Inattention and the Taxation Bias

Jérémy Boccanfuso (University of Bologna)
Antoine Ferey (LMU Munich)

Discussion Paper No. 323

April 13, 2022

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Inattention and the Taxation Bias*

Jérémy Boccanfuso[†] Antoine Ferey[‡]

First version: October 2019
This version: August 2021

Abstract

This paper shows that agent inattention to taxes generates a time-inconsistency problem in the choice of tax policy. In equilibrium, inattention leads to inefficiently high tax rates and a taxation bias emerges. Combining structural and sufficient statistics approaches, we quantify the magnitude and the welfare effects of this policy distortion for US income tax rates, and find that the taxation bias is large, alters the progressivity of income taxes, and significantly reduces social welfare. Overall, our findings shed new light on the policy and welfare implications of inattention and misperceptions.

Keywords: Optimal taxation; inattention.
JEL classification code: H21; D03.

*We are grateful to Pierre Boyer, Philippe Choné, Olivier Compte, Allan Drazen, Xavier Gabaix, Ulrich Glogowsky, Andreas Haufler, Bas Jacobs, 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 wish to thank audiences at conferences (AFSE, EEA, 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) and the Deutsche Forschungsgemeinschaft through CRC TRR 190 (LMU Munich). Ferey is also grateful for the hospitality of UC Berkeley. We thank Brian Bloch for editing. This paper was awarded the *ITAX Best PhD Paper Award* at the 2019 IIPF Conference.

[†]University of Bologna, Piazza Scaravilli 2, 40126 Bologna. E-mail: jeremy.boccanfuso@unibo.it
[‡]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). It is thus usually interpreted as reducing tax-induced distortions in agents' behavior, thereby improving economic efficiency and warranting higher optimal tax rates (Goldin, 2015; Taubinsky and Rees-Jones, 2018; Farhi and Gabaix, 2020).

Yet, recent evidence shows that, far from ignoring taxes, inattentive agents tend to 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 agents' tax perceptions, and thus their behavior, are shaped both by their attention to taxes and by the priors on which they rely. As a result, inattention reduces individual responsiveness to taxes by anchoring taxpayer perceptions on their priors. The effects of inattention on economic efficiency and optimal tax rates are thus ambiguous, as they depend on these priors.

Building on this insight, the present paper analyzes the interplay between agent inattention to taxes and the government's choice of tax policy. The key take-away is that the anchoring effect of inattention creates a time-inconsistency problem in the choice of tax policy. In equilibrium, this leads to inefficiently high tax rates and a *taxation bias* emerges. This finding highlights that inattention can actually reduce economic efficiency, and echoes empirical work showing that reductions in tax salience are accompanied by increases in tax rates, even in settings where optimal tax rates are unlikely to increase (Finkelstein, 2009; Cabral and Hoxby, 2012).[1] Overall, our analysis sheds new light on the policy and welfare implications of inattention and misperceptions.

[1] Cabral and Hoxby (2012) analyze the impact of property tax salience on property tax rates through variations in tax escrow. Because tax escrow is unlikely to affect the demand for property, optimal property tax rates are unlikely to be affected, such that the higher observed tax rates provides direct evidence of a taxation bias. Finkelstein (2009) studies the adoption of electronic tolls in a setting which is a priori not immune to changes in optimal tax rates. However, her findings probably reflect a taxation bias, because we show that optimal tax rates should actually *decrease* rather than *increase* in this setting.

Our first and main contribution is to show that inattention induces a taxation bias. The analysis begins with a stylized, yet insightful, model (Section 2). We consider a representative agent economy in which a revenue-maximizing government levies a linear tax rate that negatively affects labor supply. Agents' perceived tax rate is given by a weighted average between the actual tax rate and agents' prior, where the weight attached to the former measures attention devoted to taxes (Gabaix, 2019). To build the intuition, we assume that agents' prior coincides with the actual tax rate. This implies that agents correctly perceive the tax rate. Consequently, the optimal policy is unaffected by taxpayer inattention and follows a standard inverse elasticity rule (Dupuit, 1844; Ramsey, 1927). Yet, we show that the equilibrium policy follows a modified inverse elasticity rule, where the elasticity of labor supply is scaled by agent attention to taxes. When agents are not fully attentive, the equilibrium tax rate is thus above the optimal revenue-maximizing tax rate and the government is unable to reach the top of the Laffer curve.

The intuition behind this result is that inattention anchors agents' tax perceptions on their prior. This creates an incentive for the government to implement policy deviations that taxpayers are going to partially ignore. However, because the prior adjusts in equilibrium, these policy deviations turn out to be inefficient ex post.[2] In other words, the anchoring effect of inattention makes the government's policy time-inconsistent and leads to discretionary policymaking in the absence of a credible commitment technology. Formally, the taxation bias is defined as the difference between the tax rate under discretion and that under commitment. We find that it is proportional to agent inattention.

Turning to a general framework, we derive necessary and sufficient conditions for the existence of a taxation bias (Section 3). We consider a Ramsey problem in which a welfare-maximizing government levies a linear tax rate to redistribute resources across agents with heterogeneous earnings skills. Agents' labor supply depends on their tax perceptions captured through (skill-specific probabilistic) posteriors which can be any function of actual tax policy and agent priors, thereby nesting a wide range of perception models and allowing for misperceptions in equilibrium. We show that a taxation bias arises whenever (i) agent prior influences their posterior, and (ii) actual tax policy influences agent prior

[2] We rely on a static model and therefore refer to an equilibrium adjustment. We show in the Appendix that our static model can be interpreted as the steady state of a dynamic model in which taxpayers dynamically learn about the tax rate.

in equilibrium. Importantly, these conditions do not hinge on the government's objective nor on taxpayers' heterogeneity, but on agent perception formation. Reviewing empirical evidence on perception formation, we conclude that these conditions are likely to hold in practice, which supports the existence of a taxation bias.

Our second contribution is to combine structural and sufficient statistics methods to show how the magnitude of the taxation bias can be measured empirically, and to estimate this magnitude for income taxes in the US (Section 4). To this end, we develop a tractable and microfounded model in which agents' tax perceptions result from a Bayesian learning model with a choice of information (Sims, 2003; Gabaix, 2019; Mackowiak et al., 2021).[3] That is, taxpayers are endowed with a prior about the tax policy and endogenously choose how attentive they are to taxes. We allow the prior to be systematically biased to account, for instance, for the underestimation of marginal tax rates (Rees-Jones and Taubinsky, 2020). In particular, this approach captures the use of biased rules-of-thumb as default, while allowing taxpayers to improve their tax perceptions if they find it optimal to do so (Morrison and Taubinsky, 2021).

Deriving a sufficient statistics formula for the taxation bias, we show that the relevant sufficient statistics are the actual income tax rate, the elasticity of earnings with respect to the perceived marginal net-of-tax rate, the government's taste for redistribution, and the income-weighted average attention. Combining existing elasticity and attention estimates from the empirical literature with our endogenous attention model, we find that the taxation bias induces a 6 percentage point increase in US income tax rates. Since US income taxes can be well approximated by a linear tax rate of 29.5 percentage points, this corresponds to a 20 percent increase in relative terms, suggesting that the taxation bias is large.

While this estimate is sensitive to the value of attention, it is unaffected by systematic biases in agent prior, if any. This emphasizes that whether agents misperceive tax rates is (to a first order) irrelevant for the magnitude of the taxation bias. The reason is that misperceptions similarly affect tax rates under discretion and commitment, thereby leaving the difference between the two (largely) unaffected. However, accounting for

[3] We adopt a rational inattention framework, given the strong empirical support for the endogeneity of attention. For instance, Hoopes et al. (2015) find that tax deadlines or particular news events drive taxpayers' online information searching, while Taubinsky and Rees-Jones (2018) show experimentally that tripling the tax rate nearly doubles agents' attention to taxes.

misperceptions turns out to be crucial for welfare.

Our third contribution is that of studying the welfare implications of inattention, accounting both for the taxation bias and potential misperceptions (Section 5). On the one hand, taxpayer inattention leads to the emergence of a taxation bias. This decreases economic efficiency and generates welfare losses by creating a distortion in tax policy. On the other hand, taxpayer inattention tends to exacerbate misperceptions. If these misperceptions consist of agents underestimating their marginal income tax rates, inattention improves economic efficiency and generates welfare gains by reducing tax-induced distortions (Liebman and Zeckhauser, 2004; Rees-Jones and Taubinsky, 2020). As a result, the net welfare effect of inattention is ambiguous, and it can be positive if inattention is associated with relatively large underestimation.

In order to gauge the sign and magnitude of the net welfare effect for US income taxes, we rely on the work of Rees-Jones and Taubinsky (2020) who estimate that agents underestimate their marginal income tax rates by 19 percent on average. Combined with existing attention estimates, this implies that the net welfare effect of inattention is positive. More precisely, inferring the government's taste for redistribution from current taxes, we find that the net increase in welfare due to inattention is equivalent to a \$125 increase in annual consumption per taxpayer. Without a taxation bias, these welfare gains would be 22 percent larger. This estimate of the welfare losses attributable to the taxation bias is conservative, as we assume that each dollar collected by the government through higher tax rates is fully redistributed to taxpayers. As a result, the large efficiency costs associated with the taxation bias are in part offset by equity gains in our simulations.

Fourth, we extend the analysis to nonlinear tax schedules (Mirrlees, 1971; Saez, 2001), and show that the taxation bias also affects the progressivity of income taxes (Section 6). Indeed, discretionary increases in the marginal tax rate at a given income depend on the degree of attention of agents at (or close to) this income level. As a result, the positive correlation between income and attention leads to larger increases in marginal tax rates at low income levels. Importantly, this distortion in the progressivity of the tax schedule increases the welfare losses associated with the taxation bias by 135 percent.[4]

[4]This distortion in the shape of nonlinear income tax schedules also has potentially important implications for the inverse-optimum literature which aims at inferring the government's objective function from the progressivity of current tax schedules (Bourguignon and Spadaro, 2012; Lockwood and Weinzierl, 2016; Jacobs et al., 2017).

Last, we discuss the role of information asymmetries, tax complexity and competitive politics in the context of 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 increase agent attention to taxes, but they are unlikely to alleviate the time-inconsistency problem faced by the government. Consequently, it seems that existing institutions fail to offer a credible commitment technology and that solutions to the taxation bias would require major institutional changes.

This paper builds a bridge between the field of behavioral public economics (Bernheim and Taubinsky, 2018), fiscal illusion literature (Buchanan, 1967) and the literature on the discretionary nature of policymaking (Kydland and Prescott, 1977).

It adds to the growing field of behavioral public economics by showing that agents' inattention to taxes creates a time-inconsistency problem that leads to significant distortions in tax policies. Previous papers analyzing tax policy with behavioral agents all adopt a normative stance, implicitly characterizing tax policy under commitment (Goldin, 2015; Gerritsen, 2016; Allcott et al., 2018; Farhi and Gabaix, 2020; Moore and Slemrod, 2021). Incidentally, issues of time-inconsistency do not arise in these papers, because they treat agents' tax perceptions as a sufficient statistic and do not model the channels through which tax perceptions adjust. Building on recent empirical evidence (Gabaix, 2019; Rees-Jones and Taubinsky, 2020; Morrison and Taubinsky, 2021), we explicitly model tax perception formation and distinguish between agent attention to taxes and agent priors about taxes. The anchoring effect of inattention then naturally gives rise to a taxation bias. While our focus is on income taxes, a taxation bias can arise with any tax instrument. For instance, in subsequent work focusing on consumption taxes, Furukawa (2021) studies the emergence of a taxation bias through a deadweight loss analysis.[5]

The concept of fiscal illusion 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 more than would

[5]Furukawa (2021) considers a Bayesian learning model in a representative agent setting. He assumes that agents' demand for consumption is linear, which implies that the taxation bias no longer depends on agents' tax elasticity. Moreover, he does not consider potential misperceptions in equilibrium, which leads to the conclusion that eliminating inattention is always desirable.

be socially desirable (Buchanan, 1967; Oates et al., 1985). We add to this literature by showing that agents' inattention to taxes creates a taxation bias by itself. This implies that a welfarist government, that is neither inefficient nor self-aggrandizing, grows beyond the optimum whenever taxpayers are not fully attentive to taxes. 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).

Following the seminal paper 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). Yet, our analysis reveals that the anchoring effect of inattention provides a natural justification for the timing assumption which is central to this literature. This suggests that commitment issues could arise in virtually any setting where inattention and information frictions have been documented (Coibion and Gorodnichenko, 2015; Gabaix, 2019). Hence, the insight that agents' inattention leads to time-inconsistency issues probably extends to a wide variety of settings in which 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 setting in which we provide necessary and sufficient conditions for the emergence of taxation bias. Section 4 presents a structural calibrated model and derives a sufficient statistics formula for the taxation bias that we take to data. Section 5 turns to the welfare implications and Section 6 provides an extension to nonlinear tax 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

Canonical problem. Consider the canonical labor income tax problem in which the government sets a linear tax rate τ to maximize tax revenue. Agents' labor supply decisions are negatively affected by the tax and determine aggregate earnings $Y(1-\tau)$.

The tax revenue function is $\tau Y(1-\tau)$, has an inverted U-shape and is nil when the tax rate is equal to 0 or 100%. As is well-known (e.g. Piketty and Saez, 2013), the revenue-maximizing tax rate follows an inverse elasticity rule and is equal to

$$\tau^r = \frac{1}{1+e} \tag{1}$$

where e is the elasticity of aggregate earnings with respect to the net-of-tax rate. Intuitively, τ^r is such that the positive mechanical effect $M^r = Y(1-\tau^r)\, d\tau$ from a small increase in the tax rate $d\tau$ balances out the negative fiscal externality $F^r = \tau^r \frac{Y(1-\tau^r)}{1-\tau^r} e\, d\tau$ induced by the reduction in earnings.

Perceptions and inattention. Assume now that taxpayers are unable to perfectly observe the tax rate because of information frictions. They must nonetheless form an estimate of the latter 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 common prior $\hat{\tau}$ and the actual tax rate τ

$$\tilde{\tau} = \xi\tau + (1-\xi)\hat{\tau}, \tag{2}$$

where the weight $\xi \in [0,1]$ is a measure of taxpayer attention to taxes.[6] Indeed, when $\xi = 1$, taxpayers perfectly observe the actual tax rate whereas, when $\xi = 0$, they are completely inattentive and fully anchor their perception on their prior. Since earnings choices depend on the perceived tax rate, aggregate earnings now write $Y(1-\tilde{\tau})$.

Choice of tax policy. The government sets its tax policy taking agents' prior and attention into account. Formally, it maximizes tax revenue $\tau Y(1-\tilde{\tau})$ subject to the tax perception model (2), which remains a concave problem with respect to the actual tax rate τ. This leads to a policy function $\tau(\xi,\hat{\tau})$ that is decreasing in taxpayers' prior $\hat{\tau}$ and attention parameter ξ

$$\tau(\xi,\hat{\tau}) = \begin{cases} \frac{1-(1-\xi)\hat{\tau}}{\xi(1+e)} & \text{if } \hat{\tau} \geq 1 - \frac{\xi}{1-\xi} e \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

where e is the elasticity of aggregate earnings with respect to the *perceived* net-of-tax rate.[7] The solution is interior when agent attention or prior are high enough, otherwise

[6] Gabaix (2019) argues this is a unifying reduced-form framework that encompasses many behavioral biases and attention theories.

[7] This corresponds to the structural elasticity parameter pinned down by agents' preferences.

the government chooses to impose a 100% tax. Tax policy only coincides with the inverse elasticity rule when agents are fully attentive to taxes ($\xi = 1$).

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 $d\tau$ generates the same positive mechanical effect as before $M = M^r$, but a different negative fiscal externality $F = \xi F^r$. Indeed, the reduction in earnings now depends on the change in the *perceived* tax rate equal to $d\tilde{\tau} = \xi d\tau$, since inattentive taxpayers only observe a fraction ξ of the tax increase. In other words, inattention reduces agent responsiveness to tax changes, because it anchors agents' perceptions on their prior. As a result, when agents are not fully attentive, the mechanical effect outweighs the fiscal externality and the government systematically implements a higher tax rate.

Figure 1 plots tax policy $\tau(\xi, \hat{\tau})$ as a function of agents' prior for different attention levels, showing that small information frictions generate notable deviations in tax policy. If agents' prior is that the government implements the inverse elasticity rule – $\tau^r = 75\%$ assuming $e = 0.33$ – the government chooses a 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).[8]

[Figure 1 about here]

Equilibrium. Yet, 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. To build the intuition, we focus here on rational equilibria in which agents' prior is correct $\hat{\tau} = \tau^{\text{eq}}$, and defer to Section 3, the introduction of other equilibria (e.g. with misperceptions).

Plugging this equilibrium condition into the policy function (3), we show that the equilibrium tax policy follows a modified inverse elasticity rule, in which the elasticity parameter is scaled by agents' attention to taxes:

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

[8] 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.

Graphically, the equilibrium is represented by the point at which the 45-degree line ($\hat{\tau} = \tau^{\text{eq}}$) intersects the policy function (3). This reveals that small information rigidities lead to large changes in actual policy. 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. Because we consider a rational equilibrium in which agents' perceptions are correct, the Laffer curve is unaffected by taxpayers' inattention in equilibrium. Hence, the optimal equilibrium policy is the standard inverse elasticity rule, $\tau^{\star} = \tau^{r}$ (point E). As this would be the policy implemented by a social planner who can credibly commit to its tax policy, we refer to $\tau^{\star}$ as the optimal policy under commitment and it is, by definition, the optimal equilibrium policy in the presence of information frictions.

However, we have seen that in the presence of inattention, the government systematically deviates from the standard inverse elasticity rule and that the equilibrium policy τ^{eq} is instead characterized by a modified inverse elasticity rule. This is because the optimal policy $\tau^{\star}$ is time-inconsistent; even though the government would prefer to implement this policy ex ante, it will systematically deviate ex post. This time-inconsistency problem in the choice of tax policy arises because of the anchoring effect of inattention and restricts the set of implementable policies in equilibrium.[9] As a result, the government is unable to reach the top of the Laffer curve and the equilibrium tax rate is inefficiently high.

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

$$\tau^{\text{eq}} - \tau^{\star} \quad = \quad \frac{(1-\xi)e}{(1+\xi e)(1+e)} \geq 0 \tag{5}$$

implying that the taxation bias is strictly positive when agents are not fully attentive to taxes ($\xi < 1$). Moreover, the (absolute) size of the taxation bias increases with agents

[9] 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).

inattention $1-\xi$ and with the elasticity e, as they intuitively both make policy deviations relatively more attractive.

