation bias include average inattention to taxes and how this inattention varies across earnings. Second, we complement these results with estimates of the taxation bias obtained from calibrated structural models, allowing us to shed light on the impact of misperceptions and to discuss welfare implications.

### 5.1 Sufficient statistics approach

**Setting.** We maintain the assumptions introduced in Section 3. In particular, the government uses a linear tax-transfer system  $(\tau_g, R_g)$  and the perceived tax rate,  $\tilde{\tau}$ , is a sufficient statistic for earnings decisions,  $y(\tilde{\tau}; w)$ . To derive sufficient statistics results, we however make a number of additional assumptions on tax perceptions.

We assume that individuals' prior and posterior distributions about the tax rate,  $\hat{q}(\tau; w)$  and  $f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)$ , are summarized by sufficient statistics,  $\hat{\tau}(w)$  and  $\tilde{\tau}(w)$ , such as

---

perceptions' heterogeneity should *decrease* the commitment tax rate because this increases the efficiency cost of taxation, as highlighted by Taubinsky and Rees-Jones (2018).

<sup>17</sup>This is most visible in the sufficient statistics characterization of the commitment tax rate, formula (A.17) in Appendix B.4, showing it decreases with the difference between perceived and actual tax rates.

their means or medians.<sup>18</sup> Moreover, we assume that they are linked through:

$$\tilde{\tau}(w) = \xi(\hat{\tau}(w); w)\tau_g + (1 - \xi(\hat{\tau}(w); w))\hat{\tau}(w) , \quad (13)$$

where  $\tau_g$  is the actual tax rate and  $\xi(\hat{\tau}(w); w)$  measures attention to tax changes. Attention may depend on individuals' prior and ability, but it cannot depend on the tax rate. This restriction can be interpreted as an internal consistency requirement: individuals cannot know more ex ante than the information contained in their prior.

Condition (c) in Definition 1 defining equilibrium priors becomes

$$\hat{\tau}(w) \stackrel{\text{eq}}{=} \tau_g + m(\tau_g, R_g; w) , \quad (14)$$

where  $m(\tau_g, R_g; w)$  is a sufficient statistic measuring equilibrium misperceptions in priors that depends on actual tax policy and individuals' skills. As a result, misperceptions in perceived tax rates,  $\varpi(w) := \tilde{\tau}(w) - \tau_g$ , are in equilibrium equal to

$$\varpi(w) \stackrel{\text{eq}}{=} (1 - \xi(\hat{\tau}(w); w)) m(\tau_g, R_g; w) . \quad (15)$$

Misperceptions in perceived tax rates thus depend both on inattention and on prior misperceptions, and both can be heterogeneous. In the following, we use  $\xi(w)$  and  $m(w)$  as shorthand for  $\xi(\hat{\tau}(w); w)$  and  $m(\tau_g, R_g; w)$  to economize on notations.

**Taxation bias.** As before, discretionary tax policy depends on direct effects of changes in tax rates on tax perceptions holding priors constant,  $\frac{d\tilde{\tau}(w)}{d\tau_g} \Big|_{\hat{\tau}(w)} = \xi(w)$ , whereas tax policy under commitment also incorporates equilibrium effects,  $\frac{d\tilde{\tau}(w)}{d\tau_g} \stackrel{\text{eq}}{=} \Xi(w)$ , where

$$\Xi(w) := 1 + \left(1 - \xi(w)\right) \frac{dm(w)}{d\tau_g} - m(w) \frac{d\xi(w)}{d\hat{\tau}(w)} \left(1 + \frac{dm(w)}{d\tau_g}\right) . \quad (16)$$

The first term in (16) captures the fact that, in the absence of misperceptions, the equilibrium perceived tax rate moves one-for-one with the actual tax rate,  $\Xi(w) = 1$ . The remaining terms account for equilibrium adjustments of misperceptions, either because of changes in prior misperceptions or because of changes in equilibrium attention (induced by changes in equilibrium priors).

---

<sup>18</sup>To simplify the exposition, we assume unidimensional heterogeneity indexed by  $w$ , but Appendix B.4 shows that all sufficient statistics results extend to multidimensional heterogeneity in, e.g., tax perceptions, implying that these results extend to prior and posterior distributions. Also, in this sufficient statistics approach, we normalize implementation shocks to zero such that  $\tau_0 = \tau_g$ .

Denoting  $\bar{x} := \int_w x(w)f_w(w)dw$  the mean of variable  $x$  and  $g(w) := \frac{G'(V)}{p}$  the social marginal welfare weights, we characterize the taxation bias in the following proposition.<sup>19</sup>

**Proposition 4.** *In the absence of misperceptions and when the frictionless elasticity of labor supply  $e$  is uniform, the taxation bias is approximately given by*

$$\frac{\tau_g^{\text{eq}} - \tau_g^*}{\tau_g^*} \simeq (1 - \tau_g^{\text{eq}}) \frac{\overline{y(1 - \xi)}}{y\xi} = (1 - \tau_g^{\text{eq}}) \frac{\overline{1 - \xi} + \text{Cov}(y, 1 - \xi)/\bar{y}}{1 - \overline{1 - \xi} - \text{Cov}(y, 1 - \xi)/\bar{y}}. \quad (17)$$

With misperceptions and heterogeneous labor supply elasticities, this formula extends to

$$\frac{\tau_g^{\text{eq}} - \tau_g^*}{\tau_g^*} \simeq (1 - \tau_g^{\text{eq}}) \frac{\overbrace{y(1 - \xi)e}^{\mathcal{A}} + \overbrace{g\varpi y(1 - \xi)e}^{\mathcal{B}} + \overbrace{y(\Xi - 1)e}^{\mathcal{C}} + \overbrace{g\varpi y(\Xi - 1)e}^{\mathcal{D}}}{y\xi e}. \quad (18)$$

Without misperceptions and when the frictionless elasticity of labor supply is uniform, formula (17) shows that the magnitude of the taxation bias can be estimated from a restricted set of sufficient statistics: the observed tax rate,  $\tau_g^{\text{eq}}$ , average inattention,  $\overline{1 - \xi}$ , and the covariance between income and inattention,  $\text{Cov}(y, 1 - \xi)$ , scaled by average income,  $\bar{y}$ . In particular, the magnitude of the taxation bias increases with average inattention and the covariance, because both increase the potential revenue gains from discretionary tax hikes. Holding average inattention constant, this implies that the taxation bias is lower (resp. higher) when high income individuals are relatively more (resp. less) attentive to taxes. When the covariance is nil and attention is uniform, we retrieve the expression for the taxation bias from the stylized model of Section 2.

Formula (18) generalizes this result. In the absence of misperceptions,  $\varpi = 0$ , we have  $\mathcal{B} = \mathcal{C} = \mathcal{D} = 0$  such that formula (18) is a simple extension of formula (17) accounting for heterogeneity in the frictionless elasticity of labor supply. Hence,  $\mathcal{A}$  captures the aforementioned potential revenue gains from discretionary tax hikes related to inattention,  $1 - \xi$ . With misperceptions,  $\mathcal{C}$  captures similar revenue effects from commitment tax hikes related to equilibrium adjustments of misperceptions,  $\Xi - 1$ . These revenue effects thus contribute to the taxation bias in proportion to  $\mathcal{A} + \mathcal{C} = \overline{y(\Xi - \xi)}e$ .

Misperceptions also introduce misoptimization and thus a corrective motive for taxation (see Proposition 1) that depends on welfare weights, the size of misperceptions, and the extent to which taxation can correct earnings choices. Since tax changes shift tax

---

<sup>19</sup>This characterization is an approximation evaluating all sufficient statistics at the observed tax rate, interpreted as the equilibrium tax rate under discretion,  $\tau_g^{\text{eq}}$ , and assuming small misperceptions.

perceptions by a factor  $\Xi$  under commitment and by a factor  $\xi$  under discretion, these corrective effects generally differ and contribute to the taxation bias in proportion to  $\mathcal{B} + \mathcal{D} = \overline{g\varpi y(\Xi - \xi)e}$ . In particular, when individuals underestimate tax rates,  $\varpi < 0$ , these corrective motives call for larger corrective tax increases under commitment than under discretion, thereby reducing the magnitude of the taxation bias.

**Empirical estimates.** We now leverage these sufficient statistics results to estimate the magnitude of the taxation bias using formula (17). Although formula (18) yields qualitative insights on the impact of misperceptions, it is harder to connect with data as the relevant sufficient statistics are difficult to estimate empirically.

The U.S. redistributive tax-transfer system is well approximated by a linear tax-transfer schedule (Ferriere et al., 2023) and using OECD data on 2016 labor taxes we estimate the tax rate to be 29.5%. At this tax rate, recent evidence on attention to taxes point to an average inattention of  $\overline{1 - \xi} = 0.45$ .<sup>20</sup> Assuming inattention is uniform across incomes, this implies a taxation bias of 10.8 percentage points (p.p.) meaning that while the actual tax rate is 29.5%, the tax rate under commitment would be close to 19%. As reported in Table 1, we get a taxation bias of 8.1 p.p. and 13.6 p.p. when average inattention is 0.35 and 0.55 respectively.

One can however question the assumption that inattention is uniform across incomes. In the context of sales taxes, direct evidence is mixed.<sup>21</sup> In the context of income taxes, we are not aware of any estimate of the relationship between inattention and income, although some of the evidence indicates that high-income individuals are less likely to be confused about income taxes (Feldman et al., 2016; Aghion et al., 2017). If anything, this suggests that inattention decreases with income. Since this reduces the magnitude of the taxation bias, we adopt this conservative assumption and rely on a rational inattention model (described in the next subsection) to infer this covariance. When the model is calibrated to target an average inattention of  $\overline{1 - \xi} = 0.45$ , the implied covariance term

---

<sup>20</sup>Gabaix (2019) reports two point estimates (0.41 and 0.54 from DellaVigna and Pollet, 2009) at a stake of 30%, while his meta-analysis points to a higher value of 0.67. Measuring attention to sales taxes, Taubinsky and Rees-Jones (2018) find an average attention of 0.25 (respectively 0.48) for a tax rate of 7% (respectively 22%), which implies an average attention of 0.59 for a tax rate of 30% by linear extrapolation. We take the average of these four estimates as a baseline,  $\bar{\xi} = 0.55$ .

<sup>21</sup>There is evidence that attention to sales taxes increases with income (Taubinsky and Rees-Jones, 2018), is uniform (Morrison and Taubinsky, 2021), and decreases with income (Kroft et al., 2020).

is  $Cov(y, 1 - \xi)/\bar{y} = -0.18$  and the taxation bias reduces to 6.0 percentage points. We find similar reductions for the other values of average inattention (see Table 1).

Table 1: Magnitude of the taxation bias in the U.S. tax-transfer system

<i>Panel A: Sufficient statistics estimates</i>	Average inattention, $\overline{1 - \xi}$		
	0.35	0.45	0.55
Uniform inattention, $Cov(y, 1 - \xi) = 0$	8.1	10.8	13.6
Correlated inattention, $Cov(y, 1 - \xi) < 0$	4.5	6.0	7.9
<i>Panel B: Structural estimates</i>	Model 1	Model 2	Model 3
	$Cov(y, 1 - \xi) = 0$	$Cov(y, 1 - \xi) < 0$	$Cov(y, m) > 0$
Without misperceptions, $\bar{\omega} = 0$	8.7	4.7	–
With misperceptions, $\bar{\omega} = -5.6$	7.7	3.7	3.7

NOTE: All estimates represent the magnitude of the taxation bias,  $\tau_g^{eq} - \tau_g^*$ , in percentage points (p.p.). Estimates in panel A follow from formula (17) and negative covariances are estimated using Model 2 with  $Cov(y, 1 - \xi)/\bar{y}$  equal to -0.15, -0.18, and -0.21 when average inattention is 0.35, 0.45, and 0.55 respectively. Estimates in panel B are obtained from numerical simulations using the calibrated structural models from Section 5.2 where average inattention essentially coincides with its baseline value,  $\overline{1 - \xi} = 0.45$ .

## 5.2 Structural approach and welfare implications

The sufficient statistics characterization of the taxation bias is intuitive and allows to obtain empirical estimates relying on a small set of empirical moments. However, these sufficient statistics estimates are only approximations and they cannot easily account for the presence of misperceptions, nor speak to the welfare implications of the taxation bias. To shed light on these issues, we now turn to calibrated structural models.

**Models.** Given the uncertainty surrounding inattention and tax perceptions, we consider three different models in the hope to provide structural estimates that are robust to potential misspecification. In the first model (Model 1), attention is exogenous and uniform,  $Cov(y, 1 - \xi) = 0$ , and priors are either correct,  $m(w) = 0$ , or uniformly biased,  $m(w) = m$ , implying that  $Cov(y, m) = 0$  and that  $\Xi = 1$ . We use this simple benchmark to quantify the approximation made in the sufficient statistics formula (17) and to provide first welfare estimates.

In the second model (Model 2) attention is endogenous, consistent with evidence that attention increases with stakes (e.g., Hoopes et al., 2015; Gabaix, 2019). Building on

the rational inattention literature (Sims, 2003), each individual has a prior distribution and observes a noisy Gaussian signal about the tax rate, whose precision is endogenously chosen. We measure the information provided by the signal from the entropy reduction between the prior and the posterior, and posit that individuals incur a utility cost,  $\kappa > 0$ , per unit of information (bit) that they acquire. Since more productive individuals have more latitude to adjust their earnings choices, they endogenously acquire more information and attention increases with income.<sup>22</sup> It implies that  $Cov(y, 1-\xi)$  is endogenously negative. We again assume that priors are either correct,  $m(w) = 0$ , or uniformly biased,  $m(w) = m$ , such that  $Cov(y, m) = 0$ . However, in the presence of misperceptions, equilibrium adjustments in attention generally imply that  $\Xi \neq 1$ .

In the third model (Model 3), attention is exogenous and uniform, but prior misperceptions are endogenous. Relying on the ironing heuristic (De Bartolome, 1995; Liebman and Zeckhauser, 2004), we assume that individuals use their average tax rate as a prior and are partially attentive to their marginal tax rates. In line with the findings of Rees-Jones and Taubinsky (2020) that only a fraction of taxpayers use the ironing heuristic, this can be interpreted as a mixed population of fully inattentive ironers and fully attentive individuals, where the attention parameter is the share of fully attentive individuals. We thus assume that  $Cov(y, 1-\xi) = 0$ , while the ironing heuristic implies income-specific prior misperceptions,  $Cov(y, m) > 0$ , that endogenously adjust in equilibrium,  $\Xi \neq 1$ .

All models share otherwise the structure of standard optimal income tax models (e.g., Saez, 2001). We assume the existence of a representative individual at each earnings ability,  $w$ , with quasi-linear and iso-elastic utility,  $U(c, y; w) = c - (y/w)^{1+\varepsilon}/(1+\varepsilon)$ . All models can be nested within the general model of Section 3 (see Table A.1 for a formal correspondence) and we derive tax policies under discretion and under commitment from Proposition 1 and Proposition 2 respectively. Additional details and full micro-foundations are in Online Appendix D.

**Calibration.** When including misperceptions, we rely on the findings of Rees-Jones and Taubinsky (2020) that individuals underestimate their marginal tax rates by 19% on

---

<sup>22</sup>This pattern is reinforced if the cost to acquire information,  $\kappa$ , decreases with type,  $w$ , because, e.g., earnings ability also reflect cognitive abilities. In contrast, this pattern is attenuated if this cost increases with type because, e.g., a higher earnings ability also implies a higher opportunity cost of time.

average, meaning that average equilibrium misperceptions are  $\bar{\omega} = -0.19 * 29.5 = -5.6$  percentage points. In the first model, combined with an exogenous inattention parameter of 0.45, this allows us to calibrate prior misperceptions,  $m$ . In the second model with rationally inattentive taxpayers, we calibrate the cost of information,  $\kappa$ , to match an average inattention of 0.45, and similarly use equilibrium misperceptions to calibrate prior misperceptions. In the third model, assuming individuals use their average tax rates as a prior, average equilibrium misperceptions imply that the exogenous share of ironers (inattention parameter) is 0.47, very close to our baseline value of 0.45.

All models are otherwise calibrated in a similar fashion. The U.S. redistributive tax-transfer system is approximated by a linear tax rate of 29.5% financing a fixed amount of expenditures set to 17.5% of observed total income and a transfer close to \$5,000. The frictionless elasticity of labor supply,  $e$ , is set to 0.33 (Chetty, 2012; Kostøl and Myhre, 2021). Given the observed tax-transfer schedule and misperceptions, the distribution of earnings ability,  $w$ , is obtained by inverting the income distribution from 2016 March CPS data extended with a Pareto tail ( $k = 2$ ) for incomes above \$200,000 (e.g., Saez, 2001). We assume that the government has a Benthamite objective,  $G(\mathcal{V}) = \mathcal{V}^{1-\alpha}/(1-\alpha)$ , with parameter  $\alpha \in (0, 1)$  calibrated such that the equilibrium tax rate under discretion is the observed U.S. tax rate. Table A.1 reports calibrated and inferred parameters. Additional information about simulations can be found in Online Appendix E.