Discretionary policies due to inattention? This stylized model provides distinctive predictions regarding the effect of inattention on tax rates under commitment and under discretion. In the former case, the tax rate should not respond to variations in inattention, while in the latter case, the tax rate increases with inattention. These sharp predictions hinge on the assumption of a rational equilibrium, which implies that information frictions have no impact on labor supply in equilibrium and, therefore, cannot affect the optimal equilibrium policy.

Interestingly, Cabral and Hoxby (2012) estimate the effect of property tax salience on property tax rates, looking at variations in tax salience induced by tax escrow. Because tax escrow materializes after the purchase, it is unlikely to affect the demand for property. Moreover, they report that perceptions about property tax rates are unbiased in their data. As a result, property tax rates under commitment should not respond to variations in the salience of tax escrow. Yet, Cabral and Hoxby (2012) find that tax rates are higher when they are less salient. In related work, Finkelstein (2009) uses the adoption of electronic toll collection as a natural experiment and reaches a similar conclusion.[10] This suggests that policymaking is discretionary and provides suggestive evidence for the existence of a taxation bias.

Generalization. To build the intuition, we have analyzed in this section a static representative agent model in which attention is exogenous, and in which the government implements a linear tax policy to maximize tax revenue. Despite its simplicity, this 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.

[10] Finkelstein (2009) reports that individuals tend to (slightly) overestimate tolls. Such systematic misperceptions are hardly reconcilable with the rational equilibrium considered in the stylized model. Hence, the sharp prediction of the stylized model does not apply directly here. Nevertheless, the rest of the paper extends the analysis to account for potential equilibrium misperceptions. In the presence of overestimation, an increase in inattention leads in equilibrium to an increase or a decrease in the tax rate under discretion (depending on the degree of overestimation), but always leads to a decrease in the tax rate under commitment. Consequently, the observed increase in the tax rate arguably reflects the discretionary nature of policymaking and is consistent with the emergence of a taxation bias.

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 dynamically over time.[11] This type of dynamic learning provides an intuitive justification for taking agents' 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' dynamic learning – still leads to inefficiently high tax rates in the steady state. This indicates that imposing further constraints on the government is unlikely to eliminate the taxation bias, although this could undoubtedly limit the extent to which governments deviate. We discuss these issues in Section 7.

Yet, this stylized model misses important policy-relevant characteristics of the economy. In the remainder of the paper, we broaden the scope of the analysis to better comprehend the taxation bias and its welfare implications.

3 General model for the taxation bias

This section introduces a broader framework that features a welfarist government, heterogeneous agents, and a broad perception model allowing for endogenous attention and arbitrary biases. The aim of this section is to identify necessary and sufficient conditions for the emergence of a taxation bias. We find that these conditions hinge essentially on agents' tax perceptions formation.

3.1 Setup

Consider a population of heterogeneous agents indexed by their productivities w which are private information and distributed from a well-defined probability distribution function $f_w(w)$. The government levies a linear income tax that agents are unable to freely observe, because of information frictions; the tax rate τ and the demogrant R are thus random variables from the agents' point of view. We assume away income effects such that agents'

[11]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, and need not strategically incorporate this deviation into their perceptions.

perceived tax rate $\tilde{\tau}$ is a sufficient statistic for agents' consumption and earnings decisions that we denote $c(\tilde{\tau}; w)$ and $y(\tilde{\tau}; w)$.[12] Let $\mathcal{V}(\tilde{\tau}, \tau_0, R_0; w)$ be the associated indirect utility which depends on the actual tax rate τ_0, and the actual lump sum transfer R_0.

We try to remain as agnostic as possible regarding the tax perception formation process of taxpayers. We directly capture their (posterior) perceptions by assuming that there is a well-defined probability distribution function of perceived tax rates $f_{\tilde{\tau}}(\tau|\tau_0; w)$ that may vary with productivity w and the actual tax rate τ_0. We also assume there is a common prior or default, represented by the probability distribution function $\hat{q}(\tau)$, but we do not make any assumption about the relationship between the two.

The government sets the tax policy to maximize a social welfare function summing an increasing and weakly concave transformation $G(.)$ of taxpayers' indirect utilities $\mathcal{V}(\cdot)$. The government chooses a target tax schedule (τ_g, R_g) implemented up to the realization of an implementation shock ϑ that we introduce as an underlying source of uncertainty to ensure that, for example, rationally inattentive taxpayers have an incentive to learn in equilibrium.[13] The actual tax rate is then given by $\tau_0 = \tau_g + \vartheta$ where ϑ is white noise drawn from an exogenous distribution $f_\Theta(\vartheta)$ which is common knowledge. We assume that the actual demogrant R_0 adjusts to the realization of the implementation shock ϑ such that the government budget constraint is always binding.

The government's problem writes

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

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

where Exp is an exogenous expenditure requirement, and E_ϑ is the expectation over the implementation shock ϑ.

3.2 Tax policy under discretion

When solving problem (6), the government cannot credibly commit to a predefined tax policy because of the time-inconsistency problem discussed in Section 2. As a result, the

[12] We microfound this representation of agents' behavior in a setting with quasi-linear utilities in Section 4 and show in Online Appendix E how the results extend with income effects.

[13] By creating a source of randomness, implementation shocks allow us to formally close the model, as for instance in Matějka and Tabellini (2020).

government cannot directly change the taxpayers' prior about the tax policy (e.g. with a public announcement) and must take this prior as given when choosing its tax policy.

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

0. Agents are endowed with a common prior $\hat{q}(\tau)$ and the distribution of skills is $f_w(w)$.
1. The discretionary government sets the target tax policy (τ_g, R_g) to maximize the expected value of its objective function (6).
2. The actual tax rate $\tau_0 = \tau_g + \vartheta$ is implemented up to an implementation shock drawn from a known distribution $f_\Theta(\vartheta)$.
3. Taxpayers' posterior distribution of beliefs $f_{\tilde{\tau}}(\tau|\tau_0; w)$ adjusts, and they decide how much to consume and earn.
4. The government levies taxes and redistributes through the demogrant R_0 satisfying the resource constraint (7).

The government understands that taxpayers will gather information and adjust their decisions in reaction to its choice of tax policy. It therefore 'plays first' in the above-described sequence of events. However, it treats the prior distribution $\hat{q}(\tau)$ and the skill distribution $f_w(w)$ as predetermined state variables. As a result, the government does not have a particular strategic advantage from 'playing first' – thus reflecting 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 on the taxpayer side.

Proposition 1. *The discretionary tax policy (τ_g, R_g) is characterized by*

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

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

together with the resource constraint (7) and where p represents the social marginal cost of public funds.

The target tax rate τ_g is pinned down by first-order condition (8) which captures the (expected) effects of a marginal increase by $d\tau_g$. The first line relates to changes in allocations when the distribution of perceptions remains fixed. It corresponds to the standard mechanical and welfare effects; a marginal increase in the tax rate mechanically increases tax revenue by $yd\tau_g$ additional dollars, but reduces taxpayers' consumption and thus welfare by $\frac{G'(\mathcal{V})}{p}y\, d\tau_g$ dollars.

The second line relates to changes in the distribution of perceptions and thus captures behavioral effects from the reform. Indeed, earnings responses transit through variations in the posterior distribution $f_{\tilde{\tau}}(\tau|\tau_0, w)$ of perceived tax rates $\tilde{\tau}$. A marginal increase of $d\tilde{\tau}$ in the perceived tax rate thus reduces tax revenue by $\tau_0 y(\tilde{\tau})d\tilde{\tau}$. Moreover, because agents misoptimize, the envelope theorem no longer applies and a marginal deviation from taxpayers' perceived rate induces a welfare cost equal to $\frac{G(\mathcal{V}(\tilde{\tau}))}{p}d\tilde{\tau}$. This new welfare effect introduces a corrective motive for taxation in the presence of misperceptions common to optimal tax models with behavioral agents (Gerritsen, 2016; Farhi and Gabaix, 2020).

Condition (9) states that in the absence of income effects, social marginal welfare weights $g \equiv \frac{G'(\mathcal{V})}{p}$ average to 1 at the optimum, and the government is indifferent between having an additional dollar or redistributing an additional dollar (Saez, 2001).

3.3 Equilibrium, commitment and the taxation bias

Equilibrium. Given the structure of the problem, the only free variables are the government's target tax rate τ_g, the posterior distribution $f_{\tilde{\tau}}(\tau|\tau_0, w)$ capturing agents' perceived tax rates, and the equilibrium distribution of the common prior $\hat{q}(\tau)$. Arguably, in equilibrium, agent's prior may somehow be related to the actual tax policy. We formalize the equilibrium concept in Definition 1.

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

(a) Given the prior distribution $\hat{q}(\tau)$, the target tax rate τ_g^{eq} solves the government's problem.

(b) Given the prior distribution $\hat{q}(\tau)$, agents' posterior distribution $f_{\tilde{\tau}}(\tau|\tau_0, w)$ follows from agents' attention strategies.

(c) The common prior distribution adjusts to the target policy of the government such that $\hat{q}(\tau) = \Gamma(\tau, \tau_g^{\text{eq}})$.

This equilibrium definition is general enough to encompass many different models of perceptions. Conditions (a) and (b) ensure that neither the government nor the agents have an incentive to deviate; while condition (c) ensures that agents' prior and actual tax policy are mutually consistent in equilibrium. Specifically, the condition $\hat{q}(\tau) = \Gamma(\tau, \tau_g^{\text{eq}})$ allows us to consider a variety of equilibrium concepts, with or without perception biases.[14]

In equilibrium, the discretionary tax policy is generally suboptimal. To formalize this point, we now analyze the equilibrium welfare-maximizing tax policy. Implicitly, this is the policy described in papers studying optimal taxation with behavioral agents (e.g. Farhi and Gabaix (2020)). It coincides with the tax policy implemented by a government which, ex-ante, can credibly commit to implementing a given tax policy and change taxpayers' prior about the policy. We therefore refer to it as the tax policy under commitment.

Commitment policy. A government which can credibly commit to a predefined policy has to take into account the impact of its choice of tax policy on agents' prior. Hence, the commitment tax policy solves the government's problem (6) subject to the additional feasibility condition that agents' prior and actual tax policy realizations are mutually consistent in equilibrium (condition (c) in Definition 1).

Proposition 2. *The commitment tax policy $(\tau_g^\star, R_g^\star)$ is given by*

$$(\tau_g^\star) \quad : \quad E_\vartheta\Big[\int \Big\{ \underbrace{\int \Big[-\frac{G'(\mathcal{V})}{p} y + y\Big] f_{\tilde{\tau}}(\tau|\tau_0; w) d\tau}_{\text{mechanical \& welfare effects}} + \underbrace{\int \Big[\frac{G(\mathcal{V})}{p} + \tau_0 y\Big] \frac{df_{\tilde{\tau}}(\tau|\tau_0; w)}{d\tau_g} d\tau}_{\text{direct \& equilibrium behavioral effects}} \Big\} f_w(w) dw\Big] = 0 \tag{10}$$

$$(R_g^\star) \quad : \quad E_\vartheta\Big[\iint \Big[\frac{G'(\mathcal{V})}{p} - 1\Big] f_{\tilde{\tau}}(\tau|\tau_0; w) f_w(w) d\tau dw\Big] = 0 \tag{11}$$

[14]For instance, in a rational equilibrium in which taxpayers' prior coincides with the actual distribution of tax rates, $\Gamma(\tau, \tau_g)$ is the pdf of $\tau_g^{\text{eq}} + \vartheta$.

together with the resource constraint (7) and where p represents the social marginal cost of public funds. This is the policy implemented in a commitment equilibrium.

Proposition 2 characterizes the commitment tax policy. The main difference between Propositions 1 and 2 is that the change in agents' perceived tax rate $\frac{df_{\tilde{\tau}}(\tilde{\tau}|\tau_0;w)}{d\tau_g}$ in equation (10) now reflects both the direct and equilibrium impact of the change in tax policy. In other words, it encapsulates the equilibrium adjustment of agents' prior.

Taxation bias. The discrepancy between the discretionary and commitment equilibria constitutes a taxation bias. It is a measure of the deviation from the ex-post optimal feasible tax policy $\tau_g^\star$.

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

The taxation bias arises from the government's inability to credibly commit to a tax policy, because of the time-inconsistency that can result from an equilibrium adjustment in taxpayers' perceptions. These deviations reflect the government's inability to account for equilibrium adjustments in agents' perceptions ex-ante. Proposition 3 provides necessary and sufficient conditions for the existence of a taxation bias.

Proposition 3. *When both equilibria exist and are unique, there is a taxation bias if and only if*

$$E_\vartheta\left[\iint\left(\frac{G(\mathcal{V})}{p}+\tau_0 y\right)\Bigg(\underbrace{\frac{df_{\tilde{\tau}}(\tau|\tau_g^\star+\vartheta;w)}{d\hat{q}(\tau)}}_{\text{(i)}}\underbrace{\frac{d\mathit{\Gamma}(\tau,\tau_g)}{d\tau_g}}_{\text{(ii)}}\Bigg)f(w)d\tau dw\right]\neq 0 \tag{12}$$

Intuitively, the left-hand side of (12) represents expected *equilibrium* behavioral effects upon a marginal tax change. Hence, Proposition 3 shows there is a taxation bias whenever equilibrium responses are non-zero. More specifically, (12) decomposes equilibrium responses into two distinct channels implying that a taxation bias emerges when both (i) posterior perceptions are affected by prior beliefs, and (ii) prior beliefs are affected by the government choice of tax policy in equilibrium.

3.4 Drivers of the taxation bias

Proposition 3 shows that a taxation bias emerges under two conditions. We discuss each of these in turn, and argue they are likely to be satisfied in practice.

(i) Should we expect posterior perceptions to be affected by prior beliefs? From a theoretical perspective, this prediction is central to Bayesian updating and adopted in many theories of perceptions and attention (Gabaix, 2019). It also arises in models in which agents anchor their perceptions on a particular heuristic or default, which have been discussed comprehensively in the context of taxation (Bernheim and Taubinsky, 2018). Even in instances in which the anchor is independent of prior beliefs – e.g. models of salience (Chetty et al., 2009) or sparsity (Gabaix, 2014) in which the anchor is often assumed to be nil – prior beliefs affect posterior perceptions through the (ex-ante) allocation of attention. Indeed, prior beliefs drive the choice of attention and, therefore, any changes in prior beliefs affect posterior perceptions through a change in taxpayer attention.

Empirically, a large behavioral literature analyzes perception formation, both in the lab and in the field, and supports this theoretical prediction (Coibion and Gorodnichenko, 2012, 2015; Armantier et al., 2016; Khaw et al., 2017). Moreover, there is compelling evidence that agents' attention increases with taxes (Hoopes et al., 2015; Taubinsky and Rees-Jones, 2018; Morrison and Taubinsky, 2021), providing support for models in which attention is endogenous and influenced by prior beliefs. Gabaix (2019) provides a meta-analysis of attention variations across stakes.

(ii) Should we expect prior beliefs to be affected by tax policy in equilibrium? The idea that equilibrium variables are affected by actual policy goes back to the Lucas critique (Lucas Jr, 1976). Hence, prior beliefs about tax policy are, arguably, not deep parameters immune to changes in tax policy. Beyond its theoretical appeal, this claim is supported by the host of empirical evidence showing that taxpayers' beliefs adjust with time and experience (Sausgruber and Tyran, 2005; Fochmann and Weimann, 2013), and are formed using heuristic linearizations of the actual tax system (Liebman and Zeckhauser, 2004; Rees-Jones and Taubinsky, 2020). More generally, people seem to report beliefs that are surprisingly consistent with actual tax policy (Chetty et al., 2009; Cabral and Hoxby, 2012; Taubinsky and Rees-Jones, 2018), although political partisanship may bias some of these beliefs when it comes to redistributive tax policy (Alesina et al., 2020; Stantcheva, 2020).[15]

[15]Arguably, this dependence should also translate into earnings choices. Chetty (2012) documents a systematic difference between micro (capturing direct adjustments) and macro (capturing total adjust-

4 Application and empirical magnitude

Having emphasized that empirical evidence on tax perceptions and observed tax policies point to the existence of a taxation bias, we now gauge its magnitude. To this end, we consider a restricted but fully tractable version of the general model from Section 3, where we provide microfoundations of taxpayer behavior. Calibrating this structural model to the US economy, we then carry out numerical simulations. Lastly, we derive a simple sufficient statistics formula for the taxation bias that we take to data.

4.1 A tractable and microfounded model

We consider a setting in which agents make two types of decision. First, they decide on their earnings y and consumption c, given their perceptions of the income tax. Second, they decide on their attention to taxes ξ which shape these perceptions. Relying on standard assumptions, we present here the microfoundations of the tractable model used in the empirical applications.

Earnings-consumption choice. We assume that taxpayers use a linear representation $T^{\ell}(y)$ of the (potentially nonlinear) tax schedule $T(y)$ when making their consumption and earnings decisions. Consequently, taxpayers only have to form estimates of the marginal tax rate τ and the intercept R.[16]

Assumption 1 (linear representation)**.** *Individuals use a linear representation of the tax schedule $T^{\ell}(y) = \tau y - R$.*

Without income effects, the tax level does not influence earnings choices which depend only on agents' average perceived marginal tax rate $\tilde{\tau}$. We thus assume away income effects in the body of the paper and leave the case with income effects to Online Appendix E. More specifically, we assume agents hold quasi-linear and iso-elastic preferences in our

ments) estimates of the elasticity of taxable income, and rationalizes this difference by the existence of adjustment rigidities such as information frictions at the micro level. The difference between these two adjustments may well reflect an equilibrium adjustment in the prior.

[16]Beyond the fact that a linear approximation is usually a good approximation of existing tax schedules (Piketty and Saez, 2013), empirical evidence suggests that people tend to use linear representations of tax schedules (Liebman and Zeckhauser, 2004; Rees-Jones and Taubinsky, 2020).

applications, such that $U(c, y; \omega) = c - (y/\omega)^{1+\varepsilon}/(1+\varepsilon)$ and the first-order condition for earnings is[17]

$$y(\tilde{\tau}; \omega) = \omega \left[\omega \left(1 - \tilde{\tau}\right)\right]^{\frac{1}{\varepsilon}} \tag{14}$$

Assumption 2 (slack budget)**.** *Consumption adjusts such that agents exhaust their actual budget, i.e.,* $c(\tilde{\tau}; w) = y(\tilde{\tau}; w) - T(y(\tilde{\tau}; w))$.

Assumption 2 guarantees that agents' budget constraint holds, by assuming that consumption adjusts.[18] This uniquely determines agent 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). Agents' indirect utility function is then $V(\tilde{\tau}, \tau_0, R_0; \omega)$ with τ_0 and R_0 the actual parameters of the linearized tax schedule around agents' bliss point.

Attention choice. Turning to formation process of agent perceptions, we build on the rational inattention literature (Sims, 2003) and adopt a Bayesian learning model with a choice of information (following Gabaix (2019) classification). We choose this model due to its extensive use in economics, its well-understood microfoundations and because it generates predictions that are consistent with empirical evidence.

Let $\hat{q}(\tau)$ be agents' prior probability distribution about the tax rate τ, which is a random variable from the agent perspective. Given their prior, agents choose how much information to collect about the (unknown) actual tax rate τ_0. This information is modeled as an unbiased signal s whose precision depends on agent efforts to acquire and process information, i.e., their attention to taxes.

Assumption 3 (tractable Gaussian learning)**.** *We assume that:*

- *the signal s is drawn from a Gaussian distribution with mean τ_0 and variance σ^2,*

[17] Formally, τ is a random variable from the agent perspective. Denoting $\tilde{q}(\tau)$ its perceived probability distribution, we have $\tilde{\tau} \equiv E_{\tilde{q}(.)}[\tau]$ and the agents' problem writes

$$\max_{c,y} \int U(c, y; w)\ \tilde{q}(\tau) d\tau \text{ s.t. } c \leq R + (1-\tau)y \tag{13}$$

[18] See Reck (2016) and Farhi and Gabaix (2020) for a discussion of alternative budget adjustment rules.

- *agents choose their attention strategies relying on a quadratic approximation of their indirect utility.*

Assumption 3 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 κ 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 $\hat{q}(\tau)$ is Gaussian with mean $\hat{\tau}$ and variance $\hat{\sigma}^2$, the optimal attention strategy $\xi(\kappa,\hat{\sigma},\hat{\tau};\omega)$ is given by*

$$\xi(\kappa,\hat{\sigma},\hat{\tau};\omega) = \max\left(0, 1 + \frac{\kappa}{\hat{\sigma}^2 \int \frac{\partial^2 y(\tau;\omega)}{\partial \tilde{\tau}^2} \hat{q}(\tau;\hat{\tau},\hat{\sigma}) d\tau}\right). \tag{15}$$

Attention decreases with the information cost κ and with the precision of the prior $1/\hat{\sigma}$. Moreover, attention increases with the expected responsiveness of agents' 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. In the empirical applications we focus on a particular class of equilibria that we call biased Gaussian equilibria.

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

- *the implementation shock ϑ is a Gaussian white noise with variance σ_ϑ^2,*
- *the prior distribution $\hat{q}(\tau)$ adjusts in equilibrium to the expected distribution of the actual tax rate τ_0 up to a bias b,*

such that the equilibrium prior is Gaussian with mean $\tau_g + b$ and variance σ_ϑ^2.

The Gaussian assumption for implementation shocks allows us to enjoy the tractability of the Gaussian learning model, while introducing an exogenous and additive bias b induces a simple parametrization of misperceptions.[19]

Since the bias b in the prior is the only source of misperception, attention to taxes determines the degree of misperception in equilibrium. Indeed, the equilibrium posterior mean $\mu \equiv E_s[\tilde{\tau}(s,\sigma)] = \tau_g + \xi\vartheta + (1-\xi)b$ is, in expectation, equal to

$$E_\vartheta[\mu] = \tau_g + (1-\xi)b \tag{16}$$

where the bias b is scaled by inattention $1-\xi$ such that inattention determines the magnitude of potential perception biases (if any).

4.2 Numerical illustrations in a calibrated model

Tax policies. Corollary 1 characterizes the tax rates under discretion τ_g^{eq} and under commitment $\tau_g^\star$ in biased Gaussian equilibria. These optimal rates follow respectively from applying Propositions 1 and 2 to the Gaussian setup.

Corollary 1. *Using a first-order approximation of perceived tax rates in the integrand, and introducing social marginal welfare weights as $g(\tilde{\tau}) \equiv \frac{G'(\mathcal{V}(\tilde{\tau}))}{p}$,*

- *the Gaussian discretionary equilibrium tax policy (τ_g^{eq},R_g^{eq}) solves*

$$E_\vartheta\Big[\int \Big\{(1-g)\,y + \Big(g(1-\xi)(b-\vartheta)+\tau_0\Big)\frac{dy}{d\tilde{\tau}}\xi\Big\}\Big|_{\tilde{\tau}=\mu} dF_w(w)\Big] = 0 \tag{17}$$

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

- *the Gaussian commitment equilibrium tax policy ($\tau_g^\star$,$R_g^\star$) solves*

$$E_\vartheta\Big[\int \Big\{(1-g)\,y + \Big(g(1-\xi)(b-\vartheta)+\tau_0\Big)\frac{dy}{d\tilde{\tau}}\Big(1-\frac{d\xi}{d\tau_g}(b-\vartheta)\Big)\Big\}\Big|_{\tilde{\tau}=\mu} dF_w(w)\Big] = 0 \tag{18}$$

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

[19] More generally, our framework allows for parametric biases, e.g., a function that depends on the tax policy. Nevertheless, we believe that much of the explanatory gains from allowing for biased perceptions are already well-captured when introducing an exogenous and additive bias b.

Corollary 1 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 ξ in (17), whereas direct and equilibrium behavioral effects are proportional to $1-\frac{d\xi}{d\tau_g}(b-\vartheta)$ in (18) and thus match one-for-one with tax increases up to potential debiasing effects. As a result, equilibrium behavioral effects are proportional to $(1-\xi)-\frac{d\xi}{d\tau_g}(b-\vartheta)$ and thus generically non-zero in the presence of inattention, leading to a taxation bias in this environment.

Going back to our earlier discussion of the drivers of the taxation bias, the inattention term $1-\xi$ reflects the fact that (i) posterior perceptions are affected by prior beliefs through attention ξ, while (ii) equilibrium prior beliefs move one-for-one with the target tax policy. Normalizing the implementation shock to zero, the debiasing term $\frac{d\xi}{d\tau_g}b$ captures the fact that (i) misperceptions in the posterior are caused by a potential bias b in the prior, while (ii) attention ξ is affected in equilibrium by the target tax rate τ_g through adjustment of the prior.

Numerical simulations. We carry out numerical simulations in a calibrated US economy. The skill distribution is specified by inverting the distribution of earnings from the 2016 March CPS data, and the existing tax schedule. We extend the earnings distribution with a Pareto tail ($k = 2$) for incomes above \$200,000 (Saez, 2001). We assume that the government's objective is given by a logarithmic transformation of agents' utilities, and we set agents' structural elasticity parameter to $e = 1/\varepsilon = 0.33$ (Chetty, 2012).

[Figure 2 about here]

The left panel of Figure 2 represents equilibrium tax rates under discretion (solid line) and commitment (dashed line) for different values of the information cost κ. They correspond to a variation from full attention to average attention levels below 0.5, given the implementation shocks that we consider.[20] When agents are fully attentive, equilibrium tax rates under discretion and under commitment coincide, and are equal to 46% in this calibrated economy. However, as the information cost κ increases, agents become less attentive and the tax rates diverge. Equilibrium tax rates under discretion increase

[20]Equation (15) shows that attention ξ depends on the ratio $\frac{\kappa}{\hat{\sigma}^2}$. We here fix the value of the variance of implementation shocks which determines $\hat{\sigma}^2$, and consider variations of the information cost κ.

rapidly, whereas equilibrium tax rates under discretion remain fairly stable. The diverging paths of equilibrium tax rates under discretion and commitment suggest that even small information frictions may generate a significant taxation bias, as can be seen on the right panel of Figure 2.

Surprisingly, the magnitude of the taxation bias in numerical simulations seems almost unaffected by the presence of a bias b in agents' prior. Indeed, comparing results with (grey lines) and without (black lines) a bias, shows that the presence of bias has a relatively similar effect on equilibrium tax rates under both discretion and commitment, thereby leaving the taxation bias almost unaffected. Hence, the magnitude of the taxation bias seems relatively insensitive to the introduction of perception bias.

This perhaps counter-intuitive and empirically important result appears very clearly in the sufficient statistics formula to which we now turn to. Moreover, the sufficient statistics formula allows us to estimate the magnitude of the taxation bias leveraging direct measures of agents' attention to taxes.

4.3 A sufficient statistics approach

Sufficient statistics formula. We derive the following sufficient statistics formula for the taxation bias.

Proposition 4. *With small implementation shocks ϑ and small prior biases b, a sufficient statistics formula for the taxation bias is*

$$\tau_g^{\text{eq}} - \tau_g^\star \simeq \frac{\overline{(1-\xi)y}\ e\ t^2}{\overline{(1-g)y} + \overline{(1-\xi)y}\ e\ t} \tag{19}$$

where all endogenous right-hand-side quantities are evaluated at the actual tax rate t and where we have introduced the mean operator $\bar{x} \equiv \int x(w)f(w)dw$.

This sufficient statistics formula extends the one derived in Section 2 to more realistic settings with a welfare-maximizing government and heterogeneous agents. Everything else equal, the formula shows that the size of taxation bias increases with the (structural) elasticity of labor supply e, with the square of the actual tax rate t and decreases with the government redistributive tastes.[21]

[21] Intuitively, the taxation bias increases with government redistributive tastes, as it relatively increases the incentives to implement unanticipated tax increases. However, this first-order effect here transits through an increase in the actual tax rate t and we obtain the inverse relationship controlling for t.

Importantly, the taxation bias is (approximately) independent of the potential bias b and depends fundamentally on the income-weighted average inattention $\overline{(1-\xi)y}$ in the population. This captures the fact that when inattention is concentrated on low income individuals – as seems to be the case generally – the government has weaker incentives to increase tax rates, because tax increases have a weaker impact on tax revenue.

Application to the US economy. To gauge the empirical magnitude of the taxation bias in the current US economy, we bring our sufficient statistics formula to US data.

The income tax and transfer system in the US is well approximated by a linear tax schedule with a tax rate t of 29.46%. At this tax rate, the meta-analysis of Gabaix (2019), which combines existing measures of attention in order to map the evolution of attention with the stakes, and recent evidence on attention to taxes point to an average attention $\overline{\xi}$ of 0.55 as a baseline.[22] Using our model of endogenous attention and the distribution of income, we then compute the associated *income-weighted* average attention $\overline{y\xi}$.

Turning to other sufficient statistics, we use a structural elasticity parameter e of 0.33 as estimated in the meta-analysis of Chetty (2012), and use an inverse optimum approach to estimate the government redistributive motive $\overline{(1-g)y}$ which is consistent with actual tax policy. More precisely, we infer this last term by inverting our expression for the equilibrium tax policy under the assumption that the bias in agents' prior is small.

With this baseline calibration of the actual US economy, we estimate that the taxation bias has an order of magnitude of 6.04 percentage points. This suggests that the actual US tax rate, approximately 29.46 percentage points, is 20% higher than would be socially optimal, given the extent of inattention to tax policy.

[Table 1 about here]

Table 1 provides a sensitivity analysis with varying average attention. For realistic attention parameters, the magnitude of the taxation bias in the actual US economy ranges

[22] 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 the attention to sales taxes in an experiment, Taubinsky and Rees-Jones (2018) find an average attention parameter of 0.25 (respectively 0.48) for a tax rate of 7% (respectively 22%). Assuming that attention increases linearly with the stake – which is conservative – this implies an average attention parameter of 0.59 for a tax rate of 30%. We take the average of these four estimates (0.55) as a baseline, and report results for alternative values to acknowledge the uncertainty surrounding the value of attention.

from 4.45 to 7.89 percentage points. Our baseline estimate of 6.04 percentage points lies in the middle of this range, suggesting that the taxation bias is likely to be large.

5 Welfare implications

This section analyzes the welfare implications of agents' inattention to taxes. Inattention leads to the emergence of a taxation bias which decreases efficiency and welfare, but inattention also allows for misperceptions that may in fact increase efficiency and welfare. The net welfare effect is thus ambiguous and depends crucially on the magnitude of inattention relative to equilibrium misperceptions. Our application to US data suggests that the net welfare effect may be positive as agents tend to underestimate marginal tax rates.

5.1 A decomposition of the welfare effects of inattention

So far, we have shown that agents' inattention to taxes generates inefficient tax increases, leading to the existence of a taxation bias. Moreover, we have shown that this taxation bias is likely to be large, given existing measures of inattention. This suggests that inattention is detrimental to welfare.

Yet, inattention is also conducive to the emergence of perception biases in equilibrium.[23] While such biases do not have a first-order impact on the size of the taxation bias, they may considerably affect efficiency and welfare. Indeed, a recurring theme in the behavioral tax literature is that agent underestimation of (marginal) tax rates reduces the efficiency cost of taxation (Liebman and Zeckhauser, 2004; Rees-Jones and Taubinsky, 2020). Therefore, inattention may also positively affect welfare if accompanied by underestimations of taxpayers' own marginal tax rates.

To illustrate how these two opposing forces affect welfare, denote SW^{eq}, $SW^{\star}$, and SW^{FI} the equilibrium social welfare under discretion, commitment and full information respectively. When agents are fully attentive, the economy is in the full information benchmark and all three measures coincide. When agents are inattentive, this is no longer the case and we can decompose the net welfare change between the welfare losses

[23] For instance, in our tractable model, the equilibrium perception bias from equation (16) is proportional to the bias in prior b scaled by inattention $1 - \xi$.

induced by the taxation bias, and the welfare gains from underestimation:

$$\Delta SW \equiv SW^{\text{eq}} - SW^{FI} \quad = \quad \underbrace{SW^{\text{eq}} - SW^{\star}}_{\text{taxation bias}} + \underbrace{SW^{\star} - SW^{\text{FI}}}_{\text{underestimation}} \tag{20}$$

To put some numbers on this decomposition, we proceed in two steps. First, we rely on numerical simulations to understand how these two forces affect welfare. Second, we leverage existing measures of inattention and tax underestimation to gauge the net welfare implications in the US economy.

5.2 Convex welfare losses and concave welfare gains

Relying on our calibrated model of the US economy, we compute these welfare effects for different values of the information cost κ, holding the prior bias b fixed. Higher information costs induce higher inattention, which increases both the taxation bias and tax underestimation in equilibrium. The results are presented in Figure 3, which depicts the welfare losses from the taxation bias (light gray area), the welfare gains from tax underestimation (dark gray area), and the net welfare impact (dashed line).

[Figure 3 about here]

Simulations indicate that the welfare gains from tax underestimation outweigh the welfare losses due to the taxation bias at low levels of inattention. Yet, these welfare gains are increasing and concave with inattention, while the welfare losses from the taxation bias are increasing and convex.[24] Hence, there is an inattention threshold below which net welfare impacts are positive, and above which they become negative.

This implies that the underestimation of (marginal) tax rates is a necessary condition for misperceptions to be welfare-improving, but not a sufficient condition. Indeed, tax underestimation increases economic efficiency for any *given* tax policy. Yet, it also signals the presence of inattention and thus that of inefficiently high tax rates due to the taxation bias. Our analysis reveals that if agents are too inattentive to tax policy, the welfare losses from the taxation bias may outweigh the welfare gains from tax underestimation. In that case, the underestimation of (marginal) tax rates is paradoxically associated with a *decrease* in economic efficiency and social welfare.

[24] Although this result holds for all reasonable calibrations, signs of the second derivatives of these welfare effects are difficult to determine analytically.

This finding casts a new light on the welfare implications of inattention and tax underestimation. In what follows, we leverage existing measures of inattention and underestimation to gauge the magnitude of net welfare effects in the actual US economy.

5.3 Net welfare impact in the US economy

As before, we consider a baseline average attention of 0.55 and use our calibrated model to compute income-weighted average attention. Among the few estimates of agent perceptions of their marginal tax rates in the US, a recent and well identified estimate is that of Rees-Jones and Taubinsky (2020) who report that taxpayers underestimate their marginal tax rates by 19% on average.[25] We use this value as a baseline and consider other values in a sensitivity analysis. Importantly, we interpret it as a measure of the average bias in agents' posterior $\overline{(1-\xi)b}$ and report the results as a function of this observable parameter, rather than as a function of the unobservable bias in agents' prior b.

For given values of average attention and posterior bias, we compute the welfare decomposition (20) as follows. First, we infer the bias in agents' prior b. Second, we use the current distribution of income and the current tax rate to calibrate the distribution of agents' skills w and the information cost κ relying on agents' first-order conditions.[26] Third, assuming the government has a Benthamite objective (i.e. $G(\mathcal{V}) = \frac{\mathcal{V}^{1-\alpha}}{1-\alpha}$), we infer the redistributive parameter $\alpha \in [0,1]$ consistent with the current tax rate relying on the sufficient statistics formula (A.21). We can then simulate counterfactual policies, and use the calibrated Benthamite objective to compute the different welfare effects.

[Table 2 about here]

[25] They identify that a sizable fraction of taxpayers use the ironing heuristic, that is, taxpayers who linearize the income tax schedule with their average tax rate rather than their marginal tax rate. The former is smaller than the latter because of the transfers received at low income and because of increasing marginal tax rates. The average measure used in our calibration thus masks some heterogeneity in perceptions, but it is consistent with earlier findings from Fujii and Hawley (1988) and De Bartolome (1995) who find agents to underestimate on average their marginal tax rates by 12% and 19% respectively. In contrast, estimates of tax overestimation all relate to average tax rates (Gideon, 2017; Ballard et al., 2018; Stantcheva, 2020).

[26] This is achieved through an iterative procedure to match average attention in the population.

The results are reported in Table 2 where welfare changes are expressed in terms of annual consumption-equivalent variations. Panel A reports the efficiency gains associated with tax underestimation. They increase social welfare by the same amount as an annual increase of \$153 in all agents' consumption in the baseline. Panel B reports the welfare losses associated with the taxation bias. The losses decrease social welfare, by the same amount as an annual decrease of \$28 in all agents' consumption in the baseline. Panel C reports the net welfare impact of inattention which is thus positive in the baseline. Net the increase in social welfare is equivalent to an annual increase of \$125 in all agents' consumption. Since average annual consumption per person in the US is \$43,114, this represents a 0.3% increase in average consumption.

5.4 Discussion

The baseline calibration suggests that the welfare gains from tax underestimation outweigh the welfare costs from the taxation bias such that the net effect is positive. This result relies heavily on the relatively large tax underestimation estimated by Rees-Jones and Taubinsky (2020) in the US. Yet, high underestimation is incompatible with high attention such that the presence of a taxation bias reduces the (net) welfare gains associated with tax underestimation. In the baseline, we estimate that it reduces the welfare gains of underestimation by almost 20%.