**Empirical estimates.** We report structural estimates of the taxation bias in panel B of Table 1. When attention is uniform across incomes, the baseline sufficient statistics estimate is 10.8 p.p. whereas the baseline structural estimate is 8.7 p.p. (Model 1). This difference underlines that evaluating all sufficient statistics at the observed tax rate leads to an upward error in our sufficient statistics estimates. This structural estimate reduces to 7.7 p.p. when introducing exogenous and uniform misperceptions in priors ( $\mathcal{B} < 0$  and  $\mathcal{C} = \mathcal{D} = 0$ ). These structural estimates become 4.7 p.p. and 3.7 p.p. meaning that they decrease by 4 p.p. each in the rational inattention model (Model 2) where the covariance between inattention and income is negative. Finally, the taxation bias also reduces to 3.7 p.p. in the ironing model (Model 3) where misperceptions are endogenous to the tax-transfer system ( $\mathcal{B} < 0, \mathcal{C} < 0$ , and  $\mathcal{D} > 0$ ).

We carry out a sensitivity analysis and report further estimates in Table A.2. Overall,

our estimates point to a taxation bias in the U.S. redistributive tax-transfer system ranging between 3 and 8 percentage points. This is saying that between a tenth and a fourth of actual tax liabilities are attributable to a taxation bias.

**Welfare implications.** The inefficiently high tax rates implied by the presence of a taxation bias create a welfare loss. However, this does not necessarily imply that inattention to taxes decreases welfare. Indeed, inattention may also be conducive to taxpayers underestimating their tax rates, which reduces the efficiency cost of taxation and results in a countervailing positive welfare effect. The net welfare effect of inattention is thus ambiguous when individuals underestimate tax rates.

To illustrate how these two opposing forces affect welfare, denote  $SW^{\text{eq}}$ ,  $SW^*$ , and  $SW^{\text{FI}}$  the equilibrium social welfare under discretion, commitment, and full information respectively. All three measures coincide when taxpayers are fully attentive, but it is no longer the case in the presence of inattention. We thus decompose the net welfare change between the welfare loss induced by the taxation bias and the welfare gain from tax underestimation as follows:

$$\Delta SW \equiv SW^{\text{eq}} - SW^{\text{FI}} = \underbrace{SW^{\text{eq}} - SW^*}_{\text{taxation bias}} + \underbrace{SW^* - SW^{\text{FI}}}_{\text{underestimation}} \quad (19)$$

Relying on the previous structural models calibrated with an average inattention of 0.45 and an average tax underestimation of  $-5.6$ , we estimate the welfare effect of inattention. In Model 1 featuring exogenous and uniform attention and misperceptions, the welfare loss from the taxation bias is equivalent to decreasing the consumption of all taxpayers by \$82 per year, while the welfare gain from underestimation is equivalent to increasing consumption by \$179. Consequently, the net welfare effect is \$98, and the taxation bias reduces the potential welfare gain from tax underestimation by 46%. In Models 2 and 3, welfare losses are relatively smaller (\$34 and \$13 per year, respectively) and reduce potential welfare gains from tax underestimation by about 20%.

Figures A.1, A.2, and A.3 represent for Model 1, 2, and 3 respectively how tax rates under discretion and commitment (top left panels), the magnitude of the taxation bias (top right panels) and welfare effects (bottom panels) vary with the degree of inattention. When inattention is low, the taxation bias is rather small, and the welfare gains from tax underestimation dominate. The net welfare effect is therefore positive and it

increases with inattention. However, as inattention gets larger, the tax rate under discretion increases faster than the tax rate under commitment such that the size of the taxation bias and the associated welfare losses increase faster than the gains from tax underestimation. Consequently, in the three models that we consider, the net welfare effect of inattention decreases with the degree of inattention, and turns negative when inattention is sufficiently high.

## 6 Extension to nonlinear tax-transfer schedules

In this section, we extend the analysis to nonlinear tax-transfer schedules. When inattention varies across incomes, the taxation bias affects the progressivity of the tax-transfer system, which tends to magnify its welfare impact.

**Setting.** We introduce a nonlinear tax-transfer schedule  $T_g(y)$ , but the setup is otherwise unchanged. In particular, the perceived marginal tax rate,  $\tilde{\tau}(w)$ , remains a sufficient statistics for earnings decisions,  $y(\tilde{\tau}(w); w)$ , as we assume that taxpayers use a locally linear representation of the nonlinear schedule. This linear representation raises a new question: in the continuum of marginal tax rates,  $\{T'_g(y)\}_y$ , which marginal tax rate are taxpayers attentive to?

We assume that an individual with ability  $w$  gathers information about the marginal tax rate,  $T'_g(\hat{y}(w))$ , at her ex ante optimal earnings level,  $\hat{y}(w) := y(\hat{\tau}(w); w)$ , such that

$$\tilde{\tau}(w) = \xi(\hat{\tau}(w); w)T'_g(\hat{y}(w)) + (1 - \xi(\hat{\tau}(w); w))\hat{\tau}(w) . \quad (20)$$

This can be interpreted as an internal consistency requirement: individuals cannot know more ex ante than the information contained in their prior. This gives a novel allocative role to the prior as it determines the marginal tax rate individuals pay attention to.<sup>23</sup>

Introducing nonlinear schedules does not affect the definition of equilibrium but we assume that there is in equilibrium a strictly increasing mapping between earnings,  $y$ , and ability,  $w$ , a requirement known as the monotonicity condition. Absent misperceptions,

---

<sup>23</sup>For instance, a fully attentive taxpayer will perfectly observe  $\tilde{\tau}(w) = T'_g(\hat{y}(w))$  and choose earnings  $y_{w|\tilde{\tau}(w)} = y(T'_g(\hat{y}(w)); w)$ , whereas in the full information benchmark she would choose earnings as the solution to a fixed-point problem characterized by  $y_w = y(T'_g(y(.)); w)$ . These earnings choices coincide in the absence of prior misperceptions,  $m(w) = 0$ , but will differ with prior misperceptions,  $m(w) \neq 0$ .

the Spence-Mirrlees single crossing condition ensures that the monotonicity condition holds when  $T_g''(\cdot)$  is smooth enough, as in the full information benchmark.<sup>24</sup> With misperceptions, attention  $\xi(w)$  and prior misperceptions  $m(w)$  must also be smooth enough to prevent sharp changes in misperceptions across individuals (see Online Appendix F).

**Taxation bias.** We solve for the optimal tax schedule using a perturbation approach. That is, we consider a reform that consists of a small increase  $\Delta\tau^r$  in marginal tax rates within a small bandwidth of earnings  $[y^r - \Delta y, y^r]$  and derive its impact on the objective function of the government. Following the literature, this reform triggers fiscal and welfare effects. In this setting, analyzing the impact of a reform requires a careful identification of the individuals affected by the reform.

Mechanical fiscal and welfare effects capture changes in taxes and welfare for taxpayers with actual earnings above  $y^r$ . Using the monotonicity condition, this corresponds to individuals with an ability level above  $w^r$ , defined by  $y(\tilde{\tau}(w^r); w^r) := y^r$ . Behavioral effects come from taxpayers who are learning the marginal tax rate affected by the reform. That is, individuals whose ex ante optimal earnings level,  $\hat{y}$ , belongs to  $[y^r - \Delta y, y^r]$  and whose ability,  $w$ , belongs to  $[\hat{w}^r - \Delta\hat{w}, \hat{w}^r]$  defined by  $y(\hat{\tau}(\hat{w}^r); \hat{w}^r) := y^r$ . The two cut-offs  $w^r$  and  $\hat{w}^r$  coincide in the absence of misperceptions, but differ with misperceptions.

Assuming misperceptions are small such that the two cut-offs (approximately) coincide, we characterize optimal nonlinear schedules under discretion and under commitment in terms of sufficient statistics. We then compute the difference to obtain the following sufficient statistics characterization of the taxation bias:<sup>25</sup>

**Proposition 5.** *With a nonlinear tax-transfer schedule, a general sufficient statistics formula for the taxation bias at each earnings level,  $y$ , is*

$$\frac{T_g^{\text{eq}}(y) - T_g^{*'}(y)}{T_g^{*'}(y)} \simeq (1 - T_g^{\text{eq}}(y)) \frac{\overbrace{(1 - \xi(y))}^{\mathcal{A}} + \overbrace{g(y)\varpi(y)(1 - \xi(y))}^{\mathcal{B}} + \overbrace{(\Xi(y) - 1)}^{\mathcal{C}} + \overbrace{g(y)\varpi(y)(\Xi(y) - 1)}^{\mathcal{D}}}{\xi(y)}. \quad (21)$$

Formula (21) is a generalization to nonlinear tax-transfer schedules of formula (18). Both formulas are similar and the interpretation of the different terms remains the same,

<sup>24</sup>Following Jacquet and Lehmann (2017), we assume that  $T_g(\cdot)$  is twice differentiable and that each taxpayer's optimization problem admits a unique global maximum with strict second-order conditions.

<sup>25</sup>As in the linear case, this characterization is an approximation evaluating all sufficient statistics at the observed schedule, interpreted as the equilibrium policy under discretion, and relying on small misperceptions. Moreover, it can also be extended to settings with multidimensional heterogeneity.

except that each term is now earnings-specific. As a result, inattention or misperceptions at a given earnings level influence deviations in marginal tax rates at that earnings level. Hence, a correlation between inattention and income, or between misperceptions and income, implies that the taxation bias affects the progressivity of tax-transfer schedules.

**Numerical illustration.** We illustrate this effect on progressivity when inattention decreases with income. We simulate the optimal nonlinear tax-transfer schedules under discretion and under commitment in the rational inattention model (Model 2) introduced in Section 5. The model is calibrated to match the baseline value for average inattention,  $\overline{1-\xi} = 0.45$ , under the actual U.S. tax-transfer schedule, which implies as before that  $Cov(y, 1-\xi)/\bar{y} = -0.18$ .<sup>26</sup> We assume away misperceptions such that the commitment policy does not depend on attention and is the same as in, e.g., Saez (2001), with the recognizable U-shape pattern of marginal tax rates.

The difference between the two nonlinear schedules depicted in Figure 2 represents the magnitude of the taxation bias. Since inattention decreases with income, so does the taxation bias: it leads to an increase in marginal tax rates of 17, 7, and 2 percentage points at the first decile, the median, and the top decile of the income distribution where inattention is 0.69, 0.38, and 0.16 respectively.<sup>27</sup>

This distortion in tax-transfer progressivity magnifies the taxation bias and its welfare impact. With a linear tax-transfer schedule, the taxation bias is 4.7 p.p. (Table 1) and it imposes annual welfare loss of \$40 per taxpayer. With a nonlinear tax-transfer schedule, the average taxation bias across earnings is 8.8 p.p. and it imposes an annual welfare loss of \$89, implying that the welfare loss more than doubles.

## 7 Discussion

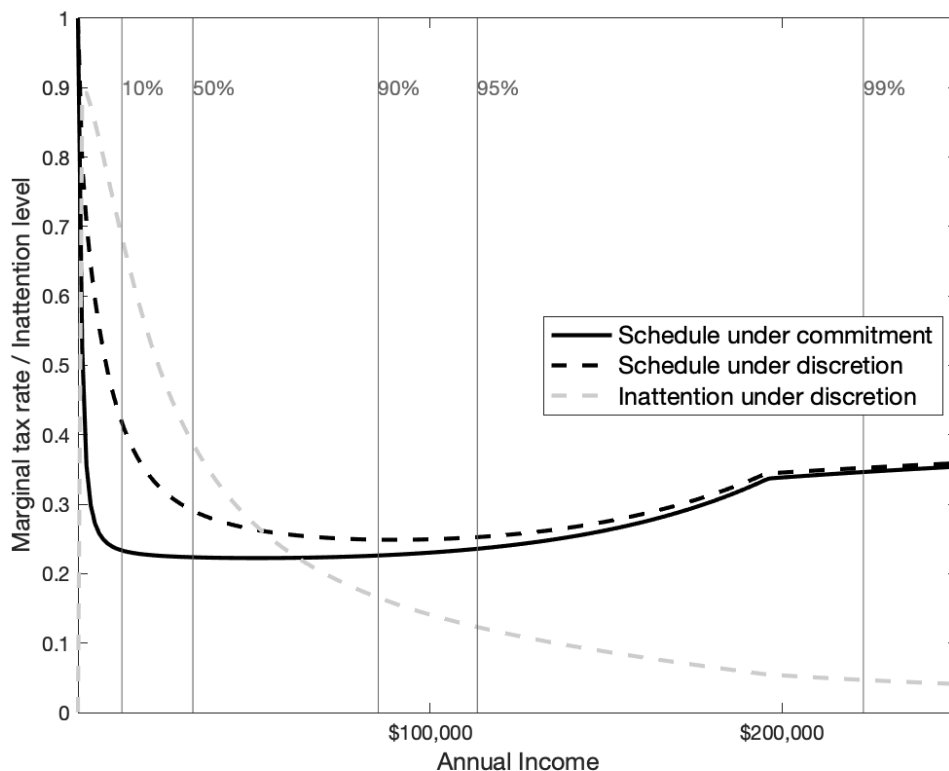
In this final section, we discuss the role of information asymmetries, tax design, and competitive politics. This leads us to the conclusion that eliminating the taxation bias requires a commitment technology that can only be delivered through institutional changes.

---

<sup>26</sup>We again use a linear approximation of the actual U.S. tax-transfer schedule and calibrate the government's objective such that the equilibrium linear tax rate under discretion is the actual tax rate.

<sup>27</sup>Inattention goes asymptotically to 0 at the origin of the income distribution because the value of collecting information becomes very large as marginal tax rates go asymptotically to 1.

Figure 2: Implications for the progressivity of nonlinear schedules



NOTE: This Figure represents the schedule of marginal tax rates across income levels under commitment (solid black line) and under discretion (dashed black line) as well as the level of inattention under discretion (dashed grey line) simulated using the calibrated rational inattention model (Model 2) introduced in Section 5.

**Information asymmetries.** If deviations from the commitment tax policy were observable ex post, it would be easy to solve the government’s commitment problem through political constraints or reputational concerns. These deviations cannot, however, be observed because of information asymmetries. For instance, the government is privately informed about its spending needs (Sleet, 2004) and its redistributive objective.<sup>28</sup> These information asymmetries, and the lack of accountability they imply, are well recognized and motivate recurring initiatives for higher fiscal transparency initiated, for instance, by the OECD and the IMF. However, governments do not have strong incentives to increase fiscal transparency (Alt et al., 2006), they rather have incentives to manipulate their communication about taxation (Albornoz et al., 2014). As a result, there are good reasons not to expect voluntary fiscal transparency.

<sup>28</sup>This motivates the inverse-optimum approach inferring policymakers’ welfare weights from actual tax policy (Bourguignon and Spadaro, 2012; Lockwood and Weinzierl, 2016; Jacobs et al., 2017).

**Tax design.** The time-inconsistency of the government is a byproduct of inattention. Therefore, any improvement in tax design to decrease information costs and increase attention is likely to reduce the magnitude of the taxation bias. Typical examples would be increases in tax salience, or the simplification of tax schedules shown by Abeler and Jäger (2015) to increase attention to tax reforms.<sup>29</sup> Nevertheless, only perfect attention can resolve the commitment problem of the government, which even the most salient and simplest tax system could fail to provide. Moreover, imperfect attention can be socially preferable if taxpayers underestimate marginal tax rates, or if tax complexity allows to better account for taxpayer heterogeneity (Kaplow, 1994; Kleven and Kopczuk, 2011).

**Competitive politics.** Because of information asymmetries and the complexity of tax design, “[t]here is a limit to how well-informed the electorate can be expected to be for the purpose of holding the government to account and ensuring good quality tax policy.” (Mirrlees Review, Alt et al., 2010, p. 3). One may nonetheless wonder whether competitive politics could alleviate the commitment problem of the government.

This seems theoretically unlikely, because we consider the behavior of a welfarist government such that any decision-maker (e.g. the median voter) with similar redistributive tastes would face the same time-inconsistency problem and make the same policy choices. Moreover, Matějka and Tabellini (2021) show in related work that electoral competition with inattentive voters also leads to policy distortions.<sup>30</sup> Empirically, evidence on the effect of competitive politics suggests that it fosters the use of less salient tax instruments (Bracco et al., 2019) and that it exacerbates time-inconsistency problems, when time-discounting factors are heterogeneous across voters (Jackson and Yariv, 2014, 2015).

---

<sup>29</sup>Initiatives to block tax reforms on the grounds that they are excessively complex or to promote simplification of the tax code include rulings from the French Constitutional Council (2005; 2012) and a Report of the President’s Advisory Panel on Federal Tax Reform (2005) in the United States.

<sup>30</sup>In the presence of a wedge between the objective of the government and that of a welfare-maximizing social planner, the taxation bias reflects the deviation from the commitment policy based on the government tastes for redistribution. Since we infer these tastes from current tax policy, our estimation of the magnitude of the taxation bias is robust to the existence of such a wedge. Our welfare analysis, however, implicitly assumes that both objectives coincide. For instance, if the government chooses to implement lower redistribution than would be socially optimal, the welfare effect of a taxation bias becomes positive.

**Commitment technologies.** In conclusion, we cannot expect voluntary fiscal transparency, and improvements in tax design or competitive politics are unlikely to resolve the commitment problem of the government in its choice of tax policy. Solving this problem thus requires tangible commitment technologies. Building on the large literature on the inflation bias, we mention two potential solutions.

First, imposed fiscal transparency may allow an independent third party to observe deviations from the commitment policy. In such a scenario, public reports of policy deviations could solve the commitment problem through reputational concerns (Barro and Gordon, 1983b).<sup>31</sup> Second, in a dynamic setting, the tax policy implemented by an infinitely patient fiscal authority that internalizes taxpayers' learning coincides with the commitment tax policy (see Online Appendix C.2). Delegating the conduct of tax policy to such a fiscal authority could therefore eliminate the taxation bias.

Our point is not to argue in favor of either solution, but instead to highlight that both would require substantial institutional changes that pose important practical and political challenges. We leave the analysis of these issues to future work.