Performing the same analysis without inferring the redistributive taste of the government – and assuming a log social welfare function – drastically increases the consumption-equivalent associated to the welfare gains, losses and net change (Panel E). The relative effect of the taxation bias is, nevertheless, relatively unaffected as we conclude that it reduces the welfare gains of underestimation by about 23%.

The welfare impacts of the taxation bias are transmitted through two different channels. While inefficient tax increases have welfare costs, they also induce increases in redistribution, because the additional tax revenue collected is used to finance a larger lump-sum transfer, by assumption. As a result, the efficiency costs associated with the taxation bias are largely compensated for by equity gains in our simulations. In contrast, if we assumed that some of the additional tax revenue was spent inefficiently, captured by politicians or led to higher government expenditures, the welfare cost of the taxation bias would be larger, which may eradicate the previously reported positive net effect.

To illustrate this point and picture the sizable increase in tax revenue induced by the taxation bias, Panel D of Table 2 reports the optimal size of the government as a fraction of its actual size. In the baseline calibration, we estimate that tax revenue should be 13% lower in the US, implying that inattention considerably inflates the size of the government through the taxation bias. This finding is connected to the theory of fiscal illusion in the social choice literature, which posits that inattention to taxes lead voters to choose higher levels of taxes, redistribution, and spending than they would otherwise (Mill, 1848; Buchanan, 1967; Oates et al., 1985; Dollery and Worthington, 1996; Sausgruber and Tyran, 2005).

Furthermore, we focus here on the welfare costs of the taxation bias associated with the income tax. However, a taxation bias is likely to affect many other tax instruments. For instance, we have already argued that this seems to be the case for property taxes and road tolls. In a related work, Furukawa (2021) analyzes the welfare implications of the taxation bias in the context of consumption taxes. Therefore, computing the overall deadweight loss associated with the taxation bias would require aggregating the impacts over all tax instruments. We leave that to future research, and maintain our focus on income taxation to next analyze how the taxation bias affects the progressivity of nonlinear tax schedules.

6 Extension to nonlinear taxation

In this section, we consider the implications for nonlinear tax schedules. We find that the taxation bias is heterogeneous across income levels, with a larger wedge at low and middle incomes. This distortion in tax progressivity flattens the U-shape pattern of marginal tax rates and more than doubles the welfare losses associated with the taxation bias.

6.1 Introducing nonlinear tax schedules

We introduce a nonlinear tax schedule $T(y)$, but the setup introduced in Section 4 is otherwise unchanged. In particular, we maintain Assumption 1 that individuals use a linear representation of the tax schedule $T^{\ell}(y) = \tau y - R$ which now raises a new question: in the continuum of marginal tax rates $\{T'(y)\}_y$, what is the marginal tax rate that agent w gathers information about?

Absent income effects, the perceived marginal tax rate $\tilde{\tau}_w$ remains a sufficient statistic for labor supply and uniquely determines earnings $y(\tilde{\tau}_w; w)$. We denote $\hat{\tau}_w$ agent w prior about its marginal tax rate and make the following assumption:

Assumption 4 (prior reliance)**.** *Taxpayer w gathers information about the actual marginal tax rate $\tau_0(\hat{\tau}_w, w) \equiv T_0'(\hat{y}_w)$ at her ex ante optimal earnings level $\hat{y}_w \equiv y(\hat{\tau}_w; w)$.*

Essentially, Assumption 4 guarantees the internal consistency of the perception-formation process by ensuring that agents have no additional information ex ante than that contained in their prior. Moreover, it gives a novel allocative role to the prior, as taxpayers linearize the tax schedule around the income level they deem optimal ex ante.[27]

The introduction of a nonlinear tax schedule does not fundamentally affect the equilibrium concepts. Two refinements are nevertheless necessary. First, for the sake of simplicity, we assume that implementation shocks ϑ uniformly affect marginal tax rates at all earnings levels y, such that $T_0'(y) = T_g'(y) + \vartheta$. Second, Definition 3 characterizing biased Gaussian equilibria should be amended to account for type-specific prior distributions as follows

- *The type-specific prior distribution $\hat{q}_w(\tau)$ adjusts in equilibrium to the expected distribution of the actual tax rate $T_0'(\hat{y}_w)$ up to a bias b.*

That is, each taxpayer's prior is consistent with her marginal tax rate of interest up to an arbitrary perception bias b. Incidentally, the prior average $\hat{\tau}_w \equiv E_{\hat{q}_w(.)}[\tau]$ is thus necessarily type-specific in equilibrium when the government implements a nonlinear tax schedule. While natural in our context, this poses a potential challenge for the resolution of this nonlinear tax model.

Indeed, type-specific priors $\hat{\tau}_w$ potentially threaten the existence of an increasing mapping between earnings y and skills w, a requirement known as the monotonocity condition. In the Online Appendix, we show that under standard assumptions, the monotonicity condition is expected to hold when $T_g''(.)$ is smooth enough, just as in the full informa-

[27] To illustrate this new allocative role, consider the limit where the information cost κ goes to zero. Perceptions are then perfect $\tilde{\tau}_w = \tau_0(\hat{\tau}_w, w)$ and each agent chooses earnings $y_{w|\hat{\tau}_w} = y(\tau_0(\hat{\tau}_w, w); w)$. This is in contrast to the full information case in which earnings are the solution to a fixed-point problem characterized by $y_w = y(T_0'(y(.)); w)$. Both income concepts coincide in a rational equilibrium ($b = 0$), but will differ in biased equilibria ($b \neq 0$).

tion benchmark.[28] As a result, we solve for the optimal tax schedule, assuming that the monotonicity condition is verified and check ex post that it holds.

6.2 ABC formula with inattentive taxpayers

The government chooses a target nonlinear tax schedule $T_g(.)$ that consists of a continuum of marginal tax rates $\{T_g'(y)\}_y$ and a tax level, indexed by $T_g(0)$. The target tax policy is implemented up to an implementation shock ϑ on marginal tax rates, and the tax level adjusts to satisfy the government budget constraint. The government problem writes as follows:

$$\max_{T_g'(.),T_g(0)} \quad E_\vartheta\Big[\iint G\Big(\mathcal{V}(\tilde{\tau}_w, T_0(.);\kappa,w)\Big)\ f_{\tilde{\tau}}(\tau|\tau_0(\hat{\tau}_w,w);w)\ f_w(w)\ d\tau dw\Big] \tag{21}$$

$$\text{s.t.} \quad \iint T_0(y^\star(\tilde{\tau}_w;w))\ f_{\tilde{\tau}}(\tau|\tau_0(\hat{\tau}_w,w);w)\ f_w(w)\ d\tau dw \geq Exp \tag{22}$$

where Exp is an exogenous expenditure requirement, $f_{\tilde{\tau}}(\tau|\tau_0(\hat{\tau}_w,w);w)$ is the posterior distribution of agent w perceived tax rate, and the indirect utility function is

$$\mathcal{V}(\tilde{\tau}_w, T_0(.);\kappa,w) = y(\tilde{\tau}_w;w) - T_0(y(\tilde{\tau}_w;w)) - v\big(y(\tilde{\tau}_w;w);w\big) - \kappa\mathcal{I}(\sigma). \tag{23}$$

We solve this problem 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 tax perturbation literature, this reform triggers three effects: mechanical, welfare and behavioral. In this setting, analyzing the impact of a reform calls for a careful identification of the agents affected by the reform.

The standard mechanical and welfare effects capture changes in taxes and welfare for agents with actual earnings above y^r. Following from the aforementioned monotonicity condition, it corresponds to all agents with a productivity level above w^r, defined such that $y(\tilde{\tau}_{w^r};w^r) \equiv y^r$. In contrast, the behavioral effect comes from taxpayers who are learning the marginal tax rate affected by the reform. That is, all agents whose ex ante optimal earnings level $\hat{y}$ belongs to $[y^r-\Delta y, y^r]$. Using the monotonicity condition, we

[28] Following Jacquet and Lehmann (2017), we assume that: (i) the tax function $T_g(.)$ is twice differentiable, (ii) the optimization program of each taxpayer admits a unique global maximum, (iii) agents' second-order conditions hold strictly.

can identify these agents as those with a productivity $w \in [\hat{w}^r - \Delta\hat{w}, \hat{w}^r]$ implicitly defined such that $y(\hat{\tau}_{\hat{w}^r}; \hat{w}^r) \equiv y^r$. Since the two cut-offs w^r and $\hat{w}^r$ almost surely differ, these different effects do not involve the same individuals.[29]

We report the discretionary and commitment equilibrium tax schedules in Proposition 5. We assume small Gaussian implementation shocks to obtain interpretable conditions, relegating general conditions to the Online Appendix.

Proposition 5 (ABC formula)**.** *Assuming small Gaussian implementation shocks, the equilibrium nonlinear tax schedule is, to a first-order approximation, characterized by*

$$\frac{T_g'(y(\mu_{\hat{w}^r}; \hat{w}^r)) + g(\hat{w}^r)|_{\tilde{\tau}=\mu_{\hat{w}^r}} \Big(\mu_{\hat{w}^r} - T_g'(y(\mu_{\hat{w}^r}; \hat{w}^r))\Big)}{1 - \mu_{\hat{w}^r}} \tag{24}$$

$$= \frac{1}{e \frac{d\mu_{\hat{w}^r}}{d\tau_g}|_{\tilde{\tau}=\mu_{\hat{w}^r}}} \frac{1}{y(\mu_{\hat{w}^r}; \hat{w}^r)} \frac{\frac{dy(\hat{\tau}_{\hat{w}^r}; \hat{w}^r)}{dw}}{f_w(\hat{w}^r)} \int_{w^r}^{\infty} \Big(1 - g(w)|_{\tilde{\tau}=\mu_w}\Big) f_w(w)\, dw$$

together with the transversality condition $\int g(w)_{|\tilde{\tau}_w=\mu_w}\, dF(w) = 1$ and the government budget constraint (22). All endogenous quantities are evaluated at their equilibrium values. Moreover, the posterior average perceived marginal tax rate is $\mu_w \equiv \xi\tau_0(\hat{\tau}_w, w) + [1-\xi]\hat{\tau}_w$ such that $\frac{d\mu_w}{d\tau_g} = \xi$ under discretion and $\frac{d\mu_w}{d\tau_g} = 1 + \frac{d\xi}{d\tau_g}[\tau_0(\hat{\tau}_w, w) - \hat{\tau}_w]$ under commitment.

Under commitment and absent perception biases ($b = 0$), the ABC formula boils down to the one derived in Diamond (1998) and the usual interpretation prevails. The presence of perception biases ($b \neq 0$) has several effects. First, it creates a wedge between w^r and $\hat{w}^r$ that is new to this nonlinear setting. Second, it adds a welfare effect to the LHS, related to the failure of the envelope theorem. Third, it modifies the inverse elasticity term on the RHS to account for inattention ($\frac{d\mu_w}{d\tau_g}$ term).

As before, the emergence of a taxation bias is due to the time-inconsistency of tax policy. Because inattention anchors taxpayers' perceptions, an increase $\Delta\tau^r$ in marginal tax rates only increases perceived marginal tax rates by $\xi\Delta\tau^r$. Yet, adjustments in the priors lead to larger changes in tax perceptions in equilibrium. As a result, marginal tax rates are again inefficiently high in equilibrium.

In this nonlinear setting, the taxation bias at a level of earnings y^r is driven by the attention of agents of type $\hat{w}^r$. Surprisingly, these agents may not even be located

[29]The two cut-offs coincide only when $\hat{\tau}_w = \tilde{\tau}_w$. That is, when $b = 0$ and $\vartheta = 0$. Since we focus on Gaussian implementation shocks here, it is never the case (a.s.).

at earnings y^r in the presence of perception biases. More importantly, because attention varies across the earnings distribution, the magnitude of the taxation bias becomes earnings-specific, thereby, altering the progressivity of nonlinear tax schedules.

6.3 Numerical illustration

To illustrate this property, we represent nonlinear tax schedules under discretion (dashed black line) and under commitment (solid black line) in Figure 4. Simulations are carried out assuming no perception biases ($b = 0$), such that the nonlinear tax schedule under commitment corresponds to the textbook optimal nonlinear tax schedule of Saez (2001), and we retrieve the known U-shape pattern of marginal tax rates.

[Figure 4 about here]

Because of the taxation bias, equilibrium marginal tax rates are higher than optimal. This pattern is much more pronounced at low income levels; marginal tax rates increase respectively by 9, 5, and 2 percentage points for agents located at the first decile, the median, and the top decile of the earnings distribution.

This impact on tax progressivity reflects the gradient in attention ξ across earnings depicted in Figure 4 (grey lines). In our model, more productive agents have more latitude to choose their earnings and thus attach a higher value to information acquisition. As a result, attention increases globally with earnings.[30] Note that this pattern is obtained assuming that all individuals have the same information cost κ. Assuming that more able workers are also more efficient at collecting information would only reinforce this result.

Finally, we can compute the welfare effect of the taxation bias for nonlinear tax schedules as in Section 5. The welfare loss arising from the taxation bias is equivalent to a \$73 reduction in annual consumption for all agents. For a meaningful comparison, the consumption-equivalent welfare loss arising from the taxation bias, assuming that the government uses linear taxes instead, is \$31. Hence, accounting for the effect of the taxation bias on tax progressivity increases the welfare losses by 135%.[31] The distortion in tax progressivity induced by inattention thus has substantial welfare implications.

[30]The pattern reverts at the very beginning of the earnings distribution, because, as the marginal tax rate approaches one, we approach the origin of the labor supply function where earnings become infinitely responsive to changes in the marginal tax rates.

[31]The reason is that the shift towards large increases in marginal tax rates at low-income levels induce

7 Discussion

As consistently emphasized by the economics literature, policymaking is (at least to some extent) discretionary. Accordingly, we show that agents' inattention to taxes leads to a taxation bias. In this final section, we discuss the role of information asymmetries, tax design and competitive politics in the emergence of a taxation bias, and argue that eliminating this bias requires a commitment technology that can only be delivered through institutional changes.

Information asymmetries. Arguably, if government deviations from its commitment tax policy were easily observable ex post, there would be many ways to solve the government's commitment problem. For instance, political constraints such as "do not deceive the electorate" or reputational concerns could suffice. Yet, deviations are unlikely to be observable ex post, due to information asymmetries. Indeed, the government is privately informed about its spending needs (Sleet, 2004), or about its redistributive objective, as clearly illustrated by the inverse-optimum literature seeking to infer marginal social welfare weights from existing tax policy (Bourguignon and Spadaro, 2012; Lockwood and Weinzierl, 2016; Jacobs et al., 2017).

These information asymmetries, and the lack of accountability that they imply, are well recognized and motivate the recurring initiatives for better 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), and have incentives to manipulate their communication about (income) taxation (Albornoz et al., 2014). As a result, there are good reasons not to expect voluntary fiscal transparency from governments, which severely reduces the observability of policy deviations and the set of potential solutions to the taxation bias.

Tax design. The time-inconsistency of the government hinges on taxpayers' inattention to taxes. Therefore, any improvement in tax design that may 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 which

larger increases in tax revenue. Indeed, the optimal size of the government is 96% of equilibrium tax revenue under a linear income tax, but 89% of equilibrium tax revenue under a nonlinear income tax.

is known to increase attention to tax reforms (Abeler and Jäger, 2015).[32]

Yet, only full attention from each and every taxpayer can resolve the commitment problem of the government. Nevertheless, even the most salient and simplest tax system might remain costly to understand and internalize in one's economic choices, so that changes in tax design are unlikely to eliminate the taxation bias. Moreover, the most salient and simplest tax system may not be optimal if agents tend to underestimate marginal tax rates (Rees-Jones and Taubinsky, 2020), or if tax complexity allows to take better account of 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 unlikely, because we consider the behavior of a welfarist government such that any agent (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 (2020) show in related work that electoral competition with inattentive voters also leads to policy distortions.

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 agents (Jackson and Yariv, 2014, 2015). More generally, the large body of research on the inflation bias and on fiscal illusion provides pervasive evidence that competitive politics leads to inefficient policymaking in the presence of information frictions.

Commitment technologies. In conclusion, voluntary fiscal transparency, improvements in tax design, and competitive politics are unlikely to resolve the commitment problem of the government in its choice of tax policy. This suggests that solving this problem might require tangible commitment technologies. Building on the large litera-

[32]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.

ture 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).[33]

Second, in a dynamic setting, the tax policy implemented by a patient fiscal authority which does not discount the future – and internalizes taxpayers' dynamic learning – coincides with the tax policy under commitment (see Online Appendix B.2). Delegating the conduct of tax policy to such a fiscal authority could thus eliminate the taxation bias.

Our point is not to argue in favor of these solutions, but instead to highlight that both would require substantial institutional changes. Obviously, such changes would also pose other practical and political challenges, and we leave the analysis of these important issues to future work.

8 Conclusion

Building on recent evidence on tax perceptions, this paper show that the anchoring effect of agents' inattention to taxes creates a time-inconsistency problem in the choice of tax policy. Ultimately, inattention induces a taxation bias, such that taxes are inefficiently high in equilibrium. We show that the taxation bias is likely to be large, that it distorts the progressivity of nonlinear tax schedules, and that it significantly reduces social welfare. Overall, our findings sheds new light on the policy and welfare implications of inattention and misperceptions.

While we focus on income taxation, it should be clear that the taxation bias is bound to appear in different tax settings, insofar as taxes are non-salient and agents inattentive. In this respect, our findings are in line with recent evidence on the effect of tax salience on tax rates gathered in the context of road tolls and property taxes (Finkelstein, 2009; Cabral and Hoxby, 2012). Moreover, and perhaps more importantly, the insight that agent inattention leads to time-inconsistency issues probably extends to a wide range of applications, in which the framework developed in this paper could be extended and

[33] Alternatively, a principal-agent optimal contract à la Walsh (1995) may also be considered if fiscal transparency cannot be imposed.

fruitfully applied.

References

Abeler, J. and Jäger, S. (2015). Complex tax incentives. *American Economic Journal: Economic Policy*, 7(3):1–28. 1, 35

Albornoz, F., Esteban, J., and Vanin, P. (2014). Market distortions and government transparency. *Journal of the European Economic Association*, 12(1):200–222. 34

Alesina, A., Miano, A., and Stantcheva, S. (2020). The polarization of reality. In *AEA Papers and Proceedings*, volume 110, pages 324–28. 17

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. 5

Alt, J., Preston, I., and Sibieta, L. (2010). The political economy of tax policy. *In stitute of Fiscal Studies (Hrsg.), Dimensions of Tax Design: The Mirrlees Review, Oxford (im Erscheinen)*. 35

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. 34

Armantier, O., Nelson, S., Topa, G., Van der Klaauw, W., and Zafar, B. (2016). The price is right: Updating inflation expectations in a randomized price information experiment. *Review of Economics and Statistics*, 98(3):503–523. 17

Ballard, C. L., Gupta, S., et al. (2018). Perceptions and realities of average tax rates in the federal income tax: evidence from michigan. *National Tax Journal*, 71(2):263–294. 27

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. 6, 9

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. 36

Bernheim, B. D. and Taubinsky, D. (2018). Behavioral public economics. In *Handbook of Behavioral Economics*, volume 1, pages 381–516. Elsevier. 5, 17

Billi, R. M. (2011). Optimal inflation for the us economy. *American Economic Journal: Macroeconomics*, 3(3):29–52. 6

Bourguignon, F. and Spadaro, A. (2012). Tax–benefit revealed social preferences. *The Journal of Economic Inequality*, 10(1):75–108. 4, 34

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. 35

Buchanan, J. M. (1967). The fiscal illusion. *Public finance in democratic process: Fiscal Institutions and individual choice.* 5, 6, 29

Cabral, M. and Hoxby, C. (2012). The hated property tax: salience, tax rates, and tax revolts. Technical report, National Bureau of Economic Research. 1, 6, 10, 17, 36

Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018. 17, 22, 24, 43, 45, 46, 47, 48, 60

Chetty, R., Looney, A., and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–77. 1, 17

Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, 120(1):116–159. 17

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. 6, 17

Conseil Constitutionnel (2005). Décision n. 2005-530 DC du 29 décembre 2005. 35

Conseil Constitutionnel (2012). Décision n. 2012-662 DC du 29 décembre 2012. 35

De Bartolome, C. A. (1995). Which tax rate do people use: Average or marginal? *Journal of Public Economics*, 56(1):79–96. 27

DellaVigna, S. and Pollet, J. M. (2009). Investor inattention and friday earnings announcements. *The Journal of Finance*, 64(2):709–749. 24

Diamond, P. A. (1998). Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95. 32

Dollery, B. E. and Worthington, A. C. (1996). The empirical analysis of fiscal illusion. *Journal of Economic Surveys*, 10(3):261–297. 29

Dovis, A. and Kirpalani, R. (2020). Fiscal rules, bailouts, and reputation in federal governments. *American Economic Review*, 110(3):860–88. 6

Dupuit, J. (1995). De la mesure de l'utilité des travaux publics (1844). *Revue française d'économie*, 10(2):55–94. 2

Eichengreen, B. (1989). The capital levy in theory and practice. Technical report, National Bureau of Economic Research. 6

Farhi, E. and Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review*, 110(1):298–336. 1, 5, 14, 15, 19

Farhi, E., Sleet, C., Werning, I., and Yeltekin, S. (2012). Non-linear capital taxation without commitment. *Review of Economic Studies*, 79(4):1469–1493. 6

Finkelstein, A. (2009). E-ztax: Tax salience and tax rates. *The Quarterly Journal of Economics*, 124(3):969–1010. 1, 6, 10, 36

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. 17

Fujii, E. T. and Hawley, C. B. (1988). On the accuracy of tax perceptions. *The Review of Economics and Statistics*, pages 344–347. 27

Furukawa, C. (2021). Tax transparency and social welfare: The role of government commitment. *Available at SSRN 3862433*. 5, 29

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4):1661–1710. 17

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, 2, 3, 5, 6, 7, 8, 17, 19, 24, 43

Gerritsen, A. (2016). Optimal taxation when people do not maximize well-being. *Journal of Public Economics*, 144:122–139. 5, 14

Gideon, M. (2017). Do individuals perceive income tax rates correctly? *Public Finance Review*, 45(1):97–117. 27

Goldin, J. (2015). Optimal tax salience. *Journal of Public Economics*, 131:115–123. 1, 5

Halac, M. and Yared, P. (2014). Fiscal rules and discretion under persistent shocks. *Econometrica*, 82(5):1557–1614. 6

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. 3, 17

Jackson, M. O. and Yariv, L. (2014). Present bias and collective dynamic choice in the lab. *American Economic Review*, 104(12):4184–4204. 35

Jackson, M. O. and Yariv, L. (2015). Collective dynamic choice: the necessity of time inconsistency. *American Economic Journal: Microeconomics*, 7(4):150–78. 35

Jacobs, B., Jongen, E. L., and Zoutman, F. T. (2017). Revealed social preferences of dutch political parties. *Journal of Public Economics*, 156:81–100. 4, 34

Jacquet, L. and Lehmann, E. (2017). Optimal income taxation with composition effects. *Journal of the European Economic Association*. 31

Kaplow, L. (1994). Accuracy, complexity, and the income tax. *National Bureau of Economic Research Working Paper Series*. 35

Khaw, M. W., Stevens, L., and Woodford, M. (2017). Discrete adjustment to a changing environment: Experimental evidence. *Journal of Monetary Economics*, 91:88–103. 17

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. 35

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. 5, 6, 9

Liebman, J. B. and Zeckhauser, R. J. (2004). Schmeduling. *National Bureau of Economic Research Working Paper Series*. 4, 17, 18, 25

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. 4, 34

Lucas Jr, R. E. (1976). Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. North-Holland. 17

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. 35

Mackowiak, B. A., Matějka, F., and Wiederholt, M. (2021). Rational inattention: A review. *Journal of Economic Literature*. 3, 51

Matějka, F. and Tabellini, G. (2020). Electoral Competition with Rationally Inattentive Voters. *Journal of the European Economic Association*. jvaa042. 12, 35

Mill, J. S. (1848). *Principles of political economy*. 5, 29

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208. 4

Moore, D. T. and Slemrod, J. (2021). Optimal tax systems with endogenous behavioral biases. *Journal of Public Economics*, 197:104384. 5

Morrison, W. and Taubinsky, D. (2021). Rules of thumb and attention elasticities: Evidence from under-and overreaction to taxes. Technical report. 1, 3, 5, 17

Oates, W. E. et al. (1985). *On the nature and measurement of fiscal illusion: A survey*. Citeseer. 6, 29

Piketty, T. and Saez, E. (2013). Chapter 7 - optimal labor income taxation. In Alan J. Auerbach, Raj Chetty, M. F. and Saez, E., editors, *Handbook of Public Economics, vol. 5*, volume 5 of *Handbook of Public Economics*, pages 391 – 474. Elsevier. 7, 18

Ramsey, F. P. (1927). A contribution to the theory of taxation. *The economic journal*, 37(145):47–61. 2

Reck, D. (2016). Taxes and mistakes: What's in a sufficient statistic? *SSRN 2268617*. 19

Rees-Jones, A. and Taubinsky, D. (2020). Measuring "schmeduling". *The Review of Economic Studies*, 87(5):2399–2438. 1, 3, 4, 5, 17, 18, 25, 27, 28, 35

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68(1):205–229. 4, 14, 22, 33, 60

Saez, E. and Stantcheva, S. (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics*, 162:120–142. 6

Sausgruber, R. and Tyran, J.-R. (2005). Testing the Mill hypothesis of fiscal illusion. *Public Choice*, 122(1-2):39–68. 17, 29

Scheuer, F. and Wolitzky, A. (2016). Capital taxation under political constraints. *American Economic Review*, 106(8):2304–28. 6

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690. 3, 19, 20

Sleet, C. (2004). Optimal taxation with private government information. *The Review of Economic Studies*, 71(4):1217–1239. 34

Stantcheva, S. (2020). Understanding tax policy: How do people reason? *National Bureau of Economic Research Working Paper Series*. 17, 27

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. 1, 3, 17, 24, 43

Walsh, C. E. (1995). Optimal contracts for central bankers. *American Economic Review*, pages 150–167. 36

Tables and figures

Table 1: Magnitude of the taxation bias for US income tax rates

Average attention	0.35	0.45	**0.55**	0.65	0.75
Taxation bias (p.p.)	10.10	7.89	**6.04**	4.45	3.05
Taxation bias (% of t)	34.27	26.76	**20.49**	15.10	10.35

NOTE: Estimates of the magnitude of the taxation bias follow from our sufficient statistics characterization (Proposition 4). The value in bold corresponds to our baseline calibration where we assume average attention is $\overline{\xi} = 0.55$ (Taubinsky and Rees-Jones, 2018; Gabaix, 2019) and we use our calibrated structural model to compute the corresponding income-weighted average attention. Other sufficient statistics are calibrated as follows: the earnings elasticity is $e = 0.33$ (Chetty, 2012), the actual US income tax rate is 29.46% (OECD), and the governments' objective function is inferred from actual tax policy.

Table 2: Welfare implications for the US economy

Panel A: Consumption-equivalent welfare gains from underestimation ($/yr)			
Underestimation		Average attention	
(% of t)	0.45	**0.55**	0.65
-19	152	**153**	153
-10	88	88	89
0	0	0	0

Panel B: Consumption-equivalent welfare losses from taxation bias ($/yr)			
Underestimation		Average attention	
(% of t)	0.45	**0.55**	0.65
-19	-46	**-28**	-16
-10	-56	-34	-19
0	-66	-40	-22

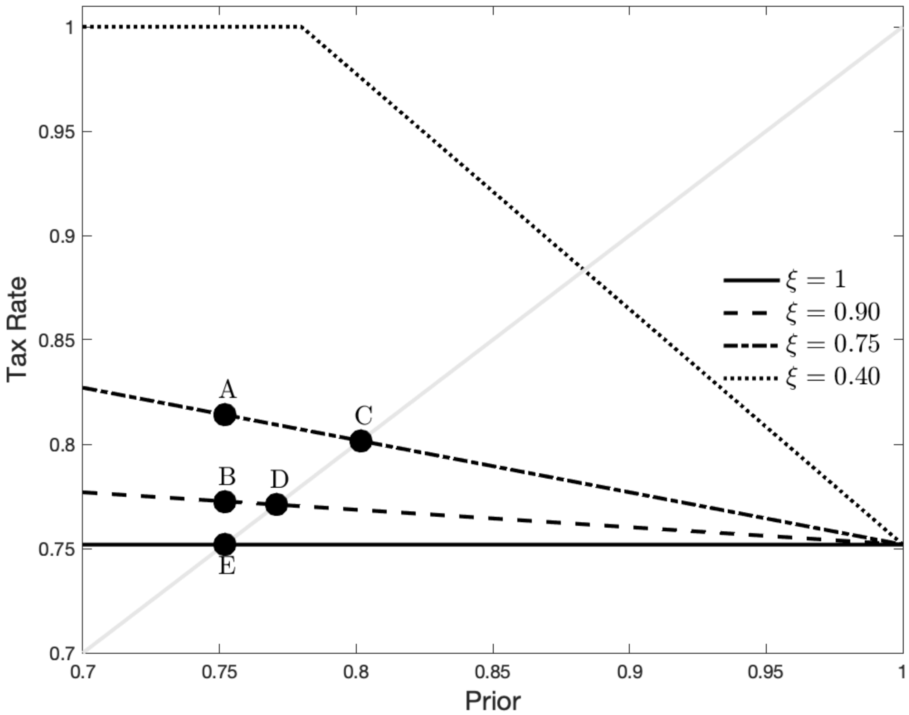
Panel C: Consumption-equivalent net welfare change ($/yr)			
Underestimation		Average attention	
(% of t)	0.45	**0.55**	0.65
-19	105	**125**	137
-10	32	55	70
0	-66	-40	-22

Panel D: Optimal government size (% of actual size)			
Underestimation		Average attention	
(% of t)	0.45	**0.55**	0.65
-19	83	**87**	91
-10	81	86	90
0	79	84	88

Panel E: Welfare decomposition with a log social welfare function ($/yr)			
Underestimation		Average attention	
(-19 % of t)	0.45	0.55	0.65
Gains from underestimation	440	403	374
Losses from taxation bias	-170	-95	-49
Net welfare change	270	308	325

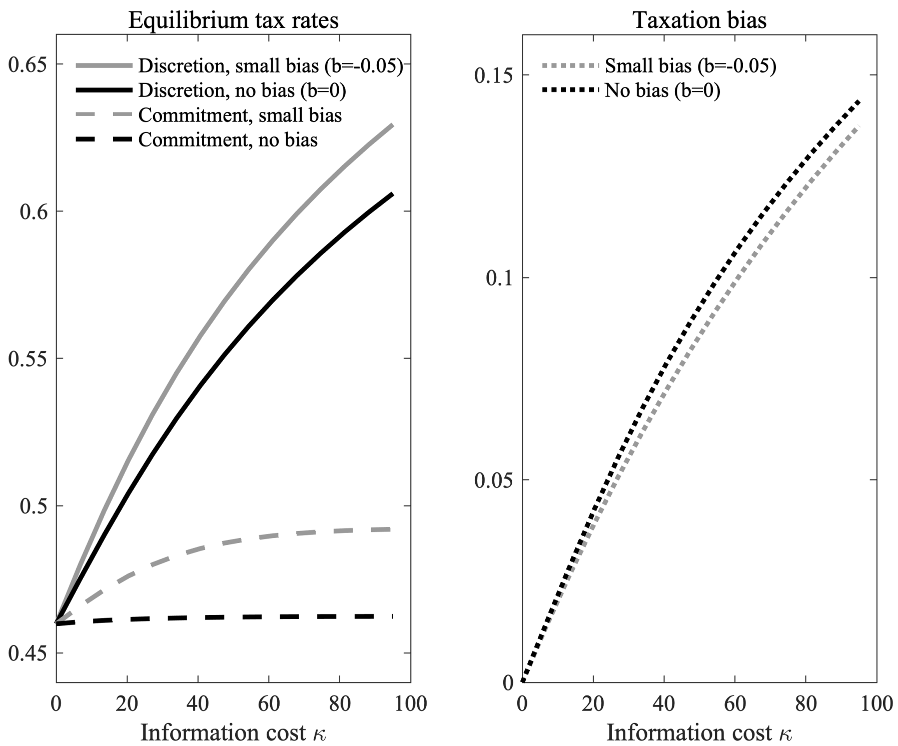
NOTE: Panels A, B and C respectively report average consumption-equivalent welfare gains, losses and net changes measured in dollars per year per individual – see welfare decomposition (20). Panel D reports the optimal government size (i.e. tax revenue) as a fraction of actual government size. Values in bold correspond to our baseline calibration. Panel E reports the previous welfare decomposition when the government's objective is given by a log social welfare function, instead of being inferred from actual tax policy, under baseline tax underestimation.

Figure 1: Policy functions and equilibrium outcomes in stylized model



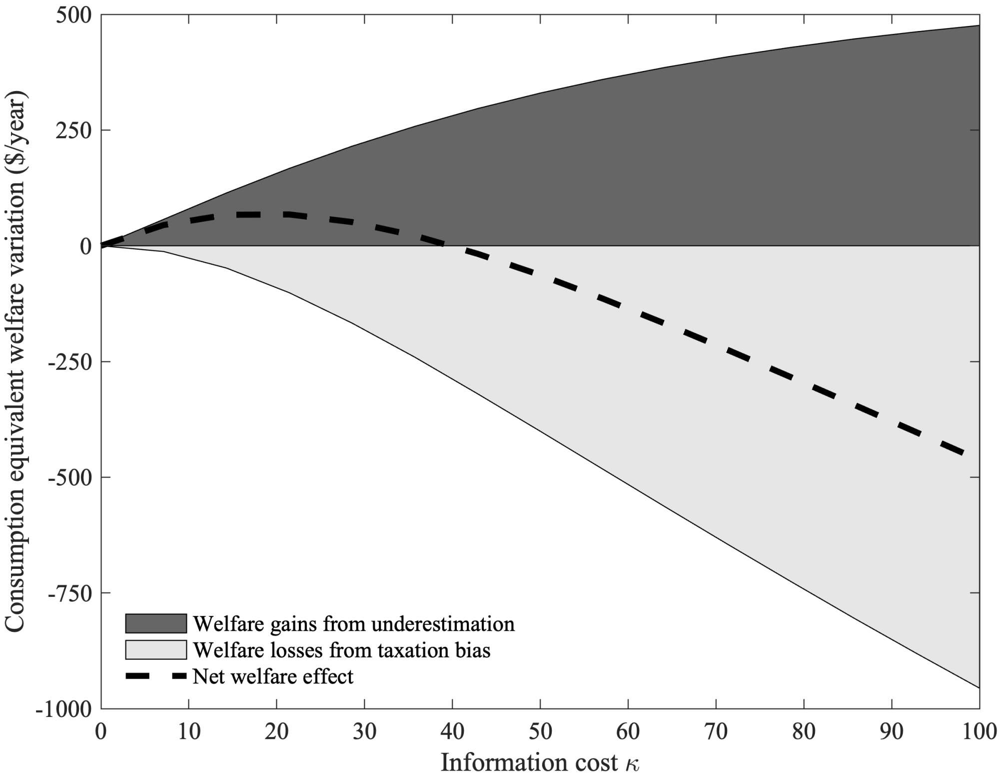
Note: Black lines represent policy functions in the stylized model, that is the government's chosen tax rate as a function of taxpayers' prior for different attention parameters ξ. The grey line represents the (rational) equilibrium condition which is that agents' prior coincides with the government's chosen tax rate. Assuming an elasticity of aggregate earnings with respect to the perceived net-of-tax rate of 0.33 (Chetty, 2012), the revenue maximizing policy is a 75% tax rate, point E. Points A and B correspond to the government's policy deviations when inattentive agents expect the government to implement the revenue maximizing tax rate, and points C and D correspond to the associated equilibrium tax rates.