## 8 Conclusion

Building on recent evidence on tax perceptions, this paper shows that the anchoring effect of inattention to taxes creates a time-inconsistency problem in the choice of tax policy. As a result, when the government cannot credibly commit to its tax policy, inattention induces inefficient increases in tax rates and a deviation from the socially optimal commitment policy. We call this deviation a taxation bias and estimate that it ranges between 3 and 8 percentage points in the U.S. redistributive tax-transfer system, implying that between a tenth and a fourth of actual tax liabilities can be attributed to a taxation bias. We also show that the taxation bias distorts the progressivity of nonlinear tax-transfer schedules, which magnifies its negative welfare impact. Overall, our findings sheds new light on the implications of inattention and misperceptions.

While we focus on the redistributive tax-transfer system that applies to labor income, a taxation bias is bound to appear in virtually all tax instruments, insofar as they are

---

<sup>31</sup>Alternatively, a principal-agent optimal contract à la Walsh (1995) may also be considered if fiscal transparency cannot be imposed.

not perfectly salient. To the extent that inattention varies across tax instruments, the taxation bias induces larger increases in taxes that individuals are the least attentive to, which suggests that the taxation bias also distorts the allocation of the tax burden across tax bases. More broadly, the insight that inattention leads to time-inconsistency issues likely extends to a wide range of applications, in which the framework developed in this paper could be fruitfully applied.

## References

- Abeler, J. and Jäger, S. (2015). Complex tax incentives. *American Economic Journal: Economic Policy*, 7(3):1–28. 1, 30
- Aghion, P., Akcigit, U., Lequien, M., and Stantcheva, S. (2017). Tax simplicity and heterogeneous learning. *NBER Working Paper*. 16, 21
- Albornoz, F., Esteban, J., and Vanin, P. (2014). Market distortions and government transparency. *Journal of the European Economic Association*, 12(1):200–222. 29
- Allcott, H., Lockwood, B., and Taubinsky, D. (2018). Ramsey strikes back: Optimal commodity tax and redistribution in the presence of salience effects. In *AEA Papers and Proceedings*, volume 108, pages 88–92. 4
- Alt, J., Preston, I., and Sibieta, L. (2010). The political economy of tax policy. *Institute of Fiscal Studies, Dimensions of Tax Design: The Mirrlees Review, Oxford*. 30
- Alt, J. E., Lassen, D. D., and Rose, S. (2006). The causes of fiscal transparency: Evidence from the us states. *IMF Staff papers*, 53(1):30–57. 29
- Barro, R. J. and Gordon, D. B. (1983a). A positive theory of monetary policy in a natural rate model. *Journal of Political Economy*, 91(4):589–610. 1, 5, 8, 16
- Barro, R. J. and Gordon, D. B. (1983b). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12(1):101–121. 31
- Bernheim, B. D. and Taubinsky, D. (2018). Behavioral public economics. In *Handbook of Behavioral Economics*, volume 1, pages 381–516. Elsevier. 4, 15
- Billi, R. M. (2011). Optimal inflation for the us economy. *American Economic Journal: Macroeconomics*, 3(3):29–52. 5
- Blesse, S., Buhlmann, F., and Doerrenberg, P. (2019). Do people really want a simple tax system? Evidence on preferences towards income tax simplification. 4

- Bourguignon, F. and Spadaro, A. (2012). Tax–benefit revealed social preferences. *The Journal of Economic Inequality*, 10(1):75–108. 29
- Bracco, E., Porcelli, F., and Redoano, M. (2019). Political competition, tax salience and accountability: Theory and evidence from Italy. *European Journal of Political Economy*, 58:138–163. 30
- Buchanan, J. M. (1967). The fiscal illusion. *Public finance in democratic process: Fiscal Institutions and individual choice*. 4
- Burman, L. E., Gale, W. G., and Rohaly, J. (2005). The expanding reach of the individual alternative minimum tax. *Journal of Economic Perspectives*, 17(2). 16
- Cabral, M. and Hoxby, C. (2012). The hated property tax: salience, tax rates, and tax revolts. Technical report, National Bureau of Economic Research. 2, 4, 15, 17
- Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018. 6, 7, 24
- Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–77. 1
- Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, 120(1):116–159. 15
- Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–78. 5, 15
- Conseil Constitutionnel (2005). Décision n. 2005-530 DC du 29 décembre 2005. 30
- Conseil Constitutionnel (2012). Décision n. 2012-662 DC du 29 décembre 2012. 30
- Craig, A. C. and Slemrod, J. (2022). Tax knowledge and tax manipulation: A unifying model. Technical report, National Bureau of Economic Research. 4, 17
- De Bartolome, C. A. (1995). Which tax rate do people use: Average or marginal? *Journal of Public Economics*, 56(1):79–96. 23
- DellaVigna, S. and Pollet, J. M. (2009). Investor inattention and Friday earnings announcements. *The Journal of Finance*, 64(2):709–749. 21
- Dollery, B. E. and Worthington, A. C. (1996). The empirical analysis of fiscal illusion. *Journal of Economic Surveys*, 10(3):261–297. 4
- Dovis, A. and Kirpalani, R. (2020). Fiscal rules, bailouts, and reputation in federal governments. *American Economic Review*, 110(3):860–88. 5

- Dupuit, J. (1995). De la mesure de l'utilité des travaux publics (1844). *Revue française d'économie*, 10(2):55–94. 1
- Eichengreen, B. (1989). The capital levy in theory and practice. Technical report, National Bureau of Economic Research. 5, 16
- Farhi, E. and Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review*, 110(1):298–336. 1, 2, 4, 10, 12, 14
- Farhi, E., Sleet, C., Werning, I., and Yeltekin, S. (2012). Non-linear capital taxation without commitment. *Review of Economic Studies*, 79(4):1469–1493. 5
- Feenberg, D. R. and Poterba, J. M. (2004). The alternative minimum tax and effective marginal tax rates. *National Tax Journal*, 57(2):407–427. 16
- Feldman, N. E., Katus ěÁk, P., and Kawano, L. (2016). Taxpayer confusion: Evidence from the child tax credit. *American Economic Review*, 106(3):807–835. 21
- Ferriere, A., Grübener, P., Navarro, G., and Vardishvili, O. (2023). On the optimal design of transfers and income-tax progressivity. 21
- Finkelstein, A. (2009). E-ztax: Tax salience and tax rates. *The Quarterly Journal of Economics*, 124(3):969–1010. 2, 4, 15, 18
- Fochmann, M. and Weimann, J. (2013). The effects of tax salience and tax experience on individual work efforts in a framed field experiment. *FinanzArchiv: Public Finance Analysis*, 69(4):511–542. 15
- Furukawa, C. (2022). Tax transparency and social welfare: The role of government commitment. *Available at SSRN 3862433*. 4
- Gabaix, X. (2019). Behavioral inattention. In *Handbook of Behavioral Economics, Vol 2, edited by D Bernheim, S DellaVigna and D Laibson, V.2*, pages 261–343. Elsevier. 1, 4, 5, 7, 21, 22, 56
- Gale, W. G. and Potter, S. R. (2002). An economic evaluation of the economic growth and tax relief reconciliation act of 2001. *National Tax Journal*, 55(1):133–186. 16
- Gerritsen, A. (2016). Optimal taxation when people do not maximize well-being. *Journal of Public Economics*, 144:122–139. 4, 12, 14
- Goldfayn-Frank, O. and Wohlfart, J. (2020). Expectation formation in a new environment: Evidence from the German reunification. *Journal of Monetary Economics*, 115:301–320. 15
- Goldin, J. (2015). Optimal tax salience. *Journal of Public Economics*, 131:115–123. 1,

2, 4, 17

- Haaland, I., Roth, C., and Wohlfart, J. (2020). Designing information provision experiments. 15
- Halac, M. and Yared, P. (2014). Fiscal rules and discretion under persistent shocks. *Econometrica*, 82(5):1557–1614. 5
- Harvey, R. P. and Tempalski, J. (1997). The individual AMT: Why it matters. *National Tax Journal*, 50(3):453–473. 16
- Hoopes, J. L., Reck, D. H., and Slemrod, J. (2015). Taxpayer search for information: Implications for rational attention. *American Economic Journal: Economic Policy*, 7(3):177–208. 22
- Jackson, M. O. and Yariv, L. (2014). Present bias and collective dynamic choice in the lab. *American Economic Review*, 104(12):4184–4204. 30
- Jackson, M. O. and Yariv, L. (2015). Collective dynamic choice: the necessity of time inconsistency. *American Economic Journal: Microeconomics*, 7(4):150–78. 30
- Jacobs, B., Jongen, E. L., and Zoutman, F. T. (2017). Revealed social preferences of Dutch political parties. *Journal of Public Economics*, 156:81–100. 29
- Jacquet, L. and Lehmann, E. (2017). Optimal income taxation with composition effects. *Journal of the European Economic Association*. 27
- Kaplow, L. (1994). Accuracy, complexity, and the income tax. *National Bureau of Economic Research Working Paper Series*. 30
- Kleven, H. J. and Kopczuk, W. (2011). Transfer program complexity and the take-up of social benefits. *American Economic Journal: Economic Policy*, 3(1):54–90. 30
- Kostøl, A. R. and Myhre, A. S. (2021). Labor supply responses to learning the tax and benefit schedule. *American Economic Review*, 111(11):3733–66. 6, 24
- Kroft, K., Laliberté, J.-W. P., Leal-Vizcaíno, R., and Notowidigdo, M. J. (2020). Salience and taxation with imperfect competition. Technical report, National Bureau of Economic Research. 21
- Kydland, F. E. and Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–491. 1, 4, 8
- Liebman, J. B. and Zeckhauser, R. J. (2004). Schmeduling. *National Bureau of Economic Research Working Paper Series*. 23
- Lockwood, B. B. and Weinzierl, M. (2016). Positive and normative judgments implicit

- in us tax policy, and the costs of unequal growth and recessions. *Journal of Monetary Economics*, 77:30–47. 29
- Mack III, C., Breaux, J., Frenzel, W. E., Garrett, E., Lazear, E. P., Muris, T. J., Poterba, J. M., Rossotti, C. O., and Sonders, L. A. (2005). Simple, fair, and pro-growth: Proposals to fix America’s tax system. Report of the President’s Advisory Panel on Federal Tax Reform. 30
- Mackowiak, B. A., Matějka, F., and Wiederholt, M. (2021). Rational inattention: A review. *Journal of Economic Literature*. 57
- Malmendier, U. (2021). Fbbva lecture 2020 exposure, experience, and expertise: Why personal histories matter in economics. *Journal of the European Economic Association*, 19(6):2857–2894. 15
- Matějka, F. and Tabellini, G. (2021). Electoral competition with rationally inattentive voters. *Journal of the European Economic Association*, 19(3):1899–1935. 3, 10, 30
- Mill, J. S. (1848). *Principles of political economy*. 4
- Moore, D. T. and Slemrod, J. (2021). Optimal tax systems with endogenous behavioral biases. *Journal of Public Economics*, 197:104384. 4, 17
- Morrison, W. and Taubinsky, D. (2021). Rules of thumb and attention elasticities: Evidence from under-and overreaction to taxes. *Review of Economics and Statistics*, pages 1–45. 1, 4, 21
- Ramsey, F. P. (1927). A contribution to the theory of taxation. *The economic journal*, 37(145):47–61. 1
- Reck, D. (2016). Taxes and mistakes: What’s in a sufficient statistic? *SSRN 2268617*. 10
- Rees-Jones, A. and Taubinsky, D. (2020). Measuring “schmeduling”. *The Review of Economic Studies*, 87(5):2399–2438. 1, 4, 23
- Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68(1):205–229. 12, 23, 24, 28, 61
- Saez, E. and Stantcheva, S. (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics*, 162:120–142. 5
- Schenk, D. H. (2011). Exploiting the salience bias in designing taxes. *Yale J. on Reg.*, 28:253. 3, 16, 17
- Scheuer, F. and Wolitzky, A. (2016). Capital taxation under political constraints. *Amer-*

- ican Economic Review*, 106(8):2304–28. 5
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690. 23, 56
- Sleet, C. (2004). Optimal taxation with private government information. *The Review of Economic Studies*, 71(4):1217–1239. 29
- Taubinsky, D. and Rees-Jones, A. (2018). Attention variation and welfare: theory and evidence from a tax salience experiment. *The Review of Economic Studies*, 85(4):2462–2496. 4, 18, 21
- Tax Policy Center (2023). What is the AMT? *Briefing Book: Key Elements of the U.S. Tax System*. 16
- Taylor, J. B. (2000). Reassessing discretionary fiscal policy. *Journal of economic Perspectives*, 14(3):21–36. 16
- Wagner, R. E. (1976). Revenue structure, fiscal illusion, and budgetary choice. *Public choice*, pages 45–61. 4
- Walsh, C. E. (1995). Optimal contracts for central bankers. *American Economic Review*, pages 150–167. 31

# A Additional tables and figures

Table A.1: Model specifications, calibration, and results

	Model 1	Model 2	Model 3
	<i>Exogenous inattention</i>	<i>Rational inattention</i>	<i>Ironing heuristic</i>
<b>A – Model specifications</b>			
$f_{\vartheta}(\vartheta)$	$\delta(\vartheta)$	$\mathcal{N}(0, \sigma_{\vartheta})$	$\delta(\vartheta)$
$f_{\bar{\tau}}(\tau \tau_0, \hat{q}; w)$	$\delta(\tau - \xi\tau_0 - (1 - \xi)E_{\hat{q}}[\tau])$	$\delta(\tau - \xi\tau_0 - (1 - \xi)E_{\hat{q}}[\tau])$	$\delta(\tau - \xi\tau_0 - (1 - \xi)E_{\hat{q}}[\tau])$
$\Gamma(\tau, \tau_g; w)$	$\delta(\tau - \tau_g - m)$	$\mathcal{N}(\tau_g + m, \sigma_{\vartheta})$	$\delta\left(\tau - \frac{T(y)}{y}\right)$
Eq misperception ( $\varpi$ )	$(1 - \xi)m$	$(1 - \xi)m$	$-(1 - \xi)\frac{R}{y}$
Attention ( $\xi$ )	exogenous	$\max\left(0, 1 + \frac{\kappa}{\sigma_{\vartheta}^2 \int \frac{d^2 y(\tau; w)}{d\tau^2} \hat{q}(\tau; w) d\tau}\right)$	exogenous
<b>B – Calibration</b>			
U.S. tax rate ( $\tau_g^{eq}$ )	29.5%	29.5%	29.5%
Expenditures ( $Exp$ )	17.5% of income	17.5% of income	17.5% of income
Earnings elast ( $e$ )	0.33	0.33	0.33
Avg attention ( $\bar{\xi}$ )	0.55	0.55	<i>0.53</i>
Shock dispersion ( $\sigma_{\vartheta}$ )	–	0.025	–
Information cost ( $\kappa$ )	–	<i>\$2.59 per bit</i>	–
Avg eq misperception ( $\bar{\varpi}$ )	-5.6 p.p.	-5.6 p.p.	-5.6 p.p.
Prior misperception ( $m$ )	<i>-12.4 p.p.</i>	<i>-12.4 p.p.</i>	$-\frac{R}{y}$
Gov't objective ( $\alpha$ )	<i>0.11</i>	<i>0.18</i>	<i>0.11</i>
<b>C – Results</b>			
Optimaleq tax rate ( $\tau_g^*$ )	<i>21.8%</i>	<i>25.8%</i>	<i>25.7%</i>
Tax bias	<i>7.7 p.p.</i>	<i>3.7 p.p.</i>	<i>3.7 p.p.</i>
Welfare cost	<i>\$82 per taxpayer</i>	<i>\$34 per taxpayer</i>	<i>\$13 per taxpayer</i>

NOTE: Panel A provides a summary of model specifications that can be mapped into Propositions 1 and 2 to obtain optimal tax policies – see details in Online Appendix D. Panel B reports calibrated parameters with inferred parameters in italics. Panel C reports the main results.