Figure 2: Equilibrium taxes and taxation bias in calibrated structural model



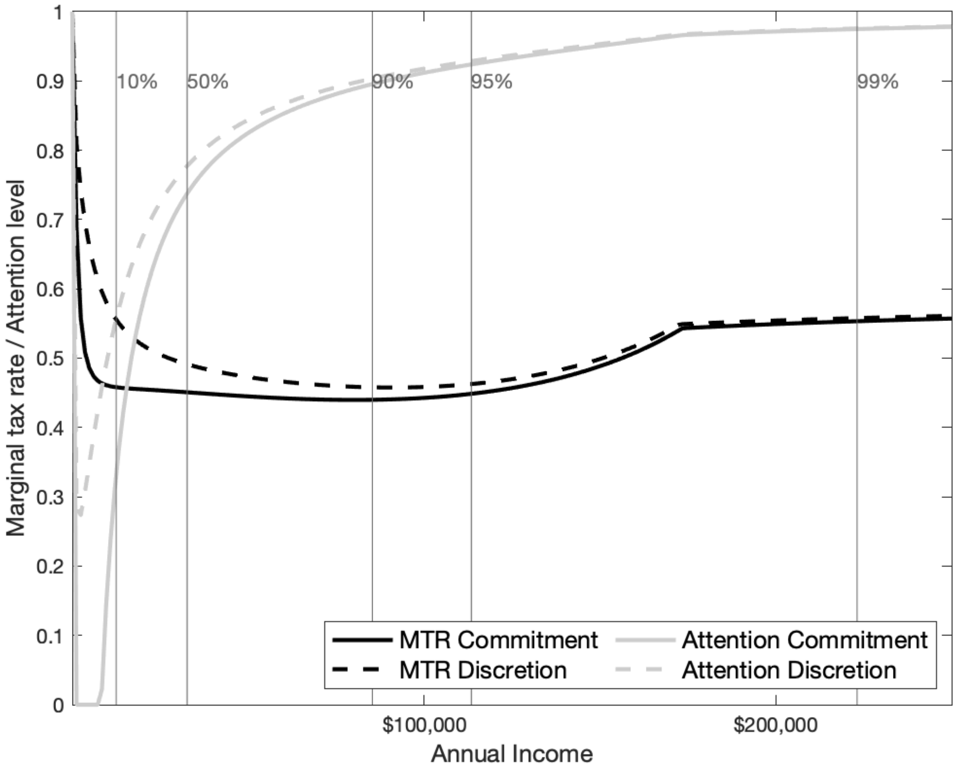
NOTE: The left panel reports equilibrium tax rates under discretion and commitment for different values of the information cost κ expressed in \$/bit/year in a calibrated structural model of the US economy. The right panel reports the associated taxation bias. Gaussian implementation shocks ϑ have a standard deviation equal to 0.05. b is the bias in agents' prior. The government has a log social welfare function and its policy under discretion and under commitment follow from Corollary 1. Taxpayers have quasi-linear iso-elastic preferences with $e = 0.33$ (Chetty, 2012). The distribution of skills $f_w(w)$ is calibrated using 2016 CPS data and a Pareto tail for high incomes.

Figure 3: Welfare implications in calibrated structural model



NOTE: This Figure represents the welfare decomposition from equation (20) in a calibrated structural model of the US economy for different values of the information cost κ assuming agents' prior underestimates marginal income tax rates by 5 percentage points ($b = -0.05$). The standard deviation for Gaussian implementation shocks is equal to 0.05. The government has a log social welfare function and its (discretionary) policy follows from Corollary 1. Taxpayers have quasi-linear iso-elastic preferences with $e = 0.33$ (Chetty, 2012). The distribution of skills $f_w(w)$ is calibrated using 2016 CPS data and a Pareto tail for high incomes.

Figure 4: Implications for the progressivity of nonlinear tax schedules



Note: This Figure represents the schedule of marginal tax rates (black lines) and the value of attention (grey lines) across income levels under discretion (dashed lines) and commitment (solid lines). We assume an information cost $\kappa = 16$\$/bit/year and no perception bias $b = 0$. The government has a log social welfare function and policies follow from Proposition 5. Taxpayers have quasi-linear iso-elastic preferences with $e = 0.33$ (Chetty, 2012). The distribution of skills $f_w(w)$ is calibrated using 2016 CPS data and a Pareto tail for high incomes.

A Appendix

A.1 Solution to the stylized model with imperfect information

The government seeks to maximize tax revenue taking the prior as given. Its problem writes $\max_\tau \tau Y(1-\tilde{\tau})$ such that $\tilde{\tau} = \xi\tau + (1-\xi)\hat{\tau}$, $\{\tau, \hat{\tau}\} \in [0,1]^2$ and $\xi \in (0,1)$. The associated Lagrangian is $\mathscr{L}(\tau, \lambda) = \tau Y(1 - \xi\tau - (1-\xi)\hat{\tau}) + \lambda(\tau - 1)$. Following from the first order Kuhn and Tucker conditions, $\tau = 1$ 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 also sufficient since the problem is convex under the assumption that $\tau Y(1-\tau)$ is concave.

At the rational equilibrium, the prior is correct $\hat{\tau} = \tau^\star$. Guess that the rational equilibrium is interior. Hence, $\tau^\star = \frac{1}{1+\xi e}$. Because $e > 0$, it implies that $\hat{\tau} > 1 - \frac{\xi}{1-\xi}e$ in equilibrium, thus confirming the guess. It is then straightforward to prove that $\tau^\star Y(1 - \tau^\star) < \tau^r Y(1-\tau^r)$ where $\tau^r \equiv \frac{1}{1+e}$ as $\tau^r = \arg\max_{\tau\in[0,1]} \tau Y(1-\tau)$. Moreover, the taxation bias $\tau^\star - \tau^r = \frac{(1-\xi)e}{(1+\xi e)(1+e)}$ is strictly positive for all $\xi \in (0,1)$.

A.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 where all agents' responses are taken into account (including equilibrium adjustments) to obtain Proposition 2 and from which Proposition 1 naturally follows.

The Lagrangian associated to problem (6) writes

$$\begin{aligned} \mathscr{L}(\tau_g, R, p) \;=\; & E_\vartheta\Big[\iint \Big[G\Big(\mathcal{V}(\tilde{\tau}, \tau_g + \vartheta, R, \kappa; w)\Big) \\ & +p\Big((\tau_g+\vartheta)y(\tilde{\tau}; w) - R_0 - E\Big)\Big] f_{\tilde{\tau}}(\tilde{\tau}|\tau_g+\vartheta; w) f_w(w) d\tilde{\tau} dw\Big] \end{aligned} \tag{A.1}$$

The first-order condition associated with the choice of the marginal tax rate τ_g is

$$\begin{aligned} \frac{1}{p}\frac{d\mathscr{L}}{d\tau_g} \;=\; & E_\vartheta\Big[\int \Big\{ \int \Big[\frac{G'(\mathcal{V})}{p}\frac{d\mathcal{V}}{d\tau_g} + y\Big] f_{\tilde{\tau}}(\tilde{\tau}|\tau_g+\vartheta; w) d\tilde{\tau} \\ & + \int \Big[\frac{G(\mathcal{V})}{p} + (\tau_g+\vartheta)y - R_0 - E\Big] \frac{df_{\tilde{\tau}}(\tilde{\tau}|\tau_g+\vartheta; w)}{d\tau_g} d\tilde{\tau}\Big\} f_w(w) dw\Big] \end{aligned} \tag{A.2}$$

where $\frac{df_{\tilde{\tau}}(\tilde{\tau}|\tau_g+\vartheta;w)}{d\tau_g}$ is the change in the posterior distribution of perceived tax rate for type w and captures agents' responses to tax changes.

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

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

This is equation (10) from Proposition 2 and characterizes the commitment tax rate. Equation (8) from Proposition 1 which characterizes the tax rate chosen by a discretionary government is obtained when agents' responses to a change in the tax rate are computed holding agents' prior $\hat{q}$ constant. That is $\frac{df_{\tilde{\tau}}(\tilde{\tau}|\tau_g+\vartheta;w)}{d\tau_g}\Big|_{\hat{q}(.)}$ replaces $\frac{df_{\tilde{\tau}}(\tilde{\tau}|\tau_g+\vartheta;w)}{d\tau_g}$ in equation (A.3).

The first-order condition associated with the choice of the demogrant R is

$$\frac{1}{p}\frac{d\mathscr{L}}{dR}=E_\vartheta\Big[\iint\Big[\frac{G'(\mathcal{V})}{p}\frac{d\mathcal{V}}{dR}-1\Big]f_{\tilde{\tau}}(\tilde{\tau}|\tau_g+\vartheta;w)f_w(w)d\tilde{\tau}dw\Big] \tag{A.4}$$

By quasi-linearity we have $\frac{d\mathcal{V}}{dR}=1$. The optimality condition $\frac{1}{p}\frac{d\mathscr{L}}{dR}=0$ thus writes

$$E_\vartheta\Big[\iint\Big[\frac{G'(\mathcal{V})}{p}-1\Big]f_{\tilde{\tau}}(\tilde{\tau}|\tau_g+\vartheta;w)f_w(w)d\tilde{\tau}dw\Big]=0 \tag{A.5}$$

This is equation (9) from Proposition 1 and equation (11) from Proposition 2.

A.3 Proof of Proposition 3

$\tau_g^\star$ solves equation (10) which involves both *direct* and *equilibrium* perception changes $\frac{df_{\tilde{\tau}}(\tau|\tau_g^\star+\vartheta;w)}{d\tau_g}$, whereas τ_g^{eq} solves equation (8) which involves only *direct* perception changes $\frac{df_{\tilde{\tau}}(\tau|\tau_g^\star+\vartheta;w)}{d\tau_g}|_{\hat{q}(.)}$. Realizing that

$$\frac{df_{\tilde{\tau}}(\tau|\tau_g+\vartheta;w)}{d\tau_g}-\frac{df_{\tilde{\tau}}(\tau|\tau_g^\star+\vartheta;w)}{d\tau_g}\Big|_{\hat{q}(.)}=\frac{df_{\tilde{\tau}}(\tau|\tau_g^\star+\vartheta;w)}{d\hat{q}(\tau)}\underbrace{\frac{d\hat{q}(\tau)}{d\Gamma(\tau,\tau_g)}}_{=1}\frac{d\Gamma(\tau,\tau_g)}{d\tau_g} \tag{A.6}$$

condition (12) implies that the left hand-side of equation (8) is different from 0 when evaluated at $\tau_g^\star$. Hence, it directly follows from the existence and uniqueness of the discretionary equilibrium that $\tau_g^{\text{eq}}\neq\tau_g^\star$ if and only if (12) holds.

A.4 Proof of Lemma 1

An agent chooses the signal precision – or equivalently its standard error σ – that maximizes her expected utility

$$\max_{\sigma} \quad \iint V\big(\tilde{\tau}(s,\sigma),\tau,R;w\big)\ \phi(s;\tau,\sigma)\ \hat{q}(\tau)\ dsd\tau - \kappa\mathcal{I}(\sigma) \tag{A.7}$$

where

$$\mathcal{I}(\sigma) \equiv H(\hat{q}(\tau)) - E_{p(s)}\Big[H(\tilde{q}(\tau|s;\sigma))\Big] \tag{A.8}$$

$H(q(\tau)) \equiv -\int q(\tau)\log_2(q(\tau))d\tau$ is the differential entropy (in bits) of the probability distribution $q(\tau)$ and $E_{p(s)}[.]$ 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 3, the taxpayer relies on a second order Taylor approximation of her indirect utility around τ_0 when solving for her attention choice. We have

$$V^2_{\tau_0}(\tilde{\tau},\tau_0,R_0;w) = V(\tau_0,\tau_0,R_0;w) + (\tilde{\tau}-\tau_0)\frac{\partial V}{\partial\tilde{\tau}}\Big|_{\tilde{\tau}=\tau_0} + \frac{(\tilde{\tau}-\tau_0)^2}{2}\frac{\partial^2 V}{\partial\tilde{\tau}^2}\Big|_{\tilde{\tau}=\tau_0} \tag{A.9}$$

where $\frac{\partial V}{\partial\tilde{\tau}}|_{\tilde{\tau}=\tau_0} = 0$ 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,

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

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

$$\mathcal{I}(\sigma) = \frac{1}{2}\big(\log(2\pi e\hat{\sigma}^2) - \log(2\pi e\tilde{\sigma}^2)\big) = \frac{1}{2}\log\frac{\hat{\sigma}^2}{\tilde{\sigma}^2} \tag{A.10}$$

where $\frac{1}{2}\log(2\pi e\sigma^2)$ is the differential entropy (in bits) of a Gaussian distribution with variance σ^2. Therefore, in a Gaussian model, problem (A.7) becomes

$$\max_{\tilde{\sigma}\geq\hat{\sigma}} \quad \tilde{\sigma}^2\int\frac{\partial^2 y}{\partial\tilde{\tau}^2}\Big|_{\tilde{\tau}=\tau}\phi(\tau;\hat{\tau},\hat{\sigma})d\tau - \kappa\log\frac{\hat{\sigma}^2}{\tilde{\sigma}^2} \tag{A.11}$$

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). It shows that the perceived tax rate is $\tilde{\tau} = \xi s + (1-\xi)\hat{\tau}$ where $\xi \in [0,1]$ is a measure of the attention level such that

$$\xi = \max\left(0, 1 + \frac{\kappa}{\hat{\sigma}^2 \int \frac{\partial^2 y}{\partial \tilde{\tau}^2}\Big|_{\tilde{\tau}=\tau} \phi(\tau; \hat{\tau}, \hat{\sigma}) d\tau}\right) \tag{A.12}$$

at the optimum.

A.5 Proof of Corollary 1

Conditions (A.3) and (A.5) apply to any learning leading to a differentiable posterior distribution of perceptions $f_{\tilde{\tau}}(\tau|\tau_0; w)$ with positive support on $[0,1]$, where $\tau_0 = \tau_g + \vartheta$. Further insights may be gained by focusing on biased Gaussian equilibria (Definition 3). Indeed, in this case $f_{\tilde{\tau}}(\tau|\tau_0; w)$ is a Gaussian pdf $\phi(\tau; \mu, \sigma^2)$ with mean $\mu = \xi\tau_0 + (1-\xi)\hat{\tau}$ and variance $\sigma^2 = \sigma^{\star 2}$. We can thus express agents' responses to tax reforms in terms of changes in the true tax rate τ_0, changes in the prior mean $\hat{\tau}$ and induced changes in attention ξ that correspond to changes in the precision of the signal $\sigma^\star$. To do so, we use a first-order approximation of the objective at the mean μ and exploit the following Lemma.

Lemma 2. *Let $\psi(x)$ be a differentiable real-valued function, $\psi_a(x) = \psi(a) + (x-a)\psi'(a)$ its first-order Taylor approximation evaluated at a and $\phi(x; \mu, \sigma^2)$ the pdf of the Gaussian distribution with mean μ and variance σ^2. Then,*

$$\int_{\mathbb{R}} \psi_\mu(x)\phi(x; \mu, \sigma^2) dx = \psi(\mu) \tag{A.13}$$

$$\int_{\mathbb{R}} \psi_\mu(x) \frac{\partial \phi(x; \mu, \sigma^2)}{\partial \mu} dx = \psi'(\mu) \tag{A.14}$$

$$\int_{\mathbb{R}} \psi_\mu(x) \frac{\partial \phi(x; \mu, \sigma^2)}{\partial \sigma} dx = 0 \tag{A.15}$$

Proof. Equation (A.13) directly follows from $\int_{\mathbb{R}} (x-\mu)\phi(x; \mu, \sigma^2) = 0$ by definition of the mean. To prove equation (A.14), realize that $\int_{\mathbb{R}} \frac{\partial \phi(x;\mu,\sigma^2)}{\partial \mu} dx = 0$ and $\frac{\partial \phi(x;\mu,\sigma^2)}{\partial \mu} = \frac{x-\mu}{\sigma^2}\phi(x; \mu, \sigma^2)$ so that $\int_{\mathbb{R}} (\psi(\mu) + (x-\mu)\psi'(\mu)) \frac{\partial \phi(x;\mu,\sigma^2)}{\partial \mu} dx = \frac{\psi'(\mu)}{\sigma^2} \int_{\mathbb{R}} (x-\mu)^2 \phi(x; \mu, \sigma^2) = \psi'(\mu)$. Equation (A.15) follows from the fact that $\int_{\mathbb{R}} \frac{\partial \phi(x;\mu,\sigma^2)}{\partial \sigma} dx = 0$ such that the integral of a constant is nil and that $\frac{\partial \phi(x;\mu,\sigma^2)}{\partial \sigma}$ is symmetric such that the integral of x is also nil by a symmetry argument. □

Rewriting equation (A.3) as

$$E_\vartheta\Big[\int\Big\{\int\Big[-\frac{G'(\mathcal{V})}{p}y+y\Big]\phi(\tau;\mu,\sigma^2)d\tau \tag{A.16}$$
$$+\int\Big[\frac{G(\mathcal{V})}{p}+\tau_0 y\Big]\left(\frac{d\phi(\tau;\mu,\sigma^2)}{d\mu}\frac{d\mu}{d\tau_g}+\frac{d\phi(\tau;\mu,\sigma^2)}{d\sigma}\frac{d\sigma}{d\tau_g}\right)d\tau\Big\}f_w(w)dw\Big]=0$$

allows us to apply Lemma 2 to obtain with $\mu=\xi\tau_0+(1-\xi)\hat{\tau}$

$$E_\vartheta\Big[\int\Big\{\Big[-\frac{G'(\mathcal{V})}{p}y+y\Big]\Big|_{\tilde{\tau}=\mu} \tag{A.17}$$
$$+\Big[\Big(\frac{G'(\mathcal{V})}{p}(\tilde{\tau}-\tau_0)+\tau_0\Big)\frac{dy}{d\tilde{\tau}}\frac{d\mu}{d\tau_g}\Big]\Big|_{\tilde{\tau}=\mu}\Big\}f_w(w)dw\Big]=0$$

since taking $\psi(\tau)=\left[\frac{G(\mathcal{V})}{p}+\tau_0 y\right](\tau)$ implies $\psi'(\mu)=\left[\left(\frac{G'(\mathcal{V})}{p}(\tilde{\tau}-\tau_0)+\tau_0\right)\frac{dy}{d\tilde{\tau}}\right](\mu)$ by the modified envelope condition. Recall that $\mu=\xi\tau_0+(1-\xi)\hat{\tau}$. Now, in equilibrium we have by definition that $\hat{\tau}=\tau_g+b$ meaning $\mu=\tau_g+\xi\vartheta+(1-\xi)b$ and $\mu-\tau_0=(1-\xi)(b-\vartheta)$. Hence, in equilibrium,

$$E_\vartheta\Big[\int\Big\{\Big[-\frac{G'(\mathcal{V})}{p}y+y\Big]\Big|_{\tilde{\tau}=\tau_g+\xi\vartheta+(1-\xi)b} \tag{A.18}$$
$$+\Big[\Big(\frac{G'(\mathcal{V})}{p}(1-\xi)(b-\vartheta)+\tau_g+\vartheta\Big)\frac{dy}{d\tilde{\tau}}\frac{d\mu}{d\tau_g}\Big]\Big|_{\tilde{\tau}=\tau_g+\xi\vartheta+(1-\xi)b}\Big\}f_w(w)dw\Big]=0$$

Last, we characterize taxpayers' average response to tax reforms $\frac{d\mu}{d\tau_g}$ as computed under discretion and commitment. Under discretion, the policymaker takes agent's priors and thus attention strategies as given, hence $\frac{d\mu}{d\tau_g}=\xi\frac{d\tau_0}{d\tau_g}=\xi$ which yields equation (17). Under commitment, the policymaker internalizes the equilibrium condition that priors and thus attention strategies adjust to the tax policy such that $\frac{d\mu}{d\tau_g}=\xi\frac{d\tau_0}{d\tau_g}+(1-\xi)\frac{d\hat{\tau}}{d\tau_g}+\frac{d\xi}{d\tau_g}(\tau_0-\hat{\tau})=1+\frac{d\xi}{d\tau_g}(\vartheta-b)$ in equilibrium. This yields equation (18).