Table A.2: Sensitivity analysis

Average inattention $\overline{1 - \xi}$	0.35		<b>0.45</b>		0.55	
	Tax bias	Welfare cost	Tax bias	Welfare cost	Tax bias	Welfare cost
Model 1	5.7	\$48	7.7	\$82	9.8	\$128
Model 2	2.7	\$20	3.7	\$34	4.9	\$53
Model 3 <sup>a</sup>	2.2	\$6	3.4	\$12	5.1	\$21
Average misperception $\bar{\omega}$	0		-2.3		<b>-5.6</b>	
	Tax bias	Welfare cost	Tax bias	Welfare cost	Tax bias	Welfare cost
Model 1	8.7	\$122	8.2	\$101	7.7	\$82
Model 2	4.7	\$40	4.2	\$38	3.7	\$34
Model 3 <sup>b</sup>	-	-	1.2	\$2	3.7	\$13
Frictionless elasticity $e$	0.25		<b>0.33</b>		0.50	
	Tax bias	Welfare cost	Tax bias	Welfare cost	Tax bias	Welfare cost
Model 1	7.7	\$62	7.7	\$82	7.7	\$126
Model 2	3.6	\$26	3.7	\$34	3.6	\$48
Model 3	3.6	\$10	3.7	\$13	3.9	\$23
Observed rate $\tau_g^{eq}$	0.25		<b>0.30</b>		0.35	
	Tax bias	Welfare cost	Tax bias	Welfare cost	Tax bias	Welfare cost
Model 1	6.8	\$60	7.7	\$82	8.6	\$114
Model 2	3.6	\$28	3.7	\$34	3.6	\$40
Model 3	1.9	\$4	3.7	\$13	6.6	\$39
Redistributive motive $\alpha$	0.12		0.30		0.99 <sup>c</sup>	
	Tax bias	Welfare cost	Tax bias	Welfare cost	Tax bias	Welfare cost
Model 1	7.9	\$89	10.0	\$194	10.7	\$365
Model 2	3.4	\$25	5.2	\$43	4.6	\$47
Model 3	4.1	\$18	6.1	\$61	6.3	\$118

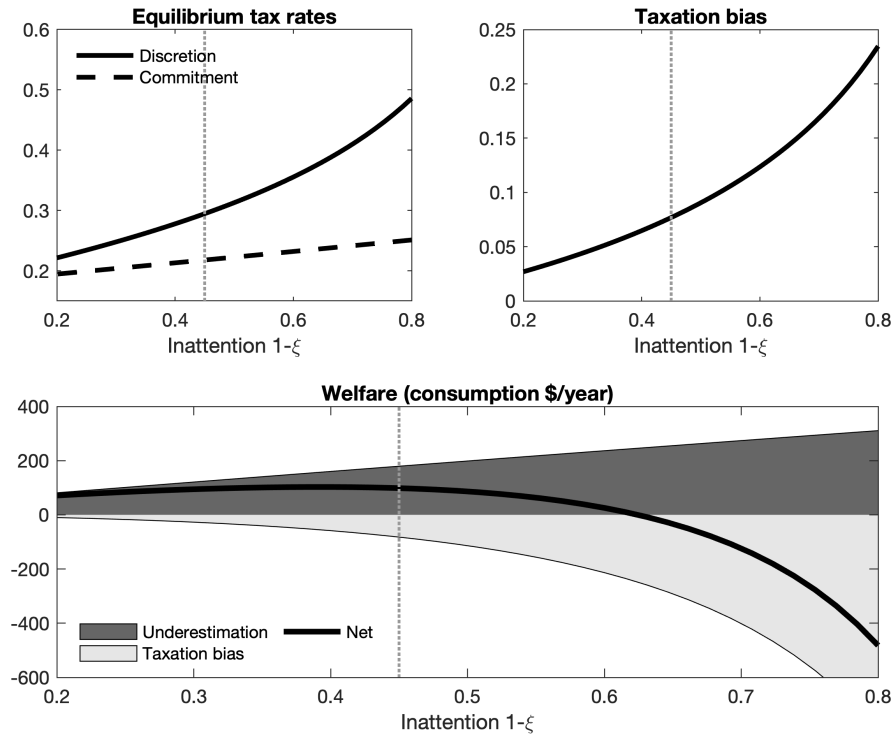
<sup>a</sup>In this case, average misperception adjusts to match average inattention.

<sup>b</sup>In this case, average inattention adjusts to match average misperception.

<sup>c</sup>In this case, the objective of the government approaches a logarithmic social welfare function.

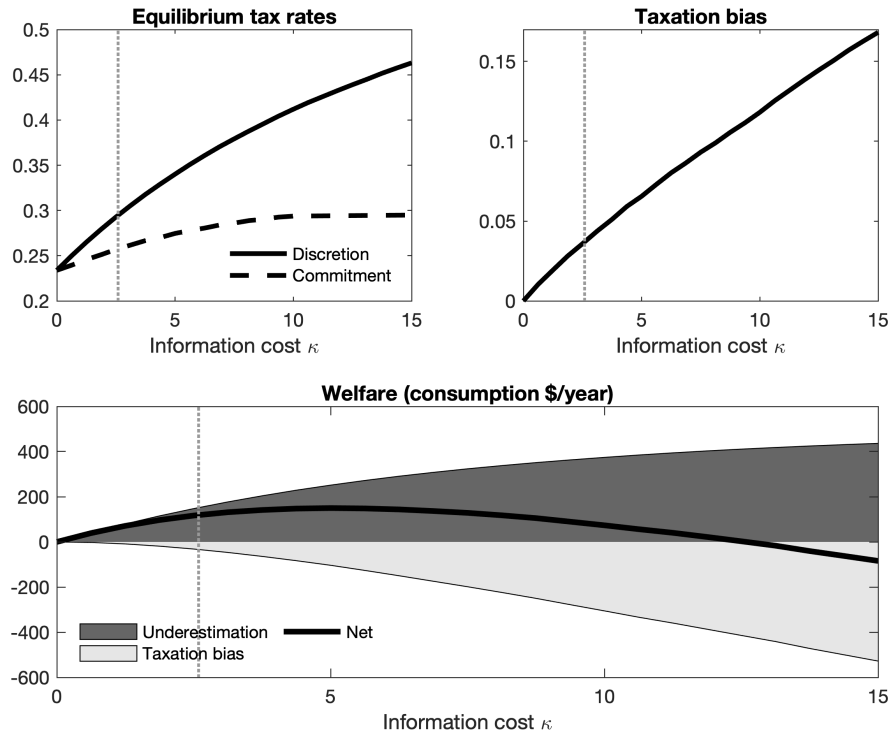
NOTE: This Table presents a sensitivity analysis of the magnitude and welfare cost of the taxation bias following a change in one parameter, keeping other parameters at their benchmark value highlighted in bold. Models are always calibrated such that the observed tax rate is the equilibrium tax rate under discretion, except when we look at changes in the redistributive motive  $\alpha$ , in which case we compare equilibria under discretion and commitment implied by this particular value of  $\alpha$ .

Figure A.1: Tax rates and welfare with exogenous inattention (Model 1)



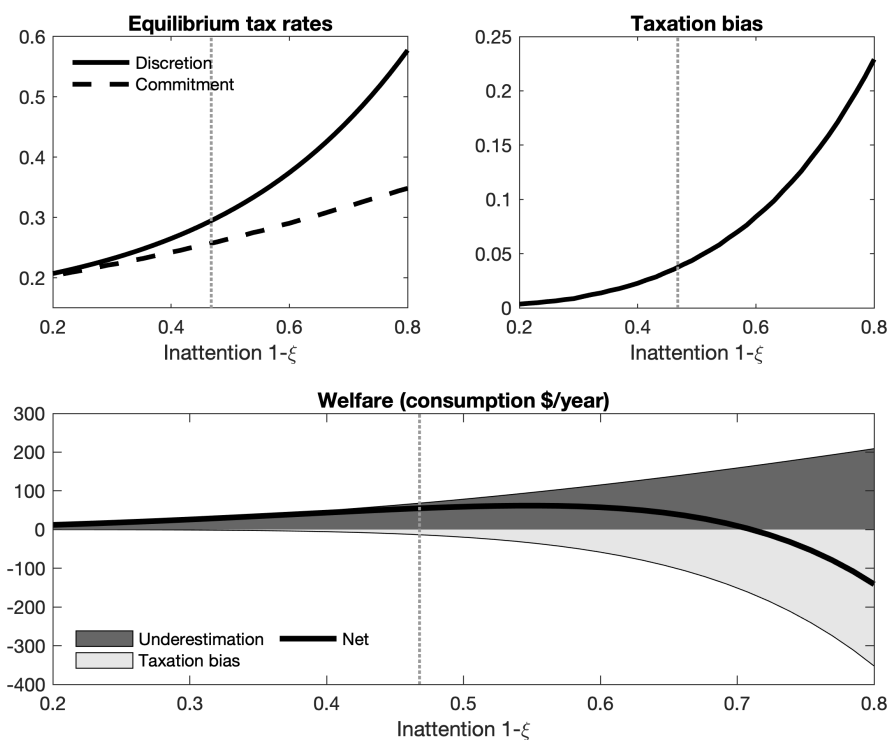
NOTE: The top left panel reports equilibrium tax rates under discretion and under commitment in a calibrated structural Model 1 (exogenous inattention) of the U.S. economy for different values of inattention. The vertical line indicates the baseline value for average inattention,  $\overline{1-\xi} = 0.45$ . The top right panel reports the associated taxation bias. The bottom panel provides a welfare decomposition of the welfare gains due to tax underestimation and welfare loss due to the taxation bias. Features of the model and its calibration are summarized in Table A.1. See Online Appendix D and E for detailed information about the model specifications and simulations respectively.

Figure A.2: Tax rates and welfare with rationally inattentive taxpayers



NOTE: The top left panel reports equilibrium tax rates under discretion and under commitment in a calibrated structural Model 2 (rational inattention) of the U.S. economy for different values of information cost  $\kappa$  expressed in \$/bit/year. The vertical line indicates the baseline value for average inattention under discretion,  $\bar{1} - \bar{\xi} = 0.45$ . The top right panel reports the associated taxation bias. The bottom panel provides a welfare decomposition of the welfare gains due to tax underestimation and welfare loss due to the taxation bias. Features of the model and its calibration are summarized in Table A.1. See Online Appendix D and E for detailed information about the model specifications and simulations respectively.

Figure A.3: Tax rates and welfare with ironers



NOTE: The top left panel reports equilibrium tax rates under discretion and under commitment in a calibrated structural Model 3 (ironing heuristic) of the U.S. economy for different values of inattention. The vertical line indicates the baseline value for average inattention inferred from the calibration,  $\overline{1-\xi} = 0.47$ . The top right panel reports the associated taxation bias. The bottom panel provides a welfare decomposition of the welfare gains due to tax underestimation and welfare loss due to the taxation bias. Features of the model and its calibration are summarized in Table A.1. See Online Appendix D and E for detailed information about the model specifications and simulations respectively.

## B Proofs of main results

### B.1 Proofs for the stylized model of Section 2

Under discretion, the government seeks to maximize tax revenue taking the prior as given. The associated Lagrangian is  $\mathcal{L}(\tau, \lambda) = \tau Y(1 - \xi\tau - (1 - \xi)\hat{\tau}) + \lambda(\tau - 1)$ , with multiplier  $\lambda$  on the constraint that  $\tau \leq 1$ . First order conditions imply that  $\tau = 1$  if and only if  $\hat{\tau} \leq 1 - \frac{\xi}{1-\xi}e$  and  $\tau = \frac{1-(1-\xi)\hat{\tau}}{\xi(1+e)}$  otherwise. These conditions are sufficient since the problem is convex under the assumption that  $\tau Y(1 - \tau)$  is concave. In a rational equilibrium, the prior is correct,  $\hat{\tau} = \tau^{\text{eq}}$ . Guess that the rational equilibrium is interior, then  $\tau^{\text{eq}} = \frac{1}{1+\xi e}$ . Because  $e > 0$ , we have in equilibrium that  $\hat{\tau} > 1 - \frac{\xi}{1-\xi}e$ , which confirms the guess.

Under commitment, the government seeks to maximize tax revenue while accounting for the equilibrium adjustment of prior. Its problem writes  $\max_{\tau} \tau Y(1 - \tilde{\tau})$  subject to the constraints that  $\tilde{\tau} = \xi\tau + (1 - \xi)\hat{\tau}$  and that  $\hat{\tau} = \tau$ . As the constraints imply that perceptions must be correct, solving this problem shows that the policy under commitment coincides with the inverse elasticity rule,  $\tau^* = \tau^r = \frac{1}{1+e}$ .

By definition,  $\tau^* = \tau^r$  maximizes  $\tau Y(1 - \tau)$ , equilibrium tax revenue is thus lower under discretion  $\tau^{\text{eq}} Y(1 - \tau^{\text{eq}}) \leq \tau^* Y(1 - \tau^*)$  and the taxation bias is

$$\tau^{\text{eq}} - \tau^* = \frac{1}{1 + \xi e} - \frac{1}{1 + e} = \frac{1}{1 + e} \frac{1}{1 + \xi e} (1 - \xi)e = \tau^* \frac{\xi e}{1 + \xi e} \frac{1 - \xi}{\xi} = \tau^* (1 - \tau^{\text{eq}}) \frac{1 - \xi}{\xi}.$$

### B.2 Proofs of Propositions 1 and 2

We here prove both propositions at the same time since the only difference between the two problems is in the nature of responses to tax changes that are taken into account. We thus solve the general problem (6)-(7).

The Lagrangian associated with this problem is, with multiplier  $p$ ,

$$\mathcal{L}(\tau_g, R_g, p) = E_{\vartheta} \left[ \iint \left[ G(\mathcal{V}(\tau, \tau_g + \vartheta, R_0; w)) \right. \right. \tag{A.1} \\ \left. \left. + p \left( (\tau_g + \vartheta)y(\tau; w) - R_0 - \text{Exp} \right) \right] f_{\tilde{\tau}}(\tau | \tau_g + \vartheta, \hat{q}; w) f_w(w) d\tau dw \right].$$

The first-order condition associated with the choice of the marginal tax rate,  $\tau_g$ , is

$$\frac{1}{p} \frac{d\mathcal{L}}{d\tau_g} = E_{\vartheta} \left[ \int \left\{ \int \left[ \frac{G'(\mathcal{V})}{p} \frac{d\mathcal{V}}{d\tau_g} + y \right] f_{\tilde{\tau}}(\tau | \tau_g + \vartheta, \hat{q}; w) d\tau \right. \right. \tag{A.2} \\ \left. \left. + \int \left[ \frac{G(\mathcal{V})}{p} + (\tau_g + \vartheta)y - R_0 - \text{Exp} \right] \frac{df_{\tilde{\tau}}(\tau | \tau_g + \vartheta, \hat{q}; w)}{d\tau_g} d\tau \right\} f_w(w) dw \right].$$

By definition,  $\int f_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)d\tau = 1$ , which implies that  $\int \frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g} d\tau = 0$ . Moreover, the quasi-linearity of utility implies that  $\frac{d\mathcal{V}}{d\tau_g} = -y(\tilde{\tau}; w)$ . Therefore, the optimality condition,  $\frac{1}{p} \frac{d\mathcal{L}}{d\tau_g} = 0$ , is

$$E_{\vartheta} \left[ \int \left\{ \int \left[ -\frac{G'(\mathcal{V})}{p} y + y \right] f_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w) d\tau + \int \left[ \frac{G(\mathcal{V})}{p} + (\tau_g + \vartheta)y \right] \frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g} d\tau \right\} f_w(w) dw \right] = 0. \quad (\text{A.3})$$

Under discretion, the government solves problem (6)-(7) taking the prior as given. Hence, replacing  $\frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g}$  by  $\frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}}$  yields condition (8) from Proposition 1 characterizing the discretionary tax rate.

Under commitment, the government solves problem (6)-(7) subject to the additional constraint that,  $\hat{q}(\tau; w) \stackrel{\text{eq}}{=} \Gamma(\tau, \tau_g; w)$ . Plugging this equilibrium relationship in the above and noting that in this case,

$$\frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g} = \left\{ \frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}} + \frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\hat{q}} \frac{d\Gamma(\tau, \tau_g; w)}{d\tau_g} \right\}_{\hat{q} \stackrel{\text{eq}}{=} \Gamma}, \quad (\text{A.4})$$

yields condition (10) from Proposition 2 characterizing the commitment tax rate.

The first-order condition associated with the choice of the demogrant,  $R_g$ , is

$$\frac{1}{p} \frac{d\mathcal{L}}{dR_g} = E_{\vartheta} \left[ \iint \left[ \frac{G'(\mathcal{V})}{p} \frac{d\mathcal{V}}{dR_0} - 1 \right] \frac{dR_0}{dR_g} f_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w) f_w(w) d\tau dw \right], \quad (\text{A.5})$$

By quasi-linearity, we have that  $\frac{d\mathcal{V}}{dR_0} = 1$ , and since the actual transfer,  $R_0$ , is equal to the target transfer,  $R_g$ , plus a term that depends on implementation shocks, we have that  $\frac{dR_0}{dR_g} = 1$ .<sup>IV</sup> Therefore, the optimality condition,  $\frac{1}{p} \frac{d\mathcal{L}}{dR_g} = 0$ , is

$$E_{\vartheta} \left[ \iint \left[ \frac{G'(\mathcal{V})}{p} - 1 \right] f_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w) f_w(w) d\tau dw \right] = 0. \quad (\text{A.6})$$

This gives condition (9) from Proposition 1 and condition (11) from Proposition 2.

### B.3 Proof of Proposition 3

The commitment tax rate,  $\tau_g^*$ , is given by (10) which involves both *direct* and *equilibrium* perception changes,  $\frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g}$ . The discretion tax rate,  $\tau_g^{\text{eq}}$ , is given by (8) which involves only *direct* perception changes,  $\frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}}$ . Now, for all ability,  $w$ ,

$$\frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g} - \frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}} = \left\{ \frac{df_{\tilde{\tau}}(\tau|\tau_g + \vartheta, \hat{q}; w)}{d\hat{q}} \frac{d\Gamma(\tau, \tau_g; w)}{d\tau_g} \right\}_{\hat{q} \stackrel{\text{eq}}{=} \Gamma}. \quad (\text{A.7})$$

<sup>IV</sup>Formally, this holds only approximately given the concavity of earnings decisions: positive and negative implementation shocks do not have symmetric effects on tax revenue and on the level of transfer.

Condition (12) thus implies that the left hand-side of (8) is not 0 when evaluated at the commitment tax rate,  $\tau_g^*$ . The existence and uniqueness of the discretionary equilibrium therefore implies that there exists a taxation bias,  $\tau_g^{\text{eq}} \neq \tau_g^*$ , if and only if (12) holds.

## B.4 Proof of Proposition 4

As highlighted in the main text, our sufficient statistics results extend to multidimensional heterogeneity in individuals characteristics, such as tax perceptions, priors, and attention. Since the proof is exactly the same with unidimensional or multidimensional heterogeneity, we directly provide a proof in this more general setting. Within a productivity level  $w$ , we index this (or these) additional dimension(s) of individual heterogeneity by  $i$ , distributed according to the probability distribution function,  $f_i(i|w)$ .

We retain the assumption that preferences are quasi-linear in consumption but allow for individual differences in the cost of working,  $U_i(c, y; w) = c - k_i(y; w)$ , and thus in the frictionless elasticity of labor supply,  $e_i(w)$ . This gives rise to additional heterogeneity in earnings decisions,  $y_i(\tilde{\tau}_i(w); w)$ , stemming either from heterogeneity in preferences or from heterogeneity in perceived tax rates,  $\tilde{\tau}_i(w)$ . Consumption is then  $c_i(\tilde{\tau}_i(w), \tau_g, R_g; w) = (1 - \tau_g)y_i(\tilde{\tau}_i(w); w) + R_g$  and indirect utility is  $\mathcal{V}_i(\tilde{\tau}_i(w), \tau_g, R_g; w)$ .

**Tax policy.** The Lagrangian of the government problem is, with multiplier  $p$ ,

$$\mathcal{L} = \int_w \int_i \left[ G(\mathcal{V}_i(\tilde{\tau}_i(w), \tau_g, R_g; w)) + p(\tau_g y_i(\tilde{\tau}_i(w); w) - R_g - \text{Exp}) \right] f_i(i|w) di f_w(w) dw. \quad (\text{A.8})$$

Suppressing arguments to economize on notations, a change in the tax rate implies

$$\begin{aligned} \frac{1}{p} \frac{d\mathcal{L}}{d\tau_g} &= \int_w \int_i \left[ \frac{G'(\mathcal{V}_i)}{p} \frac{d\mathcal{V}_i}{d\tau_g} + y_i + \tau_g \frac{dy_i}{d\tau_g} \right] f_i(i|w) di f_w(w) dw \\ &= \int_w \int_i \left[ \left( 1 - \frac{G'(\mathcal{V}_i)}{p} \right) y_i + \left( \tau_g + (\tilde{\tau}_i - \tau_g) \frac{G'(\mathcal{V}_i)}{p} \right) \frac{dy_i}{d\tau_g} \right] f_i(i|w) di f_w(w) dw \end{aligned} \quad (\text{A.9})$$

since

$$\frac{d\mathcal{V}_i}{d\tau_g} = -y_i + (1 - \tau_g) \frac{dy_i}{d\tau_g} - \frac{dk_i}{dy} \frac{dy_i}{d\tau_g} = -y_i + (\tilde{\tau}_i - \tau_g) \frac{dy_i}{d\tau_g} \quad (\text{A.10})$$

where we use the fact that the earnings of an individual  $i$  with productivity  $w$  follow from

$$\max_y (1 - \tilde{\tau}_i(w))y + \tilde{R}_i(w) - k_i(y; w) \quad \implies \quad 1 - \tilde{\tau}_i = \frac{dk_i}{dy}. \quad (\text{A.11})$$

Changes in earnings upon changes in the tax rate are given by

$$\frac{dy_i}{d\tau_g} = \frac{dy_i}{d\tilde{\tau}_i} \frac{d\tilde{\tau}_i}{d\tau_g} = -\frac{y_i}{1-\tilde{\tau}_i} e_i \frac{d\tilde{\tau}_i}{d\tau_g} \quad (\text{A.12})$$

such that, introducing social marginal welfare weights,  $g_i := \frac{G'(\mathcal{V}_i)}{p}$ , and equilibrium misperceptions,  $\varpi_i := \tilde{\tau}_i - \tau_g$ , optimal tax rates defined by  $\frac{d\mathcal{L}}{d\tau_g} = 0$  satisfy in equilibrium

$$\int_w \int_i \left[ (1-g_i)y_i - (\tau_g + \varpi_i g_i) \frac{y_i}{1-\tilde{\tau}_i} e_i \frac{d\tilde{\tau}_i}{d\tau_g} \right] f_i(i|w) di f_w(w) dw = 0. \quad (\text{A.13})$$

Assuming small misperceptions, we use  $\tilde{\tau}_i \simeq \tau_g$  in the denominator, implying that this denominator is no longer individual-specific and we obtain

$$(1-\tau_g) \overline{(1-g_i)y_i} \simeq \tau_g \overline{y_i e_i} \frac{d\tilde{\tau}_i}{d\tau_g} + \overline{\varpi_i g_i y_i e_i} \frac{d\tilde{\tau}_i}{d\tau_g} \quad (\text{A.14})$$

where  $\overline{x_i} := \int_w \int_i x_i(w) f_i(i|w) di f_w(w) dw$ .

The tax perception model,  $\tilde{\tau}_i = \xi_i \tau_g + (1-\xi_i) \hat{\tau}_i$ , implies that under discretion we have  $\frac{d\tilde{\tau}_i}{d\tau_g} \Big|_{\hat{\tau}_i} = \xi_i$ , whereas the equilibrium relationship,  $\tilde{\tau}_i \stackrel{\text{eq}}{=} \tau_g + (1-\xi_i)m_i$ , implies that under commitment we have,  $\frac{d\tilde{\tau}_i}{d\tau_g} \stackrel{\text{eq}}{=} \Xi_i$ , with

$$\Xi_i := 1 + (1-\xi_i) \frac{dm_i}{d\tau_g} - m_i \frac{d\xi_i}{d\tau_g}. \quad (\text{A.15})$$

Hence, we get

$$\tau_g^{\text{eq}} \simeq \frac{\overline{(1-g_i)y_i} - \overline{g_i \varpi_i y_i \xi_i e_i}}{\overline{(1-g_i)y_i} + \overline{y_i \xi_i e_i}}, \quad (\text{A.16})$$

$$\tau_g^{\star} \simeq \frac{\overline{(1-g_i)y_i} - \overline{g_i \varpi_i y_i \Xi_i e_i}}{\overline{(1-g_i)y_i} + \overline{y_i \Xi_i e_i}}. \quad (\text{A.17})$$

**Taxation bias.** We can now compute the difference,

$$\tau_g^{\text{eq}} - \tau_g^{\star} = \frac{1}{\overline{(1-g_i)y_i} + \overline{y_i \xi_i e_i}} \frac{1}{\overline{(1-g_i)y_i} + \overline{y_i \Xi_i e_i}} \quad (\text{A.18})$$

$$\begin{aligned} & \times \left\{ \left[ \overline{(1-g_i)y_i} + \overline{y_i \Xi_i e_i} \right] \left[ \overline{(1-g_i)y_i} - \overline{g_i \varpi_i y_i \xi_i e_i} \right] - \left[ \overline{(1-g_i)y_i} + \overline{y_i \xi_i e_i} \right] \left[ \overline{(1-g_i)y_i} - \overline{g_i \varpi_i y_i \Xi_i e_i} \right] \right\} \\ & = \frac{1}{\overline{(1-g_i)y_i} + \overline{y_i \xi_i e_i}} \frac{1}{\overline{(1-g_i)y_i} + \overline{y_i \Xi_i e_i}} \quad (\text{A.19}) \\ & \times \left\{ \overline{(1-g_i)y_i} \overline{g_i \varpi_i y_i (\Xi_i - \xi_i) e_i} + \overline{(1-g_i)y_i} \overline{y_i (\Xi_i - \xi_i) e_i} + \overline{y_i \xi_i e_i} \overline{g_i \varpi_i y_i \Xi_i e_i} - \overline{y_i \Xi_i e_i} \overline{g_i \varpi_i y_i \xi_i e_i} \right\} \end{aligned}$$

To simplify this expression we use the following approximations relying on the assumption that misperceptions are small. First, we use  $\tau_g^{\text{eq}} \simeq \frac{\overline{(1-g_i)y_i}}{\overline{(1-g_i)y_i} + \overline{y_i \xi_i e_i}}$  and  $\tau_g^{\star} \simeq \frac{\overline{(1-g_i)y_i}}{\overline{(1-g_i)y_i} + \overline{y_i \Xi_i e_i}}$  to

make appear  $\tau_g^{\text{eq}}$  and  $\tau_g^*$  in front of the curly bracket. Second, to cancel out the last two terms in the curly bracket, we use  $\overline{y_i \xi_i e_i} \overline{g_i \varpi_i y_i \Xi_i e_i} \simeq \overline{y_i \Xi_i e_i} \overline{g_i \varpi_i y_i \xi_i e_i}$ . We then obtain

$$\begin{aligned} \tau_g^{\text{eq}} - \tau_g^* &\simeq \frac{\tau_g^{\text{eq}}}{(1-g_i)y_i} \frac{\tau_g^*}{(1-g_i)y_i} \left\{ \overline{(1-g_i)y_i} \overline{y_i(\Xi_i - \xi_i)e_i} + \overline{(1-g_i)y_i} \overline{g_i \varpi_i y_i(\Xi_i - \xi_i)e_i} \right\} \\ &\simeq \tau_g^{\text{eq}} \tau_g^* \frac{\overline{y_i(\Xi_i - \xi_i)e_i} + \overline{g_i \varpi_i y_i(\Xi_i - \xi_i)e_i}}{(1-g_i)y_i} \end{aligned} \quad (\text{A.20})$$

Using again  $\tau_g^{\text{eq}} \simeq \frac{\overline{(1-g_i)y_i}}{(1-g_i)y_i + \overline{y_i \xi_i e_i}}$ , we further simplify this to

$$\frac{\tau_g^{\text{eq}} - \tau_g^*}{\tau_g^*} \simeq \frac{\overline{(1-g_i)y_i}}{(1-g_i)y_i + \overline{y_i \xi_i e_i}} \frac{\overline{y_i(\Xi_i - \xi_i)e_i} + \overline{g_i \varpi_i y_i(\Xi_i - \xi_i)e_i}}{(1-g_i)y_i} \quad (\text{A.21})$$

$$\begin{aligned} &\simeq \frac{\overline{y_i \xi_i e_i}}{(1-g_i)y_i + \overline{y_i \xi_i e_i}} \frac{\overline{y_i(\Xi_i - \xi_i)e_i} + \overline{g_i \varpi_i y_i(\Xi_i - \xi_i)e_i}}{\overline{y_i \xi_i e_i}} \\ &\simeq \left( 1 - \frac{\overline{(1-g_i)y_i}}{(1-g_i)y_i + \overline{y_i \xi_i e_i}} \right) \frac{\overline{y_i(\Xi_i - \xi_i)e_i} + \overline{g_i \varpi_i y_i(\Xi_i - \xi_i)e_i}}{\overline{y_i \xi_i e_i}} \\ &\simeq (1 - \tau_g^{\text{eq}}) \frac{\overline{y_i(\Xi_i - \xi_i)e_i} + \overline{g_i \varpi_i y_i(\Xi_i - \xi_i)e_i}}{\overline{y_i \xi_i e_i}}. \end{aligned} \quad (\text{A.22})$$

This is the multidimensional generalization of the sufficient statistics results presented in Proposition 4, which is obtained by slightly rearranging this expression to make terms  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ , and  $\mathcal{D}$  appear.

## B.5 Proof of Proposition 5

As in the linear case, our sufficient statistics results extend to multidimensional heterogeneity in individuals characteristics, such as tax perceptions, priors, and attention. However, since this further complicates the identification of individuals affected by reforms of nonlinear tax-transfer schedules, we present a proof focusing on unidimensional heterogeneity indexed by earnings ability,  $w$ .

**Tax policy.** The Lagrangian of the government problem is, with multiplier  $p$ ,

$$\mathcal{L} = \int_w \left[ G(\mathcal{V}(\tilde{\tau}(w), T_g(\cdot); w)) + p(T_g(y(\tilde{\tau}(w); w)) - \text{Exp}) \right] f_w(w) dw, \quad (\text{A.23})$$

and we consider a small increase in marginal tax rates,  $\Delta\tau$ , in a small bandwidth of earnings,  $[y^r - \Delta y, y^r]$ .

This reform triggers mechanical fiscal and welfare effects associated with a  $\Delta\tau\Delta y$  lump-sum increase in the tax liability of all individuals with earnings higher than  $y^r$ .

Using the monotonicity condition, this corresponds to individuals with earnings ability  $w \geq w^r$ , where  $y^r = y(\tilde{\tau}(w^r); w^r)$ . These mechanical effects are given by

$$\begin{aligned} dM &= \int_{w \geq w^r} \left[ \frac{G'(\mathcal{V})}{p} \frac{d\mathcal{V}}{dT_g} + 1 \right] \Delta\tau \Delta y f_w(w) dw \\ &= \left[ \int_{y \geq y^r} (1 - g(y)) h_y(y) dy \right] \Delta\tau \Delta y \end{aligned} \quad (\text{A.24})$$

since quasi-linear utility implies that  $\frac{d\mathcal{V}}{dT_g} = -1$  and where  $h_y(y)$  is the probability density function of the earnings distribution.

This reform also affects the tax perceptions of individuals who, given their prior, are attentive to marginal tax rates in the bandwidth of earnings  $[y^r - \Delta y, y^r]$ . Using the monotonicity condition, this corresponds to individuals with earnings ability in  $[\hat{w}^r - \Delta\hat{w}, \hat{w}^r]$ , where  $y^r = y(\hat{\tau}(\hat{w}^r); \hat{w}^r)$  and  $y^r - \Delta y = y(\hat{\tau}(\hat{w}^r - \Delta\hat{w}); \hat{w}^r - \Delta\hat{w})$ . In equilibrium, they choose earnings between  $\hat{y}^r = y(\tilde{\tau}(\hat{w}^r); \hat{w}^r)$  and  $\hat{y}^r - \Delta\hat{y} = y(\tilde{\tau}(\hat{w}^r - \Delta\hat{w}); \hat{w}^r - \Delta\hat{w})$ .

Because these individuals face higher marginal tax rates, they reduce their earnings through substitution effects. This leads to a behavioral fiscal effect given by

$$\begin{aligned} dB_F &= T'_g(\hat{y}^r) \left[ \frac{dy(\tilde{\tau}; \hat{w}^r)}{d\tilde{\tau}} \frac{d\tilde{\tau}}{dT'_g(y^r)} \Delta\tau \right] f_w(\hat{w}^r) \Delta\hat{w} \\ &= -T'_g(\hat{y}^r) \left[ \frac{\hat{y}^r}{1 - \tilde{\tau}} e(\hat{y}^r) \frac{d\tilde{\tau}}{dT'_g(y^r)} \Delta\tau \right] h_y(\hat{y}^r) \Delta\hat{y} \end{aligned} \quad (\text{A.25})$$

since for small reforms,  $f_w(\hat{w}^r) \Delta\hat{w} = h_y(\hat{y}^r) \Delta\hat{y}$ .

These individuals are also incurring a utility change proportional to their misperceptions. This behavioral welfare effect is given by

$$\begin{aligned} dB_W &= \left[ \frac{G'(\mathcal{V})}{p} \frac{d\mathcal{V}}{d\tilde{\tau}} \frac{d\tilde{\tau}}{dT'_g(y^r)} \Delta\tau \right] f_w(\hat{w}^r) \Delta\hat{w} \\ &= - \left[ g(y^r) (\tilde{\tau} - T'_g(\hat{y}^r)) \frac{\hat{y}^r}{1 - \tilde{\tau}} e(\hat{y}^r) \frac{d\tilde{\tau}}{dT'_g(y^r)} \Delta\tau \right] h_y(\hat{y}^r) \Delta\hat{y} \end{aligned} \quad (\text{A.26})$$

since, as in the linear case, individuals' earnings first-order condition implies that we have  $\frac{d\mathcal{V}}{d\tilde{\tau}} = (\tilde{\tau} - T'_g(\hat{y}^r)) \frac{dy(\tilde{\tau}; \hat{w}^r)}{d\tilde{\tau}}$ , where  $\frac{dy(\tilde{\tau}; \hat{w}^r)}{d\tilde{\tau}} = -\frac{\hat{y}^r}{1 - \tilde{\tau}} e(\hat{y}^r)$ .

Characterizing the optimal nonlinear schedule by  $\frac{d\mathcal{L}}{p} = dM + dB_F + dB_W = 0$ , we get that in equilibrium,

$$\frac{T'_g(\hat{y}^r) + g(\hat{y}^r) \varpi(\hat{y}^r)}{1 - \tilde{\tau}} \hat{y}^r e(\hat{y}^r) \frac{d\tilde{\tau}}{dT'_g(y^r)} h_y(\hat{y}^r) \Delta\hat{y} = \left[ \int_{y \geq y^r} (1 - g(y)) h_y(y) dy \right] \Delta y \quad (\text{A.27})$$

where we have introduced equilibrium misperceptions  $\varpi(\hat{y}^r)$  and simplified by  $\Delta\tau$ .

As before, we have  $\frac{d\tilde{\tau}}{dT'_g(y^r)} = \xi(\hat{y}^r)$  under discretion and  $\frac{d\tilde{\tau}}{dT'_g(y^r)} = \Xi(\hat{y}^r)$  under commitment. Assuming misperceptions are small, we can use  $\hat{y}^r \simeq y^r$  as well as  $\Delta y \simeq \Delta \hat{y}$  and  $1 - \tilde{\tau} \simeq 1 - T'_g(y^r)$  to finally obtain that, in equilibrium, the optimal marginal tax rates under discretion and under commitment satisfy respectively

$$T_g^{\text{eq}}(y^r) = \frac{E_{y \geq y^r}[1 - g(y)] - g(y^r)\varpi(y^r)e(y^r)\xi(y^r)y^r h_y(y^r)}{E_{y \geq y^r}[1 - g(y)] + e(y^r)\xi(y^r)y^r h_y(y^r)}, \quad (\text{A.28})$$

$$T_g^{\star'}(y^r) = \frac{E_{y \geq y^r}[1 - g(y)] - g(y^r)\varpi(y^r)e(y^r)\Xi(y^r)y^r h_y(y^r)}{E_{y \geq y^r}[1 - g(y)] + e(y^r)\Xi(y^r)y^r h_y(y^r)}, \quad (\text{A.29})$$

where  $E_{y \geq y^r}[1 - g(y)] := \int_{y \geq y^r} (1 - g(y))h_y(y)dy$ .