Transversality conditions follow from a direct application of Lemma 2 to equation A.5 with again $\mu=\tau_g+\xi\vartheta+(1-\xi)b$:

$$E_\vartheta\Big[\int\frac{G'(\mathcal{V})}{p}\Big|_{\tilde{\tau}=\mu}f(w)dw\Big]=1 \tag{A.19}$$

A.6 Proof of Proposition 4

Taking a small noise approximation, characterizations of equilibrium tax rates under discretion τ_g^{eq} and commitment $\tau_g^\star$ in this tractable Gaussian model write

$$\int \Big[\Big(1-\frac{G'(\mathcal{V})}{p}\Big)y + \Big(\frac{G'(\mathcal{V})}{p}(1-\xi)b + \tau_g^{\text{eq}}\Big)\frac{dy}{d\tilde{\tau}}\xi\Big]\Big|_{\tilde{\tau}=\tau_g^{\text{eq}}+(1-\xi)b} f_w(w)dw = 0$$
$$\int \Big[\Big(1-\frac{G'(\mathcal{V})}{p}\Big)y + \Big(\frac{G'(\mathcal{V})}{p}(1-\xi)b + \tau_g^{\star}\Big)\frac{dy}{d\tilde{\tau}}\Big(1-\frac{d\xi}{d\tau_g}b\Big)\Big]\Big|_{\tilde{\tau}=\tau_g^{\star}+(1-\xi)b} f_w(w)dw = 0$$

Assuming preferences are iso-elastic, $U(c,y;w) = c - \frac{(y/w)^{1+\varepsilon}}{1+\varepsilon}$, the elasticity of earnings with respect to the perceived marginal net-of-tax rate e is constant

$$\forall \tilde{\tau}, w, \quad e \equiv \frac{1-\tilde{\tau}}{y}\frac{dy}{d(1-\tilde{\tau})} = \frac{1}{\varepsilon} \iff \frac{dy}{d\tilde{\tau}} = -e\frac{y}{1-\tilde{\tau}} \tag{A.20}$$

Plugging in e we get

$$\int \Big[\Big(1-\frac{G'(\mathcal{V})}{p}\Big)y - \Big(\frac{G'(\mathcal{V})}{p}(1-\xi)b + \tau_g^{\text{eq}}\Big)e\frac{y}{1-\tilde{\tau}}\xi\Big]\Big|_{\tilde{\tau}=\tau_g^{\text{eq}}+(1-\xi)b} f_w(w)dw = 0$$
$$\int \Big[\Big(1-\frac{G'(\mathcal{V})}{p}\Big)y - \Big(\frac{G'(\mathcal{V})}{p}(1-\xi)b + \tau_g^{\star}\Big)e\frac{y}{1-\tilde{\tau}}\Big(1-\frac{d\xi}{d\tau_g}b\Big)\Big]\Big|_{\tilde{\tau}=\tau_g^{\star}+(1-\xi)b} f_w(w)dw = 0$$

To further simplify these formulas, we now make a small posterior bias approximation $|b(1-\xi)| << 1-\tau_g$. This allows us to use the approximation $\frac{1}{1-\tau_g-(1-\xi)b} \approx \frac{1}{1-\tau_g}$ and to assume $\frac{d\xi}{d\tau_g}b << 1$ to simplify some terms.[I] Defining social marginal welfare weights $g(w) \equiv \frac{G'(\mathcal{V})}{p}$ and the mean operator $\bar{x} = \int x(w)f(w)dw$ we get

$$\Big\{\overline{(1-g)y} - \frac{\tau_g^{\text{eq}}}{1-\tau_g^{\text{eq}}}\overline{y\xi}e - \frac{b}{1-\tau_g^{\text{eq}}}\overline{g(1-\xi)y\xi e}\Big\}\Big|_{\tilde{\tau}=\tau_g^{\text{eq}}+(1-\xi)b} = 0$$
$$\Big\{\overline{(1-g)y} - \frac{\tau_g^{\star}}{1-\tau_g^{\star}}\overline{y}e - \frac{b}{1-\tau_g^{\star}}\overline{g(1-\xi)ye}\Big\}\Big|_{\tilde{\tau}=\tau_g^{\star}+(1-\xi)b} = 0$$

which simplify to the compact sufficient statistics formulas

$$\tau_g^{\text{eq}} = \frac{\overline{(1-g)y}}{\overline{(1-g)y} + \overline{y\xi}\ e} - b\ \frac{\overline{g(1-\xi)y\xi}\ e}{\overline{(1-g)y} + \overline{y\xi}\ e} \tag{A.21}$$
$$\tau_g^{\star} = \frac{\overline{(1-g)y}}{\overline{(1-g)y} + \overline{y}\ e} - b\ \frac{\overline{g(1-\xi)y}\ e}{\overline{(1-g)y} + \overline{y}\ e} \tag{A.22}$$

where all endogenous quantities on the right hand-side of the equations are evaluated at respectively $\tilde{\tau} = \tau_g^{\text{eq}} + (1-\xi)b$ and $\tilde{\tau} = \tau_g^{\star} + (1-\xi)b$. In other words formulas are

[I] In our simulations we do check that $\frac{d\xi}{d\tau_g}$ does not take large values (it takes values between 0.2 and 1 in equilibrium) as a way to confirm the validity of this approximation.

expressed in terms of sufficient statistics evaluated at the optimum. These are exact when $b = 0$.

A difficulty in comparing τ_g^{eq} and $\tau_g^\star$ is that some right-hand side quantities are endogenous to the tax rate and thus evaluated at different tax rates. Because we only observe the discretionary equilibrium, we cannot use these formulas for an application to the US economy. To overcome this difficulty, we use a small taxation bias approximation such that quantities can be evaluated to a first-order approximation at the same tax rate. For small perception biases, the corrective terms are second-order and go in the same direction for both commitment and discretionary tax rates. They are thus not driving the difference between the two and we disregard them to derive the following simple sufficient statistics formula for the taxation bias

$$
\begin{aligned}
\tau_g^{\text{eq}} - \tau_g^\star &\simeq \frac{\overline{(1-g)y}}{\overline{(1-g)y} + \overline{y\xi}\ e} - \frac{\overline{(1-g)y}}{\overline{(1-g)y} + \overline{y}\ e} \\
&= \frac{\overline{(1-g)y}\left(\overline{(1-g)y} + \overline{y}\ e\right) - \overline{(1-g)y}\left(\overline{(1-g)y} + \overline{y\xi}\ e\right)}{\left(\overline{(1-g)y} + \overline{y\xi}\ e\right)\left(\overline{(1-g)y} + \overline{y}\ e\right)} \\
&= \frac{e\ \tau_g^{\text{eq}}\ \tau_g^\star}{\overline{(1-g)y}}\left(\overline{y} - \overline{y\xi}\right) = \frac{\overline{(1-\xi)y}}{\overline{(1-g)y}} e\ \tau_g^{\text{eq}}(\tau_g^{\text{eq}} - (\tau_g^{\text{eq}} - \tau_g^\star)) \\
&\simeq \frac{\overline{(1-\xi)y}}{\overline{(1-g)y} + \overline{(1-\xi)y}\ e\ \tau_g^{\text{eq}}}\ e\ (\tau_g^{\text{eq}})^2
\end{aligned}
\tag{A.23}
$$

We then use our simulations to confirm that our assumption that the sufficient statistics are evaluated at the same rate is not driving the results reported in Table 1. All the other assumptions made to derive the above sufficient statistics formula are then related to an assumption of small prior biases.

Online Appendix

(Not for publication)

This online appendix first discusses dynamic extensions to the stylized model. Second, it provides detailed information on the numerical simulations. Third, it gives all proofs and derivations for the extension to nonlinear taxation. Finally, we show how to incorporate income effects in the analysis.

B 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 online appendix conveys three important messages: (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) and long-run sophistication from the government is unlikely to entirely eliminate the taxation bias.

B.1 Dynamic learning equilibrium

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

$$\tilde{\tau}_t = \xi \tau_t + (1 - \xi)\hat{\tau}_t \tag{24}$$

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

$$\hat{\tau}_t = \tilde{\tau}_{t-1} \tag{25}$$

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. Given that taxpayers' prior $\hat{\tau}_t$ is given at the beginning of each period, this implies that first-order condition from the static problem still applies and, focusing on interior solutions, tax policy in period t is given by

$$\tau_t = \frac{1 - (1 - \xi)\tilde{\tau}_{t-1}}{\xi(1 + e)}. \tag{26}$$

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

$$\Delta\hat{\tau}_{t+1} = \xi(\tau_t - \hat{\tau}_t) \tag{27}$$

$$\Delta\tau_{t+1} = \frac{1}{\xi(1+e)} - \frac{\xi(2-\xi+e)}{\xi(1+e)}\tau_t - \frac{(1-\xi)^2}{\xi(1+e)}\hat{\tau}_t \tag{28}$$

where (27) describes the dynamic evolution of the prior and follows from (24) and (25), and where (28) describes the dynamic evolution of tax policy and follows from (24) and (26). The loci are

$$\Delta\hat{\tau}_{t+1} = 0 \iff \tau_t = \hat{\tau}_t \tag{29}$$

$$\Delta\tau_{t+1} = 0 \iff \tau_t = \frac{1 - (1-\xi)^2\hat{\tau}_t}{\xi(2-\xi+e)} \tag{30}$$

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

$$\tau^{eq} = \frac{1}{1 + \xi e} \tag{31}$$

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 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.

B.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

$$\begin{aligned} \max_{\{\tau_t\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t \tau_t Y(1-\tilde{\tau}_t) \\ s.t. \quad & \tilde{\tau}_t = \xi\tau_t + (1-\xi)\tilde{\tau}_{t-1} \end{aligned} \tag{32}$$

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 B.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

$$\mathscr{L}(\tau_t, \tilde{\tau}_t, \lambda_t) = \sum_{t=0}^{\infty} \beta^t \left[\tau_t Y(1-\tilde{\tau}_t) + \lambda_t(\tilde{\tau}_t - \xi\tau_t - (1-\xi)\tilde{\tau}_{t-1})\right] \tag{33}$$

The first-order conditions are

$$Y(1-\tilde{\tau}_t) - \lambda_t\xi = 0 \tag{34}$$

$$\tau_t \frac{\partial Y(1-\tilde{\tau}_t)}{\partial \tilde{\tau}_t} + \lambda_t - \beta\lambda_{t+1}(1-\xi) = 0 \tag{35}$$

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

$$\beta(1-\xi)\; Y(1-\tilde{\tau}_{t+1}) = \left(1 - \frac{\tau_t}{1-\tilde{\tau}_t}\xi e\right) Y(1-\tilde{\tau}_t) \tag{36}$$

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

$$\tau^{eq} = \frac{1 - \beta(1 - \xi)}{1 + \xi e - \beta(1 - \xi)} \tag{37}$$

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

$$\frac{1}{1 + e} \leq \tau^{eq} \leq \frac{1}{1 + \xi e}. \tag{38}$$

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

$$\tau^{eq} - \tau^{\star} = \frac{e(1 - \beta)(1 - \xi)}{(1 + e)[1 + \xi e - \beta(1 - \xi)]} \tag{39}$$

which means that its value is substantially smaller than in the static case. For instance, when $\xi = 0.75$, $e = 0.33$ and $\beta = 0.95$,[II] the steady state taxation bias would be of only 0.31 percentage points.

Yet, 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 $Y(1 - \tau^{\star})$ around τ^{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

$$\frac{e}{1 - \beta} \, \frac{\tau^{eq} - \tau^{\star}}{\tau^{eq}(1 - \tau^{eq})} \tag{40}$$

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 <hen considering alternative realistic values of the discount parameter β. 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 $\xi = 85\%$.

[II] $\beta = 0.95$ corresponds to a standard value in the macroeconomic literature for the discount rate of private agents when a time period refers to a year.

C Numerical simulations

Simulations are implemented using Matlab and the algorithm may be summarized as follows. We first estimate a log-normal distribution of skills that we extend with a Pareto tail. This distribution of skills is then binned into a discrete approximation. Second, we find the optimal policy of the government using an iterative routine. Starting with a guess for the optimal policy, we compute the optimal attention strategies and allocations in equilibrium (i.e. when the priors are adjusted). We then compute a new optimal policy given taxpayers' choices and iterate until convergence to a fixed point solution.

This appendix provides details on these different steps. We first present the calibration strategy for the skill distribution. Second, we explain how to solve for the optimal attention strategies and allocations for a given tax schedule. Finally, we discuss how the government's problem is solved in the linear tax setting before turning to the nonlinear case.

C.1 Skill distribution

Simulations require an exogenous distribution of skills $f_w(.)$. We fit the adjusted gross incomes from the 2016 March CPS data and focus on singles without dependent children to avoid the complexity of interrelated labor supply decisions within families. The parameters of the log-normal are 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 agents' first-order conditions for labor supply. We first use OECD data on 2016 labor taxes in the US and fit a linear tax schedule $\{\tau_{obs}, R_{obs}\}$.[III] Then, we impose a quasi-linear utility specification $u(c, y; w) = c - (y/w)^{1+\epsilon}/(1+\epsilon)$ with $e = 1/\epsilon = 0.33$ (Chetty, 2012).

[III] 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.

Assuming we are in a no bias equilibrium (i.e. rational expectation) such that agents' perceived tax rate coincide with the observed one τ_{obs}, this allows us to compute skills through $w = \left(y^{\epsilon}/(1-\tau_{obs})\right)^{\frac{1}{1+\epsilon}}$. We also use the estimated linear tax system $\{\tau_{obs}, R_{obs}\}$ together with the actual distribution of earnings to deduce an exogenous expenditure requirement E for the government budget constraint.

When considering biased perception in the welfare analysis (Table 2), the distribution of skills also depends on the bias and (endogenous) attention. We use an iterative method to simultaneously infer both given the fitted tax schedule and a level for taxpayers' average attention.

C.2 Taxpayers' behavior

Taxpayers' choices are presented in Section 4. For the simulations, we consider Gaussian implementation shocks. Under this assumption, the equilibrium prior distribution is Gaussian as well. Consequently, one may easily compute the attention parameter (ξ), income (y) and consumption (c) for each taxpayer. Given an attention cost κ, a marginal tax rate τ – that potentially varies for each individual – and an uncertainty parameter σ_{ϑ}, the attention strategy in equilibrium follows from equation (15). Gaussian integrals are approximated using Gauss-Hermite quadratures. Using an agent's first-order condition and budget constraint, we compute her income, consumption and utility for different signal realizations. These computations are made for each type of agent w. The demogrant R is computed from the government budget constraint.

C.3 Optimal linear tax

Unless stated otherwise, we assume throughout our numerical exercise that the social planner has a log objective $G(.) = \log(.)$.

In order to compute the optimal linear tax under discretion, we start with a guess $\tau_{g,0}$. Using this guess, we can deduce each taxpayer's attention strategy when the prior is adjusted to the guess $\hat{\tau}_0 = \tau_{g,0} + b$. We then consider this distribution of attention strategies as constant and use a Matlab optimization routine to find a new $\tau_{g,1}$ which maximizes social welfare for these attention strategies. We then update the prior $\hat{\tau}_1 = \tau_{g,1} + b$, recompute the attention strategies and re-optimize until convergence $|\hat{\tau}_i - \tau_{g,i+1}| \leq$

$1e^{-5}$. This method is intuitive and captures the essence of the discretionary policy: the government maximizes its objective taking attention strategies as fixed.

We also implement an alternative algorithm where instead of maximizing social welfare numerically we directly pick a new tax rate using the government FOCs in Proposition 1 under a small signals approximation. We find comparable equilibrium rates. Similarly, we compute the optimal policy under commitment using the FOCs in Proposition 2.

In the welfare analysis in Table 2 Panel A, we instead 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. We infer α so that the sufficient statistics formula in Equation (A.21) coincides with the observed income tax rate. We can then compute the counterfactual policies without information frictions and commitment, and use the calibrated Benthamite parameter to compute the respective welfares.

C.4 Optimal nonlinear tax

In order to compute the optimal nonlinear tax, we again use an iterative routine. We start with a guess – namely, a constant marginal rate – and iterate until convergence of the nonlinear tax schedule. We only present results for the unbiased equilibrium $b = 0$. We proceed in the same spirit as for the linear tax schedule:

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 deduce a new tax schedule
5. Repeat steps 1-4 until convergence.

To maintain the numerical stability of the algorithm we impose a slow adjustment of attention strategies ξ at each iteration. Indeed, marginal tax rates being 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.

D Proofs for the extension to non-linear taxation

We here provide the proofs on the monotonicity condition and Proposition 5 (ABC formula) of the main text.

D.1 Monotonicity

In this section, we demonstrate that the monotonicity condition is expected to hold for the quasi-linear and iso-elastic separable utility function that we consider in our simulations. For alternative specifications, we recommend to proceed using a guess-and-verify method. The latter is already implemented in our code and a warning is automatically displayed when the monotonicity does not hold ex post.

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

$$(FOC)_y : 1 - \tilde{\tau}_w - \frac{1}{w}\left(\frac{y}{w}\right)^{\epsilon} = 0 \tag{41}$$

Differentiating this equation with respect to w yields

$$\frac{\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon-1}\frac{dy(\tilde{\tau}_w; w)}{dw} = \frac{1+\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon} - \frac{d\tilde{\tau}_w}{dw} \tag{42}$$

Now – in expectation of the realization of the implementation shock ϑ – we also have $\tilde{\tau}_w = T_g'(y(\hat{\tau}_w; w)) + (1-\xi)b$ which allows us to get

$$\frac{d\tilde{\tau}_w}{dw} = T_g''(y(\hat{\tau}_w; w))\frac{dy(\hat{\tau}_w; w)}{dw} + \frac{d}{dw}\Big[(1-\xi)b\Big] \tag{43}$$

and we can show that

1. If agents correctly perceive marginal tax rates ($b = 0$), the equilibrium condition $\hat{\tau}_w = T_g'(y(\hat{\tau}_w; w)) + b$ becomes $\hat{\tau}_w = T_g'(y(\hat{\tau}_w; w)) = \tilde{\tau}$. We then have $\frac{dy(\hat{\tau}_w;w)}{dw} = \frac{dy(\tilde{\tau}_w;w)}{dw}$ such that plugging (43) with $b = 0$ into (42) the monotonicity condition boils down to

$$\frac{dy}{dw} = \frac{\frac{1+\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon}}{\frac{\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon-1} + T_g''(y)} \geq 0 \iff -T_g''(y) \leq \frac{\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon-1} \tag{44}$$

2. If agents exhibit a small perception bias ($b \approx 0$) such that we have $\frac{dy(\hat{\tau}_w;w)}{dw} \approx \frac{dy(\tilde{\tau}_w;w)}{dw}$ plugging (43) into (42) the monotonicity condition rewrites

$$\frac{dy}{dw} = \frac{\frac{1+\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon} - \frac{d}{dw}\Big[(1-\xi)b\Big]}{\frac{\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon-1} + T_g''(y)} \geq 0 \iff \begin{cases} -T_g''(y) \leq \frac{\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon-1} \\ \frac{d}{dw}\Big[(1-\xi)b\Big] \leq \frac{1+\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon} \end{cases} \tag{45}$$

where the equivalence comes from the fact that the other case in which we would have $-T_g''(y) \geq \frac{\epsilon}{w^2}\left(\frac{y}{w}\right)^{\epsilon-1}$ is infeasible.

Hence, the monotonicity condition will hold if the tax function $T_g(y)$ is sufficiently smooth such that its second derivative is bounded (in absolute value).

D.2 Proposition 5 (ABC formula)

We proceed with a tax perturbation approach in order to characterize the nonlinear tax schedule chosen under discretion and under commitment. Consider a tax schedule $T_g(.)$ and a reform that consists in a small increase $\Delta\tau^r$ in marginal tax rates in a small bandwidth of earnings $[y^r - \Delta y^r, y^r]$ and let us compute its impact on the government's objective (written in Lagrangian form)

$$\mathscr{L} = E_\vartheta\Bigg[\iint \Big\{G\Big(\mathcal{V}(\tilde{\tau}_w, T_0(.);\kappa, w)\Big) + p\Big(T_0(y(\tilde{\tau}_w; w)) - E\Big)\Big\}\ f_{\tilde{\tau}_w}(\tau|\tau_0; w)\ f_w(w)d\tau dw\Bigg] \quad (46)$$

where p is the multiplier associated to the government's budget constraint and is equal to the social marginal value of public funds at the optimum.

Impact of the reform For a given target tax schedule $T_g(.)$, the reform has

- a mechanical effect dM and a welfare effect dW that translate the lump-sum increase of $\Delta\tau^r\Delta y^r$ in the tax liabilities of agents $w \in [w^r, \infty[$ defined by $y(\tilde{\tau}_{w^r}; w^r) \equiv y^r$ where $E_s[\tilde{\tau}_w|\hat{\tau}_w] = \xi T_0'(y(\hat{\tau}_w; w)) + (1-\xi)\hat{\tau}_w$ with $T_0' = T_g' + \vartheta$
- a labor supply or behavioral effect dB that translates an increase $\Delta\tau^r$ in marginal tax rates that impacts the perceived marginal tax rates $\tilde{\tau}_w$ of agents $w \in [\hat{w}^r - \Delta\hat{w}^r, \hat{w}^r]$ defined by $y(\hat{\tau}_{\hat{w}^r}; \hat{w}^r) \equiv y^r$ and $y(\hat{\tau}_{\hat{w}^r}; \hat{w}^r - \Delta\hat{w}^r) \equiv y^r - \Delta y^r$

such that the total impact on the government's objective is

$$\frac{d\mathcal{L}}{p} = \frac{dM}{p} + \frac{dW}{p} + \frac{dB}{p} \quad (47)$$

with

$$\begin{aligned}\frac{dM}{p} + \frac{dW}{p} &= \int_{w^r}^{\infty} E_\vartheta\Bigg[\int \Big\{\Delta\tau^r\Delta y^r - \frac{G'(\mathcal{V}(w))}{p}\frac{\partial U}{\partial c}\Delta\tau^r\Delta y^r\Big\}\ f_{\tilde{\tau}}(\tau|\tau_0; w)\ d\tau\Bigg] f_w(w)\ dw \\ &= \int_{w^r}^{\infty} E_\vartheta\Bigg[\int \Big(1 - g(w)\Big)\Delta\tau^r\Delta y^r\ f_{\tilde{\tau}}(\tau|\tau_0; w)\ d\tau\Bigg] f_w(w)\ dw \quad (48)\end{aligned}$$

since we here have, holding $\tilde{\tau}_w$ constant,

$$d\mathcal{V} = \frac{d}{dc}\Big\{U\Big(\underbrace{y(\tilde{\tau}_w;w) - T_0(y(\tilde{\tau}_w;w))}_{c}, y(\tilde{\tau}_w;w);w\Big) - \kappa\mathcal{I}(\sigma^\star)\Big\}dc = -\frac{\partial U}{\partial c}dT_0 \quad (49)$$

and

$$\begin{aligned}\frac{dB}{p} &= \int_{\hat{w}^r-\Delta\hat{w}^r}^{\hat{w}^r} E_\vartheta\Big[\int\Big\{\frac{G(\mathcal{V}(w))}{p} + T_0(y(\tilde{\tau};w))\Big\}\ \frac{df_{\tilde{\tau}}(\tau|\tau_0;w)}{d\tau_g}\Delta\tau^r\ d\tau\Big]f_w(w)\ dw \\ &\approx E_\vartheta\Big[\int\Big\{\frac{G(\mathcal{V}(\hat{w}^r))}{p} + T_0(y(\tilde{\tau};\hat{w}^r))\Big\}\ \frac{df_{\tilde{\tau}}(\tau|\tau_0;\hat{w}^r)}{d\tau_g}\Delta\tau^r\ d\tau\Big]f_w(\hat{w}^r)\ \Delta\hat{w}^r \quad (50)\end{aligned}$$

since we here have, holding $\tilde{\tau}_w$ constant,

$$d\mathcal{V} = \frac{d}{d\tau_g}\Big\{U\Big(y(\tilde{\tau}_w;w) - T_0(y(\tilde{\tau}_w;w)), y(\tilde{\tau}_w;w);w\Big) - \kappa\mathcal{I}(\sigma^\star)\Big\}d\tau_g = 0 \quad (51)$$

Characterization of tax policy The optimality condition for the choice of tax policy $\frac{d\mathcal{L}}{p} = 0$ thus writes

$$\begin{aligned}E_\vartheta\Big[\int\Big\{\frac{G(\mathcal{V}(\hat{w}^r))}{p} + T_0(y(\tilde{\tau}_{\hat{w}^r};\hat{w}^r))\Big\}\ \frac{df_{\tilde{\tau}_{\hat{w}^r}}(\tau|\tau_0;\hat{w}^r)}{d\tau_g}\ d\tau\Big]\ \frac{f_w(\hat{w}^r)}{\frac{dy(\hat{\tau}_{\hat{w}^r};\hat{w}^r)}{dw}} \\ + \int_{w^r}^{\infty} E_\vartheta\Big[\int\Big(1-g(w)\Big)\ f_{\tilde{\tau}_w}(\tau|\tau_0;w)\ d\tau\Big]f_w(w)\ dw = 0 \quad (52)\end{aligned}$$

where we have simplified through by $\Delta\tau^r\Delta y^r$ noting that

$$y(\hat{\tau};\hat{w}^r - \Delta\hat{w}^r) \equiv y^r - \Delta y^r \quad \Longrightarrow \quad \Delta\hat{w}^r\frac{dy(\hat{\tau}_{\hat{w}^r};\hat{w}^r)}{dw} \approx \Delta y^r$$

Assuming we are in the tractable Gaussian case, the ex post (after learning) distribution of the perceived marginal tax rate is Gaussian $f_{\tilde{\tau}_w}(\tau|\tau_0;w) \sim \mathcal{N}(\mu_w,\sigma^2)$ with mean $\mu_w = \xi\tau_0 + (1-\xi)\hat{\tau}_w$ and variance parameter $\sigma = \sigma^\star$. Applying Lemma 2 we can thus rewrite the optimality condition as

$$\begin{aligned}E_\vartheta\Big[\Big[\Big\{\frac{G'(\mathcal{V}(\hat{w}^r))}{p}\Big(\tilde{\tau}_{\hat{w}^r} - T_0'(y(\tilde{\tau}_{\hat{w}^r};\hat{w}^r))\Big) \ + \ T_0'(y(\tilde{\tau}_{\hat{w}^r};\hat{w}^r))\Big\}\ \frac{dy}{d\tilde{\tau}}\frac{d\mu_{\hat{w}^r}}{d\tau_g}\Big]\Big|_{\tilde{\tau}=\mu_{\hat{w}^r}}\ \frac{f_w(\hat{w}^r)}{\frac{dy(\hat{\tau}_{\hat{w}^r};\hat{w}^r)}{dw}} \\ + \int_{w^r}^{\infty}\Big[1-g(w)\Big]\Big|_{\tilde{\tau}=\mu_w} f_w(w)\ dw\Big] = 0 \quad (53)\end{aligned}$$

where $\frac{d\mu_w}{d\tau_g} = \xi\frac{d\tau_0}{d\tau_g} = \xi$ under discretion since the government takes agents' priors as given whereas $\frac{d\mu_w}{d\tau_g} = \xi\frac{d\tau_0}{d\tau_g} + (1-\xi)\frac{d\hat{\tau}_w}{d\tau_g} + \frac{d\xi}{d\tau_g}(\tau_0 - \hat{\tau}_w)$ under commitment since the government internalizes that priors adjust to the choice of tax policy and are thus an endogenous object.

In addition, the Lagrange multiplier is – absent income effects – determined by the same transversality condition as before

$$E_\vartheta\left[\int_0^\infty \Big[1 - g(w)\Big]\Big|_{\tilde{\tau}=\mu_w} f_w(w)\ dw\right] = 0 \tag{54}$$

which can be obtained in a perturbation approach by computing the impact of a uniform lump-sum increase in taxes.

ABCD formula To obtain our ABCD formula from equation (53), let us introduce $e = \frac{1-\tilde{\tau}_w}{y(\tilde{\tau}_w;w)}\frac{dy(\tilde{\tau}_w;w)}{d(1-\tilde{\tau}_w)}$ and assume that the shock ϑ is small to use $E_\vartheta[\psi(\vartheta)] \approx \psi(E_\vartheta[\vartheta])$ regardless of function ψ's curvature such that $E_\vartheta[\psi(\tau_0)] \approx \psi(\tau_g)$. This yields

$$\begin{aligned}&\frac{T_g'(y(\mu_{\hat{w}^r};\hat{w}^r)) + g(\hat{w}^r)|_{\tilde{\tau}=\mu_{\hat{w}^r}}\Big(\mu_{\hat{w}^r} - T_g'(y(\mu_{\hat{w}^r};\hat{w}^r))\Big)}{1-\mu_{\hat{w}^r}} \\ &\qquad = \frac{1}{e\frac{d\mu_{\hat{w}^r}}{d\tau_g}|_{\tilde{\tau}=\mu_{\hat{w}^r}}}\frac{1}{y(\mu_{\hat{w}^r};\hat{w}^r)}\frac{\frac{dy(\hat{\tau}_{\hat{w}^r};\hat{w}^r)}{dw}}{f_w(\hat{w}^r)}\int_{w^r}^\infty \Big(1-g(w)|_{\tilde{\tau}=\mu_w}\Big)f_w(w)\ dw\end{aligned} \tag{55}$$

where $\mu_w = \xi T_g'(y(\hat{\tau}_w;w)) + (1-\xi)\hat{\tau}_w$ and $\frac{d\mu_w}{d\tau_g} = \xi\frac{dT_g'(y(\hat{\tau}_w;w))}{d\tau_g} = \xi$ under discretion since the government takes agents' priors as given whereas $\frac{d\mu_w}{d\tau_g} = \xi + (1-\xi)\frac{d\hat{\tau}_w}{d\tau_g} + \frac{d\xi}{d\tau_g}(T_g'(y(\hat{\tau}_w;w)) - \hat{\tau}_w)$ under commitment since the government internalizes that priors adjust to the policy rule and are thus an endogenous object.

Note that with a quasi-linear and iso-elastic separable utility function we have $y(\tilde{\tau}_w;w) = w^{1+\frac{1}{\epsilon}}(1-\tilde{\tau}_w)^{\frac{1}{\epsilon}}$ and $e = \frac{1}{\epsilon}$ such that

$$\frac{dy(\tilde{\tau}_w;w)}{dw} = \left(1+\frac{1}{\epsilon}\right)w^{\frac{1}{\epsilon}}(1-\tilde{\tau}_w)^{\frac{1}{\epsilon}} = \frac{1+e}{w}y(\tilde{\tau}_w;w) \tag{56}$$

Assuming small (or no) perception biases such that $\tilde{\tau}_w \approx \hat{\tau}_w$ and $\frac{dy(\tilde{\tau}_w;w)}{dw} \approx \frac{dy(\hat{\tau}_w;w)}{dw}$ yields

$$\frac{T_g'(y(\mu^{\text{eq}}_{\hat{w}^r};\hat{w}^r)) + g(\hat{w}^r)(1-\xi)b}{1-\tilde{\tau}^{\text{eq}}_{\hat{w}^r}} = \frac{1}{\frac{d\mu_{\hat{w}^r}}{d\tau_g}}\frac{1+e}{e}\frac{1}{\hat{w}^r f_w(\hat{w}^r)}\int_{w^r}^\infty\Big(1-g(w)\Big)f_w(w)\ dw \tag{57}$$

E Income effects

In this section of the Online Appendix, we illustrate how the (linear tax) model in the paper could be extended to account for income effects and accordingly characterize tax policy under discretion and commitment. We now have to account for the fact that the average posterior tax rate is no longer a sufficient statistics for taxpayers' earnings choices. This requires a mere reformulation of the initial problem without income effects: integration in the government's problem is now with respect to the signal distribution.

In order to introduce income effects, it will prove useful to slightly reformulate taxpayers' problem introduced in Section 4.. To this end, consider that there is a continuum of individuals at each skill w of size $f(w)$ and let $Y(\tau_0) \equiv \iint y(\cdot)\phi(s;\tau_0,\sigma^\star)dsdF(w)$ be the aggregate earnings. Then, because the government budget constraint is binding at the optimum, the demogrant writes $R(\tau_0) = \tau_0 Y(\tau_0) - E$ as the overall population remains of size one. Further, and given that a taxpayer's budget constraint binds ex post, consumption adjusts such that $c_0 = R(\tau_0) + (1-\tau_0)y$. Therefore, an agent's utility is $u(R(\tau_0) + (1-\tau_0)y, y)$ for a realization τ_0 and earnings choice y.

Given the above reformulation, the only uncertainty arises from the randomness in the realized tax rate. An individual therefore chooses the signal precision σ and income y to maximize her expected utility

$$\sup_{\sigma,y|s} \iint u(R(\tau) + (1-\tau)y, y; w)\phi(s;\tau,\sigma)\hat{q}(\tau)dsd\tau - \kappa\mathcal{I}(\sigma) \tag{58}$$

where admissible earnings policies for this individual's choice may depend on the signal s. Now, guess that the optimal attention strategy $\sigma^\star$ depends only on w, $\hat{q}(.)$, and κ. As a consequence, the optimal earnings choice $y(s,w;\sigma^\star,\hat{q}(.))$ now solves

$$\begin{aligned}\int [(1-\tau)u_c(R(\tau) + (1-\tau)y, y; w)& \\ +u_y(R(\tau) + (1-\tau)y, y; w)]f\big(\tau|s;\sigma^\star,\hat{q}(.)\big)d\tau &= 0\end{aligned} \tag{59}$$

where $f(\tau|s;\sigma^\star,\hat{q}(.)) = \frac{\phi(s;\tau,\sigma^\star)\hat{q}(\tau)}{\int \phi(s;\tau,\sigma^\star)\hat{q}(\tau)d\tau}$ from Bayes rule. Assume that a solution to equation (59) exists. In turn, it implies that

$$\sigma^\star(w,\hat{q}(.),\kappa) \quad = \quad \arg\sup_{\sigma} \iint u(R(\tau) + (1-\tau)y, y; w)\phi(s;\tau,\sigma)\hat{q}(\tau)dsd\tau - \kappa\mathcal{I}(\sigma) \tag{60}$$

thus confirming the guess on $\sigma^\star$ (when it exists). We can now define agents' indirect utility function

$$\mathcal{V}(s,\tau_0;w,\kappa,\hat{q}(.)) \equiv u(R(\tau_0) + (1-\tau_0)y, y; w) - \kappa\mathcal{I}(\sigma^\star) \tag{61}$$

Turning to the government problem, it requires a mere variation from (6)

$$\max_{\tau_g} \quad E_\vartheta\Big[\iint G\Big(\mathcal{V}\big(s,\tau_0;w,\kappa,\hat{q}(.)\big)\Big)\ \phi(s;\tau_0,\sigma^\star)\ f_w(w)d\tau dw\Big] \tag{62}$$

Note that the inner integration is now with respect to the signal distribution $\phi(s;\tau_0,\sigma^\star)$ and no longer with respect to the posterior distribution of perceived rates. This is because the perceived tax rate $\tilde{\tau}$ is no longer a sufficient statistics for earnings choices.

The first order condition for the target tax rate under discretion writes

$$E_\vartheta\Big[\int\Big\{\int\Big[G'(\mathcal{V})\frac{d\mathcal{V}}{d\tau_g}\phi(s;\tau_0,\sigma^\star)ds+\int G(\mathcal{V})\frac{d\phi(s;\tau_0,\sigma^\star)}{d\tau_g}ds\Big\}f_w(w)dw\Big]=0 \tag{63}$$

and the first order condition for the target tax rate under commitment writes

$$\begin{aligned}E_\vartheta\Big[\int\Big\{\int\Big[G'(\mathcal{V})\Big(\frac{d\mathcal{V}}{d\tau_g}+\frac{d\mathcal{V}}{d\hat{q}(.)}\frac{d\hat{q}(.)}{d\tau_g}\Big)\phi(s;\tau_0,\sigma^\star)ds \\ +\int G(\mathcal{V})\frac{d\phi(s;\tau_0,\sigma^\star)}{d\tau_g}ds\Big\}f_w(w)dw\Big]=0\end{aligned} \tag{64}$$

This characterizes tax policy under discretion and commitment in the presence of income effects. The key difference between the two equations is the fact that the commitment tax policy takes into account the adjustment in the prior $\frac{d\hat{q}(.)}{d\tau_g}$ whereas the discretion tax policy does not. This leads to a taxation bias.