**Taxation bias.** Suppressing arguments to economize on notations, the difference is

$$\begin{aligned} T_g^{\text{eq}}(y^r) - T_g^{\star'}(y^r) &= \frac{1}{E[1 - g] + e\xi y^r h_y} \frac{1}{E[1 - g] + e\Xi y^r h_y} \\ &\quad \times \left\{ [E[1 - g] - g\varpi e\xi y^r h_y][E[1 - g] + e\Xi y^r h_y] \right. \\ &\quad \left. - [E[1 - g] + e\xi y^r h_y][E[1 - g] - g\varpi e\Xi y^r h_y] \right\} \\ &= \frac{1}{E[1 - g] + e\xi y^r h_y} \frac{E[1 - g]}{E[1 - g] + e\Xi y^r h_y} (1 + g\varpi)(\Xi - \xi)e y^r h_y \quad (\text{A.30}) \end{aligned}$$

Consistent with the assumption that misperceptions are small, we use  $T_g^{\text{eq}}(y^r) \simeq \frac{E[1-g]}{E[1-g]+e\xi y^r h_y}$  and  $T_g^{\star'}(y^r) \simeq \frac{E[1-g]}{E[1-g]+e\Xi y^r h_y}$  to simplify this expression as

$$\begin{aligned} T_g^{\text{eq}}(y^r) - T_g^{\star'}(y^r) &\simeq T_g^{\star'}(y^r) \frac{e\xi y^r h_y}{E[1 - g] + e\xi y^r h_y} \frac{(1 + g\varpi)(\Xi - \xi)e y^r h_y}{e\xi y^r h_y} \\ &\simeq T_g^{\star'}(y^r) (1 - T_g^{\text{eq}}(y^r)) \frac{(1 + g\varpi)(\Xi - \xi)}{\xi}. \quad (\text{A.31}) \end{aligned}$$

Rearranging this expression to make terms  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ , and  $\mathcal{D}$  appear leads to the general sufficient statistics formula for the taxation bias presented in Proposition 5.

Note that this proof relies on the premise that the monotonicity condition holds in equilibrium, a requirement that we elaborate on in Online Appendix F.

# Online Appendix

(Not for publication)

This online appendix first presents dynamic extensions to the stylized model (Section C). Second, it provides microfoundations for the three structural models used in the empirical application. Third, it gives detailed information on the numerical simulations (Section E). Fourth, it analyzes the assumptions underlying the monotonicity condition with nonlinear tax-transfer schedules (Section F). Last, it shows how to incorporate income effects in the analysis (Section G).

## C Dynamic versions of the stylized model

In this online appendix, we demonstrate that the equilibrium of the static model presented in Section 2 coincides with the unique stable equilibrium of a dynamic model in which taxpayers dynamically learn about taxes. This allows us to show that: (i) the presence of a taxation bias extends to dynamic setups, (ii) the government does not try to systematically fool taxpayers as we converge to a unique stable equilibrium, (iii) long-run sophistication from the government does not eliminate the taxation bias.

### C.1 Dynamic learning equilibrium

**Analysis.** Consider a dynamic environment where in period  $t$ , agents' perceived tax rate is determined through,

$$\tilde{\tau}_t = \xi\tau_t + (1 - \xi)\hat{\tau}_t, \quad (\text{A.32})$$

with  $\xi$  the attention level to the actual tax rate  $\tau_t$ , and  $\hat{\tau}_t$  agents' prior about this tax rate. Taxpayers have memory and enter next period with their posterior as a prior,

$$\hat{\tau}_t = \tilde{\tau}_{t-1}. \quad (\text{A.33})$$

The problem of the government is the same as in Section 2, except that its objective is now to maximize the discounted sum of present and future tax revenues. Assuming that the government disregards the impact of current policies on future beliefs (an assumption we relax below), the problem is equivalent to one where the government sequentially maximizes tax revenues at each period  $t$ . Since taxpayers' prior  $\hat{\tau}_t$  is given at the beginning

of each period, this implies that the first-order condition from the static problem still applies and, focusing on interior solutions, tax policy in period  $t$  is given by

$$\tau_t = \frac{1 - (1 - \xi)\tilde{\tau}_{t-1}}{\xi(1 + e)}. \quad (\text{A.34})$$

The dynamics of the model therefore follows from the two differential equations:

$$\Delta\hat{\tau}_{t+1} = \xi(\tau_t - \hat{\tau}_t) \quad (\text{A.35})$$

$$\Delta\tau_{t+1} = \frac{1}{\xi(1 + e)} - \frac{\xi(2 - \xi + e)}{\xi(1 + e)}\tau_t - \frac{(1 - \xi)^2}{\xi(1 + e)}\hat{\tau}_t \quad (\text{A.36})$$

where (A.35) describes the dynamic evolution of the prior and follows from (A.32) and (A.33), and where (A.36) describes the dynamic evolution of tax policy and follows from (A.32) and (A.34). The loci are

$$\Delta\hat{\tau}_{t+1} = 0 \iff \tau_t = \hat{\tau}_t \quad (\text{A.37})$$

$$\Delta\tau_{t+1} = 0 \iff \tau_t = \frac{1 - (1 - \xi)^2\hat{\tau}_t}{\xi(2 - \xi + e)} \quad (\text{A.38})$$

Plugging the first locus into the second, the steady state tax rate is

$$\tau^{eq} = \frac{1}{1 + \xi e} \quad (\text{A.39})$$

which coincides with the equilibrium of the static stylized model analyzed in Section 2. Drawing a phase diagram using the above loci, one finds that this equilibrium is stable, i.e., that we asymptotically converge to this equilibrium for any initial prior.

**Interpretation.** There are a few things to learn from this dynamic model. First, it provides an intuitive interpretation for the requirement that a discretionary government must take agents' prior as given. Here, priors are based on agents' perception history and are thus independent of the current policy.

Second, it also provides an intuitive interpretation of the equilibrium condition that agents' prior must adjust to actual tax policy. Here, agents continuously learn from (partially) observing the current policy implemented by the government. Doing so, the economy smoothly converges to a steady state where agents' prior coincides to the tax policy implemented by the government, which is equivalent to the definition of a rational equilibrium in the static model.

Third, it should be clear that the government is not trying to systematically fool taxpayers as we converge to a stable point. Systematic fooling would lead to, e.g., systematic fluctuations or growth in tax rates, but not to a unique stable steady state.

Fourth, our characterization of the taxation bias does not implicitly require taxpayers to be sophisticated in the sense that they do not have to anticipate the governments' deviations, and to realize that a taxation bias will occur in equilibrium in order to adjust their prior. Our characterization of the taxation bias may naturally arise as an equilibrium outcome through dynamic learning, that is with naive agents gathering information on tax rates in every period.

## C.2 Dynamic learning equilibrium with long-run sophistication

In the previous dynamic model, we assume the government does not internalize the impact of current policies on future beliefs. We here relax this assumption and consider that the government internalizes taxpayers' dynamic learning and thereby exhibits a form of long-run sophistication. Formally, the government solves

$$\begin{aligned} \max_{\{\tau_t\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t \tau_t Y(1 - \tilde{\tau}_t) \\ \text{s.t.} \quad & \tilde{\tau}_t = \xi \tau_t + (1 - \xi) \tilde{\tau}_{t-1} \end{aligned} \quad (\text{A.40})$$

given its discount factor  $\beta \in [0, 1]$  and an initial condition  $\tilde{\tau}_{-1}$ . This extension of the basic dynamic model presented in the previous Section C.1 is meant to illustrate that even long-run sophistication and further self-imposed constraints will, generally, not be sufficient to resolve the taxation bias.

The current-value Lagrangian associated to the government problem writes

$$\mathcal{L}(\tau_t, \tilde{\tau}_t, \lambda_t) = \sum_{t=0}^{\infty} \beta^t [\tau_t Y(1 - \tilde{\tau}_t) + \lambda_t (\tilde{\tau}_t - \xi \tau_t - (1 - \xi) \tilde{\tau}_{t-1})] \quad (\text{A.41})$$

The first-order conditions are

$$Y(1 - \tilde{\tau}_t) - \lambda_t \xi = 0 \quad (\text{A.42})$$

$$\tau_t \frac{\partial Y(1 - \tilde{\tau}_t)}{\partial \tilde{\tau}_t} + \lambda_t - \beta \lambda_{t+1} (1 - \xi) = 0 \quad (\text{A.43})$$

After a few standard manipulations of these first order conditions, one finds that the optimal taxation path is given by

$$\beta(1 - \xi) Y(1 - \tilde{\tau}_{t+1}) = \left(1 - \frac{\tau_t}{1 - \tilde{\tau}_t} \xi e\right) Y(1 - \tilde{\tau}_t) \quad (\text{A.44})$$

A steady state in this economy corresponds to a situation where the tax policy is constant and the learning has converged. Hence, the steady state tax policy of a discretionary

government is given by

$$\tau^{eq} = \frac{1 - \beta(1 - \xi)}{1 + \xi e - \beta(1 - \xi)} \quad (\text{A.45})$$

Despite the long-run sophistication of the government, this equilibrium policy still generally differs from the socially preferable commitment policy. Indeed, we have

$$\frac{1}{1 + e} \leq \tau^{eq} \leq \frac{1}{1 + \xi e}. \quad (\text{A.46})$$

More precisely, the steady state policy is always larger than the commitment policy such that there exists a positive taxation bias, as long as the government discounts future tax revenues ( $\beta < 1$ ). Moreover, when the government does not value future tax revenues ( $\beta = 0$ ), it behaves as in the previously discussed dynamic model where the government does not internalize taxpayers' dynamic learning.

The steady state taxation bias is given by

$$\tau^{eq} - \tau^* = \frac{e(1 - \beta)(1 - \xi)}{(1 + e)[1 + \xi e - \beta(1 - \xi)]} \quad (\text{A.47})$$

which means that its value is relatively smaller than in the static case. For instance, when  $\xi = 0.75$ ,  $e = 0.33$  and  $\beta = 0.95$ ,<sup>V</sup> the steady state taxation bias would be of only 0.31 percentage points.

However, to interpret the importance of the taxation bias for the economy, one should look at the present value of this permanent deviation. Using a first order Taylor expansion of  $Y(1 - \tau^*)$  around  $\tau^{eq}$  and computing the government present value of all streams of tax revenues, one finds that the present discounted cost of this inefficiency relative to current tax revenue equals

$$\frac{e}{1 - \beta} \frac{\tau^{eq} - \tau^*}{\tau^{eq}(1 - \tau^{eq})} \quad (\text{A.48})$$

Under the aforementioned calibration, it implies that the government would be willing to give up 11% of its current income tax revenue to correct for this long term inefficiency. This value is barely affected when considering alternative values of the discount parameter  $\beta$ . It is, however, sensitive to the attention parameter and would be equal to 17% with an attention parameter  $\xi = 65\%$  and to 6% with an attention parameter  $\xi = 85\%$ .

---

<sup>V</sup> $\beta = 0.95$  corresponds to a standard value in the macroeconomic literature for the discount rate of private agents when a time period refers to a year.

## D Structural models: microfoundation

This appendix provides microfoundation for the three structural models used in Section 5.2 and leading to the functional forms in Table A.1.

### D.1 Earnings-consumption choice

Since we assume in applications that agents hold quasi-linear and iso-elastic preferences,  $U(c, y; \omega) = c - (y/\omega)^{1+\varepsilon}/(1 + \varepsilon)$ , the first-order condition for earnings is

$$y(\tilde{\tau}; \omega) = \omega [\omega (1 - \tilde{\tau})]^{\frac{1}{\varepsilon}} . \quad (\text{A.49})$$

Together with the assumption that consumption costlessly adjusts such that agents exhaust their actual budget, i.e.,  $c(\tilde{\tau}; w) = y(\tilde{\tau}; w) - T(y(\tilde{\tau}; w))$ , this uniquely determines earnings and consumption.

Graphically, in a  $y$ - $c$  diagram, perceptions of the tax schedule determine earnings (tangency condition between indifference curve and perceived budget line), while consumption adjusts to the true budget constraint (intersection with true budget line).

### D.2 Exogenous attention (Model 1)

**Perceptions.** In Model 1, we assume that attention  $\xi$  and prior misperception  $m$  are exogenous and uniform such that tax perceptions, given by equation (13), are

$$\tilde{\tau} = \xi \tau_g + (1 - \xi) \hat{\tau} ,$$

and the equilibrium condition (14) is

$$\hat{\tau} \stackrel{\text{eq}}{=} \tau_g + m ,$$

and we assume away implementation shocks, i.e.,  $f_{\Theta}(\vartheta) = \delta(\vartheta)$ , such that  $\tau_0 = \tau_g$ .

Formally, this means that  $f_{\tilde{\tau}}(\tau|\tau_g, \hat{q}; w) = \delta(\tau - \xi \tau_g - (1 - \xi) E_{\hat{q}}[\tau])$  and that the equilibrium condition is  $\hat{q}(\tau; w) \stackrel{\text{eq}}{=} \Gamma(\tau, \tau_g; w)$  with  $\Gamma(\tau, \tau_g; w) = \delta(\tau - \tau_g - m)$ , where  $\delta(\cdot)$  is the Dirac distribution.

**Tax policies.** Corollary 1 characterizes tax rates under discretion  $\tau_g^{\text{eq}}$  and under commitment  $\tau_g^*$ , following from the application of Propositions 1 and 2 to Model 1.

**Corollary 1.** Let  $\mu = \tau_g + (1 - \xi)m$  be the equilibrium perceived tax rate.

The discretionary equilibrium tax policy  $(\tau_g^{\text{eq}}, R_g^{\text{eq}})$  solves

$$\int \left\{ (1 - g)y + \left( g(1 - \xi)m + \tau_g \right) \frac{dy}{d\tilde{\tau}} \xi \right\} \Big|_{\tilde{\tau}=\mu} dF_w(w) = 0, \quad (\text{A.50})$$

together with the resource constraint (7) and  $\int g|_{\tilde{\tau}=\mu} dF_w(w) = 1$ .

The commitment equilibrium tax policy  $(\tau_g^*, R_g^*)$  solves

$$\int \left\{ (1 - g)y + \left( g(1 - \xi)m + \tau_g \right) \frac{dy}{d\tilde{\tau}} \right\} \Big|_{\tilde{\tau}=\mu} dF_w(w) = 0, \quad (\text{A.51})$$

together with the resource constraint (7) and  $\int g|_{\tilde{\tau}=\mu} dF_w(w) = 1$ .

*Proof.* We have  $f_{\tilde{\tau}}(\tau|\tau_g, \hat{q}; w) = \delta(\tau - \xi\tau_g - (1 - \xi)\hat{\tau})$  and  $\Gamma(\tau, \tau_g; w) = \delta(\tau - \tau_g - m)$ .

Hence, the direct behavioral effects from Proposition 1 are equal to

$$\begin{aligned} & \int \left[ \frac{G(\mathcal{V})}{p} + \tau_g y \right] \frac{df_{\tilde{\tau}}(\tau|\tau_g, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}(\tau; w)} d\tau \\ &= \int \left[ \frac{G(\mathcal{V})}{p} + \tau_g y \right] \frac{d\delta(\tau - \xi\tau_g - (1 - \xi)\hat{\tau})}{d\tau_g} \Big|_{\hat{\tau}} d\tau \\ &= -\xi \int \left[ \frac{G(\mathcal{V})}{p} + \tau_g y \right] \frac{d\delta(\tau - \xi\tau_g - (1 - \xi)\hat{\tau})}{d\tau} \Big|_{\hat{\tau}} d\tau \\ &= \xi \int \frac{d \left[ \frac{G(\mathcal{V})}{p} + \tau_g y \right]}{d\tau} \delta(\tau - \xi\tau_g - (1 - \xi)\hat{\tau}) d\tau \\ &\stackrel{\text{eq}}{=} \xi \int \frac{d \left[ \frac{G(\mathcal{V})}{p} + \tau_g y \right]}{d\tau} \delta(\tau - \xi\tau_g - (1 - \xi)(\tau_g + m)) d\tau \\ &\stackrel{\text{eq}}{=} \xi \left[ g \frac{d\mathcal{V}}{d\tilde{\tau}} + \tau_g \frac{dy}{d\tilde{\tau}} \right] \Big|_{\tilde{\tau}=\tau_g+(1-\xi)m} \\ &\stackrel{\text{eq}}{=} \xi \left[ \frac{dy}{d\tilde{\tau}} (g(1 - \xi)m + \tau_g) \right] \Big|_{\tilde{\tau}=\tau_g+(1-\xi)m}. \end{aligned} \quad (\text{A.52})$$

Where  $\stackrel{\text{eq}}{=}$  uses the equilibrium condition,  $\hat{\tau} \stackrel{\text{eq}}{=} \tau_g + m$ . Plugging this expression into equation (8) characterizing the discretionary tax rate (Proposition 1) gives equation (A.50).

Similarly, the direct & equilibrium behavioral effects from Proposition 1, are equal to

$$\begin{aligned} & \left\{ \int \left[ \frac{G(\mathcal{V})}{p} + \tau_g y \right] \frac{df_{\tilde{\tau}}(\tau|\tau_g, \hat{q}; w)}{d\tau_g} d\tau \right\} \Big|_{\hat{q}(\tau; w) \stackrel{\text{eq}}{=} \Gamma(\tau, \tau^{\text{eq}}; w)} \\ &\stackrel{\text{eq}}{=} \int \left[ \frac{G(\mathcal{V})}{p} + \tau_g y \right] \frac{d\delta(\tau - \xi\tau_g - (1 - \xi)(\tau_g + m))}{d\tau_g} d\tau \\ &\stackrel{\text{eq}}{=} \left[ g \frac{d\mathcal{V}}{d\tilde{\tau}} + \tau_g \frac{dy}{d\tilde{\tau}} \right] \Big|_{\tilde{\tau}=\tau_g+(1-\xi)m} \\ &\stackrel{\text{eq}}{=} \left[ \frac{dy}{d\tilde{\tau}} (g(1 - \xi)m + \tau_g) \right] \Big|_{\tilde{\tau}=\tau_g+(1-\xi)m}. \end{aligned} \quad (\text{A.53})$$

Plugging this expression into equation (10) characterizing the commitment tax rate (Proposition 1) gives equation (A.51).  $\square$

### D.3 Rational inattention (Model 2)

**Attention choice.** Building on the rational inattention literature (Sims, 2003), we adopt a Bayesian learning model with a choice of information (in Gabaix, 2019, classification). Let  $\hat{q}(\tau)$  be the common prior distribution of the tax rate  $\tau$ , which is a random variable from the agent perspective. Given this prior, taxpayers choose how much information to collect about the actual tax rate  $\tau_0$ . This information is modeled as an unbiased signal  $s$  whose precision depends on effort to acquire and process information, i.e., their attention to taxes.

**Assumption 1** (tractable Gaussian learning). *We assume that:*

- *the signal  $s$  is drawn from a Gaussian distribution with mean  $\tau_0$  and variance  $\sigma^2$ ,*
- *agents choose their attention strategies relying on a quadratic approximation of their indirect utility.*

Assumption 1 guarantees tractability. Given a prior  $\hat{q}(\tau)$  and signal realization  $s$ , the posterior distribution  $\tilde{q}(\tau|s; \sigma)$  follows from Bayes' rule. Thus, when the prior is Gaussian with mean  $\hat{\tau}$  and variance  $\hat{\sigma}^2$ , the average perceived marginal tax rate, once the signal is observed, takes a simple form  $\tilde{\tau}(s, \sigma) = \xi(\sigma)s + (1 - \xi(\sigma))\hat{\tau}$ , where  $\xi(\sigma) \equiv \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \sigma^2} \in [0, 1]$  can be interpreted as a measure of attention.

Building on the rational inattention literature (Sims, 2003), the information content transmitted through the signal is measured from the entropy reduction between the prior and the posterior, which we denote as  $\mathcal{I}(\sigma)$ . Intuitively,  $\mathcal{I}(\sigma)$  is a measure of the expected amount of information transmitted through the signal. Information being costly, we posit that taxpayers suffer a utility cost  $\kappa$  per unit of processed information.

The attention strategy of a taxpayer with productivity  $w$  thus results from a trade-off between improved private decisions thanks to more accurate information and the cost of acquiring this information. Lemma 1 characterizes the optimal attention strategy.

**Lemma 1.** *When the prior distribution is Gaussian with mean  $\hat{\tau}$  and variance  $\hat{\sigma}^2$ , i.e.,  $\hat{q}(\tau) = \phi(\tau; \hat{\tau}, \hat{\sigma})$  the optimal attention strategy  $\xi(\kappa, \hat{\sigma}, \hat{\tau}; w)$  is given by*

$$\xi(\kappa, \hat{\sigma}, \hat{\tau}; w) = \max \left( 0, 1 + \frac{\kappa}{\hat{\sigma}^2 \int \frac{\partial^2 y(\tau; w)}{\partial \tau^2} \phi(\tau; \hat{\tau}, \hat{\sigma}) d\tau} \right). \quad (\text{A.54})$$

*Proof.* An agent chooses the signal precision – or equivalently its standard error  $\sigma$  – that

maximizes her expected utility

$$\max_{\sigma} \iint V(\tilde{\tau}(s, \sigma), \tau, R; w) \phi(s; \tau, \sigma) \hat{q}(\tau) ds d\tau - \kappa \mathcal{I}(\sigma) \quad (\text{A.55})$$

where

$$\mathcal{I}(\sigma) \equiv H(\hat{q}(\tau)) - E_{p(s)}[H(\tilde{q}(\tau|s; \sigma))] \quad (\text{A.56})$$

$H(q(\tau)) \equiv -\int q(\tau) \log_2(q(\tau)) d\tau$  is the differential entropy (in bits) of the probability distribution  $q(\tau)$  and  $E_{p(s)}[\cdot]$  is the expectation taken over the marginal distribution of signals  $p(s) \equiv \int \phi(s; \tau, \sigma) \hat{q}(\tau) d\tau$ .

Following Assumption 1, the taxpayer relies on a second order Taylor approximation of her indirect utility around  $\tau_0$  when solving for her attention choice. We have

$$V_{\tau_0}^2(\tilde{\tau}, \tau_0, R_0; w) = V(\tau_0, \tau_0, R_0; w) + (\tilde{\tau} - \tau_0) \frac{\partial V}{\partial \tilde{\tau}} \Big|_{\tilde{\tau}=\tau_0} + \frac{(\tilde{\tau} - \tau_0)^2}{2} \frac{\partial^2 V}{\partial \tilde{\tau}^2} \Big|_{\tilde{\tau}=\tau_0} \quad (\text{A.57})$$

where  $\frac{\partial V}{\partial \tilde{\tau}} \Big|_{\tilde{\tau}=\tau_0} = 0$  and  $\frac{\partial^2 V}{\partial \tilde{\tau}^2} \Big|_{\tilde{\tau}=\tau_0} = \frac{\partial^2 y}{\partial \tilde{\tau}^2}$  from the first order condition of an individual's maximization program. Hence,

$$\iint V_{\tau}^2(\tilde{\tau}, \tau, R; w) \phi(s; \tau, \sigma) \phi(\tau; \hat{\tau}, \hat{\sigma}) ds d\tau = \int \left[ V(\tau, \tau, R; w) + \frac{\tilde{\sigma}^2}{2} \frac{\partial^2 y}{\partial \tilde{\tau}^2} \Big|_{\tilde{\tau}=\tau} \right] \phi(\tau; \hat{\tau}, \hat{\sigma}) d\tau$$

where  $\tilde{\sigma}^2$  is the posterior variance and we are using the fact that with a Gaussian prior and a Gaussian signal, the posterior is also Gaussian. Accordingly, the expected information reduction writes

$$\mathcal{I}(\sigma) = \frac{1}{2} (\log(2\pi e \hat{\sigma}^2) - \log(2\pi e \tilde{\sigma}^2)) = \frac{1}{2} \log \frac{\hat{\sigma}^2}{\tilde{\sigma}^2} \quad (\text{A.58})$$

where  $\frac{1}{2} \log(2\pi e \sigma^2)$  is the differential entropy (in bits) of a Gaussian distribution with variance  $\sigma^2$ . Therefore, in a Gaussian model, problem (A.55) becomes

$$\max_{\tilde{\sigma} \geq \hat{\sigma}} \tilde{\sigma}^2 \int \frac{\partial^2 y}{\partial \tilde{\tau}^2} \Big|_{\tilde{\tau}=\tau} \phi(\tau; \hat{\tau}, \hat{\sigma}) d\tau - \kappa \log \frac{\hat{\sigma}^2}{\tilde{\sigma}^2} \quad (\text{A.59})$$

This problem has been extensively studied in the literature. For instance, a step-by-step derivation of the solution is provided in Mackowiak et al. (2021). At the optimum, the perceived tax rate upon a realization of the signal  $s$ , is  $\tilde{\tau}(s) = \xi s + (1 - \xi) \hat{\tau}$  where  $\xi \in [0, 1]$  is a measure of attention such that

$$\xi = \max \left( 0, 1 + \frac{\kappa}{\hat{\sigma}^2 \int \frac{\partial^2 y}{\partial \tilde{\tau}^2} \Big|_{\tilde{\tau}=\tau} \phi(\tau; \hat{\tau}, \hat{\sigma}) d\tau} \right). \quad (\text{A.60})$$

□

Attention decreases with the information cost  $\kappa$  and the precision of the prior  $1/\hat{\sigma}$ . Moreover, attention increases with the expected responsiveness of labor-supply decisions to changes in perceptions, since this determines the gains from information acquisition. As a result, attention to taxes increases with earnings ability  $w$  and with expected prior tax rates  $\hat{\tau}$ . Intuitively, more productive agents have a greater latitude in their earnings choices, while agents become increasingly responsive as taxes increase.

**Biased Gaussian equilibria.** We consider the following class of equilibria:

**Definition 3** (biased Gaussian equilibria). *We consider situations in which*

- *the implementation shock  $\vartheta$  is a Gaussian white noise with variance  $\sigma_\vartheta^2$ ,*
- *the prior distribution  $\hat{q}(\tau)$  adjusts in equilibrium to the expected distribution of the actual tax rate  $\tau_0$  up to a misperception  $m$  common to all taxpayers,*

*such that the equilibrium condition (c) is  $\hat{q}(\tau) \stackrel{\text{eq}}{=} \phi(\tau; \tau_g + m, \sigma_\vartheta^2)$ .*

The assumption of Gaussian implementation shocks allows us to retain the tractability of the Gaussian learning model, while introducing an exogenous and additive prior misperception  $m$  induces a simple parametrization of misperceptions.

Since the bias  $m$  in the prior is the only source of misperception, attention to taxes determines the degree of misperception in equilibrium. As stated in the text, we disregard the heterogeneity induced by the idiosyncratic signals  $s$  by assuming a representative taxpayer at each skill  $w$  whose perceived tax rate is:<sup>VI</sup>

$$\tilde{\tau} = \int \tilde{\tau}(s) dF(s) = \xi\tau_0 + (1 - \xi)\hat{\tau} \quad (\text{A.61})$$

Hence, the pdf of perceived tax rates is  $f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w) = \delta(\tau - \xi\tau_0 - (1 - \xi)\hat{\tau})$ , where  $\delta(\cdot)$  is the Dirac distribution.

**Tax policies.** Corollary 2 characterizes the tax rates under discretion  $\tau_g^{\text{eq}}$  and under commitment  $\tau_g^*$  in biased Gaussian equilibria, following from the application of Propositions 1 and 2 to the Gaussian setup.

**Corollary 2.** *Let  $\mu = \tau_g + m + \xi(\vartheta - m)$  be the equilibrium perceived rate.*

*The Gaussian discretionary equilibrium tax policy  $(\tau_g^{\text{eq}}, R_g^{\text{eq}})$  solves*

$$E_\vartheta \left[ \int \left\{ (1 - g)y + \left( g(1 - \xi)(m - \vartheta) + \tau_0 \right) \frac{dy}{d\tilde{\tau}} \xi \right\} \Big|_{\tilde{\tau}=\mu} dF_w(w) \right] = 0, \quad (\text{A.62})$$

---

<sup>VI</sup>In the working paper version, we show that this assumption has no first order effect on tax policies.

together with the resource constraint (7) and  $E_{\vartheta}[\int g|_{\tilde{\tau}=\mu} dF_w(w)] = 1$ .

The Gaussian commitment equilibrium tax policy  $(\tau_g^*, R_g^*)$  solves

$$E_{\vartheta} \left[ \int \left\{ (1-g)y + \left( g(1-\xi)(m-\vartheta) + \tau_0 \right) \frac{dy}{d\tilde{\tau}} \left( 1 - \frac{d\xi}{d\tau_g} (b-\vartheta) \right) \right\} \Big|_{\tilde{\tau}=\mu} dF_w(w) \right] = 0, \quad (\text{A.63})$$

together with the resource constraint (7) and  $E_{\vartheta}[\int g|_{\tilde{\tau}=\mu} dF_w(w)] = 1$ .

*Proof.* We have  $f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w) = \delta(\tau - \xi\tau_0 - (1-\xi)\hat{\tau})$  and  $\Gamma(\tau, \tau_g; w) = \phi(\tau; \tau_g + m, \sigma_{\vartheta}^2)$ .

Using these functionals, the direct behavioral effects from Proposition 1 are equal to

$$\begin{aligned} & \int \left[ \frac{G(\mathcal{V})}{p} + \tau_0 y \right] \frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}(\tau; w)} d\tau \\ &= \int \left[ \frac{G(\mathcal{V})}{p} + \tau_0 y \right] \frac{d\delta(\tau - \xi(\tau_g + \vartheta) - (1-\xi)\hat{\tau})}{d\tau_g} \Big|_{\hat{\tau}} d\tau \\ &= -\xi \int \left[ \frac{G(\mathcal{V})}{p} + \tau_0 y \right] \frac{d\delta(\tau - \xi(\tau_g + \vartheta) - (1-\xi)\hat{\tau})}{d\tau} \Big|_{\hat{\tau}} d\tau \\ &\stackrel{\text{eq}}{=} \xi \int \frac{d \left[ \frac{G(\mathcal{V})}{p} + \tau_0 y \right]}{d\tau} \delta(\tau - \xi(\tau_g + \vartheta) - (1-\xi)(\tau_g + m)) d\tau \\ &\stackrel{\text{eq}}{=} \xi \left[ g \frac{d\mathcal{V}}{d\tilde{\tau}} + \tau_0 \frac{dy}{d\tilde{\tau}} \right] \Big|_{\tilde{\tau}=\tau_g+m+\xi(\vartheta-m)} \\ &\stackrel{\text{eq}}{=} \xi \left[ \frac{dy}{d\tilde{\tau}} (g(1-\xi)(m-\vartheta) + \tau_0) \right] \Big|_{\tilde{\tau}=\tau_g+m+\xi(\vartheta-m)} \end{aligned} \quad (\text{A.64})$$

Where  $\stackrel{\text{eq}}{=}$  uses the equilibrium condition  $\hat{\tau} \stackrel{\text{eq}}{=} \tau_g + m$  and where  $\xi$  is also evaluated at its equilibrium value  $\xi(\kappa, \hat{\sigma}, \hat{\tau}; w) \stackrel{\text{eq}}{=} \xi(\kappa, \sigma_{\vartheta}, \tau_g + m; w)$ . Plugging this expression for the direct behavioral effects into equation (8) in Proposition 1 gives equation (A.62).

Similarly, the direct & equilibrium behavioral effects from Proposition 1, simplify to

$$\begin{aligned} & \left\{ \int \left[ \frac{G(\mathcal{V})}{p} + \tau_0 y \right] \frac{df_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w)}{d\tau_g} d\tau \right\} \Big|_{\hat{q}(\tau; w) \stackrel{\text{eq}}{=} \Gamma(\tau, \tau^{eq}; w)} \\ &\stackrel{\text{eq}}{=} \int \left[ \frac{G(\mathcal{V})}{p} + \tau_0 y \right] \frac{d\delta(\tau - \tau_g - m - (1-\xi)(\vartheta - m))}{d\tau_g} d\tau \\ &\stackrel{\text{eq}}{=} \left[ g \frac{d\mathcal{V}}{d\tilde{\tau}} + \tau_0 \frac{dy}{d\tilde{\tau}} \right] \Big|_{\tilde{\tau}=\tau_g+m+\xi(\vartheta-m)} \left( 1 - \frac{d\xi}{d\hat{\tau}} \frac{d(\tau_g + m)}{d\tau_g} (\vartheta - m) \right) \\ &\stackrel{\text{eq}}{=} \left[ \frac{dy}{d\tilde{\tau}} (g(1-\xi)(m-\vartheta) + \tau_0) \right] \Big|_{\tilde{\tau}=\tau_g+m+\xi(\vartheta-m)} \left( 1 - \frac{d\xi}{d\hat{\tau}} (\vartheta - m) \right) \end{aligned} \quad (\text{A.65})$$

where, again, the  $\xi$  and its derivatives are evaluated at the equilibrium  $\xi(\kappa, \hat{\sigma}, \hat{\tau}; w) \stackrel{\text{eq}}{=} \xi(\kappa, \sigma_{\vartheta}, \tau_g + m; w)$ . Plugging this expression for the direct & equilibrium behavioral effects into equation (10) in Proposition 2 gives equation (A.63).  $\square$

Corollary 2 enables identifying the channels through which a taxation bias emerges in biased Gaussian equilibria. Direct behavioral effects in response to tax increases are proportional to attention  $\xi$  in (A.62), whereas direct and equilibrium behavioral effects are proportional to  $1 - \frac{d\xi}{d\tau_g}(m - \vartheta)$  in (A.63) and thus match one-for-one with tax increases up to potential debiasing effects. As a result, equilibrium behavioral effects are proportional to  $(1 - \xi) - \frac{d\xi}{d\tau_g}(m - \vartheta)$  and thus generically non-zero in the presence of inattention, leading to a taxation bias in this environment.

## D.4 Model with ironers (Model 3)

**Perceptions.** Let  $1 - \xi$  denote the share of ironers, which we assume to be uniform across skills. Ironers confuse their marginal and average tax rates,  $\tilde{\tau}^i = T(y)/y$ . The remaining taxpayers perfectly observe their marginal tax rate,  $\tilde{\tau}^r = \tau_0$ . There are no implementation shocks, i.e.,  $f_\Theta(\vartheta) = \delta(\vartheta)$ , such that  $\tau_0 = \tau_g$ . As stated in the text, we assume that there is a representative taxpayer at each skill  $w$  whose perceived marginal tax rate is  $\tilde{\tau} = \xi\tau_g + (1 - \xi)\tilde{\tau}^i$ . Thus, we interpret  $\tilde{\tau}^i$  as being the prior of the representative taxpayer and  $\xi$  the attention to tax reforms affecting  $\tau_g$ . Given these assumptions, it follows that  $f_{\tilde{\tau}}(\tau|\tau_0, \hat{q}; w) = \delta(\tau - \xi\tau_g - (1 - \xi)\tilde{\tau}^i)$ . The equilibrium condition then states that the average tax rate used as a prior should indeed coincide with the average tax rate of the representative taxpayer,  $\tilde{\tau}^i \stackrel{\text{eq}}{=} \tau_g - \frac{R_g}{y}$ , or equivalently,  $\Gamma(\tau, \tau_g; w) = \delta\left(\tau - \tau_g + \frac{R_g}{y}\right)$ .

**Tax policies.** Corollary 3 characterizes the equilibrium tax rates under discretion  $\tau_g^{\text{eq}}$  and under commitment  $\tau_g^*$ , following from the application of Propositions 1 and 2 to this ironing model.

**Corollary 3.** Let  $\mu = \tau_g - (1 - \xi)\frac{R_g}{y}$  be the equilibrium perceived tax rate.

The discretionary equilibrium tax policy  $(\tau_g^{\text{eq}}, R_g^{\text{eq}})$  solves

$$\int \left\{ (1 - g)y + \left[ -g(1 - \xi)\frac{R}{y} + \tau_g \right] \frac{dy}{d\tilde{\tau}} \xi \right\} \Big|_{\tilde{\tau}=\mu} dF_w(w) = 0, \quad (\text{A.66})$$

together with the resource constraint (7) and  $\int g(\tilde{\tau})|_{\tilde{\tau}=\mu} dF_w(w) = 1$ .

The commitment equilibrium tax policy  $(\tau_g^*, R_g^*)$  with ironers solves

$$\int \left\{ (1 - g)y + \left[ -g(1 - \xi)\frac{R}{y} + \tau_g \right] \frac{dy}{d\tilde{\tau}} \left( 1 - (1 - \xi)\frac{d(R/y)}{d\tau_g} \right) \right\} \Big|_{\tilde{\tau}=\mu} dF_w(w) = 0, \quad (\text{A.67})$$

together with the resource constraint (7) and  $\int g(\tilde{\tau})|_{\tilde{\tau}=\mu} dF_w(w) = 1$ .

*Proof.* We have  $f_{\tilde{\tau}}(\tau|\tau_0; w) = \delta(\tau - \xi\tau_g - (1 - \xi)E_{\hat{q}(\tau; w)}[\tau])$  and  $\Gamma(\tau, \tau_g; w) = \delta\left(\tau - \tau_g + \frac{R_g}{y}\right)$ . Plugging these formulas into Propositions 1 and 2, and simplifying the terms as in the proofs of Corollary 1 and 2, one obtains equations (A.66) and (A.67) respectively.  $\square$

## E Numerical simulations

This appendix provides details on the simulations which are implemented using Matlab. We first present the calibration strategy for the skill distribution. Second, we explain how to solve for the optimal linear tax policies with exogenous attention (Model 1), rationally inattentive taxpayers (Model 2) and ironers (Model 3). Third, we explain how we infer the social welfare function. Last, we explain how we simulate nonlinear policies.

### E.1 Skill distribution

Simulations require an exogenous distribution of skills  $f_w(\cdot)$ , which we calibrate from the observed distribution of incomes. We use adjusted gross incomes from the 2016 March CPS data, focusing on singles without dependent children to avoid the complexity of interrelated labor supply decisions within families. We approximate this distribution with a log-normal chosen to match the mean and standard deviation of the observed distribution for incomes below \$200,000. Following Saez (2001), we extend the log-normal distribution with a Pareto tail ( $k = 2$ ) for annual incomes above \$200,000. We then discretize the income distribution using evenly distributed bins over the  $[200; 200,000]$  interval and evenly distributed bins (in ln scale) over the  $[200,000; 4,000,000]$  interval. This allows us to approximate integrals with Riemann sums.

To translate this income distribution into a skill distribution, we invert taxpayers' first-order condition for labor supply given calibrated misperceptions and attention. We first use OECD data on 2016 labor taxes in the U.S. and fit a linear tax rate  $\tau^{obs}$ .<sup>VII</sup> The exogenous expenditure requirement is set to a fixed amount equal to 17.5% of aggregate gross incomes, any remaining amount of tax revenue is then used to finance the transfer  $R^{obs}$ , and is close to \$5,000 under the status-quo. Imposing a quasi-linear utility specifi-

<sup>VII</sup>We here rely on the average tax rates and wedges indicator of the OECD Tax Database. We use the tax rates associated with income taxes (at both federal and state levels) and exclude social security contributions.

cation  $u(c, y; w) = c - (y/w)^{1+\epsilon}/(1+\epsilon)$  and setting  $e = 1/\epsilon$  to its target value (0.33 in the baseline), we can infer the skill distribution from observed income and model-dependent tax perceptions.

In the model with rationally inattentive taxpayers, we obtain a first guess for the skill distribution assuming attention is homogeneous. Using this guess for  $f_w(\cdot)$ , we obtain optimal attention at each skill level and infer the attention cost  $\kappa$  to match the calibrated average attention. We then recompute the skill distribution given updated optimal attention. We iterate on the skill distribution and endogenous attention until convergence. For other models, this additional step is not required.

## E.2 Optimal policies: exogenous attention (Model 1)

Optimal policies with exogenous attention follow from Corollary 1. Given the discretized skill distribution and tax policy  $\tau_g$ , we compute the first order conditions from Corollary 1. We then rely on a nonlinear solver to find the  $\tau_g$  such that the first order conditions equal zero.

## E.3 Optimal policies: rationally inattentive taxpayers (Model 2)

Optimal policies with rationally inattentive taxpayers follow from Corollary 2. We use Gauss-Hermite quadratures for Gaussian integrations. For a given attention cost  $\kappa$  and skill distribution, we guess an optimal tax rate and compute the implied attention at each skill level. Using these values for attention, we use a nonlinear solver to find the tax policy solving Corollary 2. This step gives an updated guess for the optimal tax rate. We iterate until convergence, computing derivatives numerically.

## E.4 Optimal policies: ironers (Model 3)

Optimal policies with ironers follow from Corollary 3. For a given attention and tax rate  $\tau_g$ , we compute perceived tax rates. This requires to find a fixed point solution in the lump-sum transfer  $R$  by iterating on perceptions to find  $R = \tau_g \bar{y} - Exp$  such that perceptions and actual policies are consistent. Then, it is straightforward to compute the FOC in Corollary 3, where the derivative  $d(R/y)/d\tau_g$  is computed numerically in a way which ensures that  $R$  solves the above fixed point. We then rely on a nonlinear solver to

find the  $\tau_g$  such that the first order conditions equal zero.

## E.5 Social welfare

As stated in the text, we assume that the government has a Benthamite social welfare function  $G(\mathcal{V}) = \frac{\mathcal{V}^{1-\alpha}}{1-\alpha}$  where  $\alpha \in (0, 1)$  captures the government taste for redistribution. In each model, the value of  $\alpha$  is set so that the optimal policy under discretion equals  $\tau^{obs}$  for the set of parameters that we calibrate. This equivalent to the procedures described in E.2-E.4 excepted that we now take the value of  $\tau_g^{ea} = \tau^{obs}$  as given and the nonlinear solver is used to return of value for  $\alpha$ . We use this value of  $\alpha$  to compute the welfare weights and consumption-equivalent welfare gains or losses.

## E.6 Optimal policies: nonlinear tax-transfer schedules

Optimal nonlinear schedules follow from the characterization in Appendix, in particular the conditions for optimal marginal tax rates under discretion A.28 and commitment A.29, together with the usual resource constraint and the integral condition on marginal social welfare weights.

We compute the optimal nonlinear schedules in Model 2 in the absence of misperceptions,  $m = 0$ , through the following iterative routine:

1. Start with a guess for the nonlinear tax schedule
2. Compute the attention strategies ( $\forall w$ ) for a given adjusted prior  $\hat{\tau}_w$
3. Compute allocations given attention strategies and tax schedule
4. Solve for the government FOCs at each  $w$  to obtain a new guess
5. Repeat steps 1-4 until convergence.

To maintain the numerical stability of the algorithm we impose a slow adjustment of attention strategies  $\xi$  at each iteration. Indeed, since marginal tax rates are sensitive to attention, one shall avoid large jumps in the attention parameter. The convergence criteria we use is the infinite norm for both marginal tax rates and attention strategies.

## F Monotonicity with nonlinear tax-transfer schedules

With a quasi-linear and iso-elastic separable utility function the first-order condition defining earnings,  $y(\tilde{\tau}(w); w)$ , is

$$\frac{1}{w} \left( \frac{y}{w} \right)^\varepsilon = 1 - \tilde{\tau}(w) \quad (\text{A.68})$$

Differentiating this equation with respect to  $w$  yields

$$\frac{\varepsilon}{w^2} \left( \frac{y}{w} \right)^{\varepsilon-1} \frac{dy(\tilde{\tau}(w); w)}{dw} = \frac{1 + \varepsilon}{w^2} \left( \frac{y}{w} \right)^\varepsilon - \frac{d\tilde{\tau}(w)}{dw} \quad (\text{A.69})$$

Using the equilibrium condition  $\hat{\tau}(w) \stackrel{\text{eq}}{=} T'_g(y(\hat{\tau}(w); w)) + m$ , we have that  $\tilde{\tau}(w) \stackrel{\text{eq}}{=} T'_g(y(\tilde{\tau}(w); w)) + (1 - \xi)m$ , and thus that

$$\frac{d\tilde{\tau}(w)}{dw} = T''_g(y(\tilde{\tau}(w); w)) \frac{dy(\tilde{\tau}(w); w)}{dw} + \frac{d}{dw} [(1 - \xi)m]. \quad (\text{A.70})$$

This implies that:

1. If agents correctly perceive marginal tax rates,  $m = 0$ , the equilibrium condition is  $\hat{\tau}_w \stackrel{\text{eq}}{=} T'_g(y(\hat{\tau}(w); w)) \stackrel{\text{eq}}{=} \tilde{\tau}(w)$  such that  $\frac{dy(\hat{\tau}(w); w)}{dw} \stackrel{\text{eq}}{=} \frac{dy(\tilde{\tau}(w); w)}{dw}$ . As a result, the monotonicity condition boils down in equilibrium to

$$\frac{dy}{dw} \stackrel{\text{eq}}{=} \frac{\frac{1+\varepsilon}{w^2} \left( \frac{y}{w} \right)^\varepsilon}{\frac{\varepsilon}{w^2} \left( \frac{y}{w} \right)^{\varepsilon-1} + T''_g(y)} \geq 0 \iff -T''_g(y) \leq \frac{\varepsilon}{w^2} \left( \frac{y}{w} \right)^{\varepsilon-1} \quad (\text{A.71})$$

2. If agents have small prior misperceptions,  $m \approx 0$ , such that,  $\frac{dy(\hat{\tau}(w); w)}{dw} \approx \frac{dy(\tilde{\tau}(w); w)}{dw}$ , the monotonicity condition becomes in equilibrium,

$$\frac{dy}{dw} \stackrel{\text{eq}}{=} \frac{\frac{1+\varepsilon}{w^2} \left( \frac{y}{w} \right)^\varepsilon - \frac{d}{dw} [(1 - \xi)m]}{\frac{\varepsilon}{w^2} \left( \frac{y}{w} \right)^{\varepsilon-1} + T''_g(y)} \geq 0 \iff \begin{cases} -T''_g(y) \leq \frac{\varepsilon}{w^2} \left( \frac{y}{w} \right)^{\varepsilon-1} \\ \frac{d}{dw} [(1 - \xi)m] \leq \frac{1+\varepsilon}{w^2} \left( \frac{y}{w} \right)^\varepsilon \end{cases} \quad (\text{A.72})$$

where the equivalence comes from the fact that the other case in which we would have  $-T''_g(y) \geq \frac{\varepsilon}{w^2} \left( \frac{y}{w} \right)^{\varepsilon-1}$  is infeasible.

Hence, without misperceptions, the monotonicity condition will hold in equilibrium if the tax function  $T_g(y)$  is sufficiently smooth such that its second derivative is bounded (in absolute value). With misperceptions, the monotonicity condition will hold in equilibrium if, in addition, both inattention and prior misperceptions are sufficiently smooth across types such that across-income variations in the equilibrium misperception in perceived tax rates,  $(1 - \xi)m$ , are also bounded.

## G Extension to income effects

We here illustrate how to incorporate income effects in the general model of Section 3.

In the presence of income effects, both the perceived marginal tax rate,  $\tilde{\tau}$  and the perceived lump-sum transfer,  $\tilde{R}$ , matter for earnings decisions,  $y(\tilde{\tau}, \tilde{R}; w)$ . Given the actual tax rate,  $\tau_0$ , and lump-sum transfer,  $R_0$ , consumption costlessly adjusts such that individuals' budget constraints are binding,  $c(\tilde{\tau}, \tilde{R}, \tau_0, R_0; w) = (1 - \tau_0)y(\tilde{\tau}, \tilde{R}; w) + R_0$ . The associated indirect utility is denoted by  $\mathcal{V}(\tilde{\tau}, \tilde{R}, \tau_0, R_0; w)$ .

We assume the existence of a probability density function,  $\hat{q}(\tau, R; w)$ , capturing priors about the tax rate and the transfer. We also assume the existence of a probability density function,  $f_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)$ , capturing posteriors about the tax rate and the transfer.

As before, the government chooses the target policy,  $(\tau_g, R_g)$ , and the presence of an implementation shock,  $\vartheta$ , implies that the actual tax rate is  $\tau_0 = \tau_g + \vartheta$  and that the actual transfer adjusts to satisfy the government's resource constraint.

The Lagrangian associated with this problem is, with multiplier  $p$ ,

$$\begin{aligned} \mathcal{L}(\tau_g, R_g, p) = E_\vartheta \left[ \iiint \left[ G(\mathcal{V}(\tilde{\tau}, \tilde{R}, \tau_0, R_0; w)) \right. \right. \\ \left. \left. + p((\tau_0)y(\tau, R; w) - R_0 - Exp) \right] f_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w) f_w(w) d\tau dR dw \right]. \end{aligned} \quad (\text{A.73})$$

The first-order condition associated with the choice of the marginal tax rate,  $\tau_g$ , is

$$\begin{aligned} \frac{1}{p} \frac{d\mathcal{L}}{d\tau_g} = E_\vartheta \left[ \int \left\{ \iint \left[ \frac{G'(\mathcal{V})}{p} \frac{d\mathcal{V}}{d\tau_g} + y \right] f_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w) d\tau dR \right. \right. \\ \left. \left. + \iint \left[ \frac{G(\mathcal{V})}{p} + \tau_0 y - R_0 - Exp \right] \frac{df_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)}{d\tau_g} d\tau dR \right\} f_w(w) dw \right]. \end{aligned} \quad (\text{A.74})$$

By definition,  $\iint f_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w) d\tau dR = 1$ , thus  $\int \frac{df_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)}{d\tau_g} d\tau dR = 0$ . Moreover, the quasi-linearity of utility implies that  $\frac{d\mathcal{V}}{d\tau_g} = -y(\tilde{\tau}, \tilde{R}; w)$ . Therefore, the optimality condition,  $\frac{1}{p} \frac{d\mathcal{L}}{d\tau_g} = 0$ , is

$$\begin{aligned} E_\vartheta \left[ \int \left\{ \iint \left[ -\frac{G'(\mathcal{V})}{p} y + y \right] f_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w) d\tau dR \right. \right. \\ \left. \left. + \iint \left[ \frac{G(\mathcal{V})}{p} + \tau_0 y \right] \frac{df_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)}{d\tau_g} d\tau dR \right\} f_w(w) dw \right] = 0. \end{aligned} \quad (\text{A.75})$$

Under discretion, the government takes the prior as given. Hence, replacing in the above  $\frac{df_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)}{d\tau_g}$  by  $\frac{df_{\tilde{\tau}, \tilde{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}}$  yields a characterization of the discretionary tax rate in the presence of income effects.

Under commitment, the government takes into account the additional constraint that,  $\hat{q}(\tau, R; w) \stackrel{\text{eq}}{=} \Gamma(\tau, R, \tau_g, R_g; w)$ , where we here extend the equilibrium condition (c) in Definition 1 to this setting with income effects. Plugging this equilibrium relationship in the above and noting that in this case,

$$\begin{aligned} & \frac{df_{\hat{\tau}, \hat{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)}{d\tau_g} & (A.76) \\ & = \left\{ \frac{df_{\hat{\tau}, \hat{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)}{d\tau_g} \Big|_{\hat{q}} + \frac{df_{\hat{\tau}, \hat{R}}(\tau, R | \tau_0, R_0, \hat{q}; w)}{d\hat{q}} \frac{d\Gamma(\tau, R, \tau_g, R_g; w)}{d\tau_g} \right\}_{\hat{q} \stackrel{\text{eq}}{=} \Gamma} , \end{aligned}$$

yields a characterization of the commitment tax rate in the presence of income effects.

These characterizations of optimal tax rates,  $\tau_g$ , in the presence of income effects are conceptually similar to those obtained in the absence of income effects ( and the same is true for optimal transfers,  $R_g$ ). As a result, the existence of a taxation bias similarly depends on the existence of equilibrium (i.e., long-run) responses to tax reforms.