Income Taxation with Elasticity Heterogeneity

John Sturm — Princeton
André Sztutman — Carnegie Mellon

August 25, 2023

Abstract

How should income taxes account for heterogeneity in elasticities of taxable income? We address this question with a test that passes if and only if there exists a weighted utilitarian planner for whom taxes are locally optimal. Our test incorporates standard sufficient statistics and a novel ingredient: the variance of elasticities conditional on income. Theoretically, we show that the test fails when these variances are sufficiently high. Empirically, we find they are indeed large in a panel of US tax returns. We thereby conclude, without taking a stance on redistributive preferences, that there are welfare-improving tax reforms.

Author contacts: sturmjohna@gmail.com and asztutma@andrew.cmu.edu. We are extremely grateful to Ivan Werning for many helpful comments, questions, and suggestions. We also thank Alberto Abadie, Daron Acemoglu, Kristoffer Berg, Arnaud Costinot, Peter Diamond, Dave Donaldson, Amy Finkelstein, Joel Flynn, James Hines Jr., Jon Gruber, Christian Moser, Jim Poterba, Ali Shourideh, Robert Townsend, and Charles Wolock for their comments and suggestions. We are also grateful to seminar participants at the MIT Macro and Public Finance Lunch, the European Winter Meeting of the Econometric Society, the NTA Annual Conference on Taxation, and at the North American Summer Meeting of the Econometric Society. Finally, we thank Jon Gruber, Emmanuel Saez, and Dan Feenberg for their advice about working with the University of Michigan Income Tax Panel.

1 Introduction

It is uncontroversial that different people respond to incentives differently. Tax responses are no exception. Indeed, heterogeneity in tax responsiveness is a consistent finding of the empirical literature on the elasticity of taxable income (ETI). These differences reflect variation in elasticities along many characteristics including gender, age, and sophistication [Blau and Kahn, 2007, Vere, 2011, Bastani and Waldenström, 2021].

This paper explores the possibility of efficiency-enhancing income tax reforms that differentially tax these heterogeneous households without directly conditioning taxes on their elasticities. Intuitively, if taxes could directly condition on household ETI, then one could easily reduce deadweight losses associated with income taxation by imposing lower taxes on more elastic households and higher taxes on less elastic households. While we assume taxes cannot explicitly condition on ETI, we show that a similar type of reform is possible whenever there is enough heterogeneity in ETI. In particular, a tax reform can improve welfare by sorting households into different parts of the tax schedule where they face different marginal taxes. The same ETI heterogeneity that provides an efficiency motive for differential taxation also facilitates this sorting because it causes households with the same initial income to respond to tax changes differently.

This idea—that ETI heterogeneity can open the door to tax improvements—may appear to contradict a well-known fact in the optimal taxation literature: To first order, only the mean elasticity at each income level affects a tax reform's revenue impact [Jacquet and Lehmann, 2015, Scheuer and Werning, 2018]. However, we show that a planner's *second-order condition*, which indicates whether the planner is indeed maximizing a social welfare function, depends intimately on the structure of household heterogeneity. The second-order condition captures how the effects of an infinitesimal variation in taxes *change* when one repeats that variation for a second time. One key source of such changes is that, due to each household's response to the first variation, it experiences the second tax variation at a new income level. Notably, if households with the same initial income respond to the first variation differently and the tax change differs across incomes, then the second variation can even have different effects on households who faced the same taxes ex-ante. In particular, we show that taxes themselves can be used as an instrument to simultaneously "sort" households on ETI and "extort" them by raising taxes disproportionately on the unresponsive.

Our theoretical results encapsulate these ideas in a simple, planner-agnostic "rationalizability test" for the tax schedule, formulated in terms of locally observable sufficient statistics. Our test passes if and only if *there exists a social planner* within a broad class for whom the tax schedule is optimal among all nearby schedules. While our test—by virtue of remaining agnostic about planner preferences—is relatively permissive, it is at the same time stricter than the familiar test for constrained Pareto efficiency. This is

because we impose some mild restrictions on planner preferences: We assume that the social planner maximizes a weighted sum of household utilities, where the weight on any household may be endogenous to the tax schedule so long as (a) it changes smoothly in that household's utility and (b) the average weight at each income is continuous in income. These assumptions are satisfied by most welfarist objectives studied in the optimal taxation literature. By contrast, a "Pareto planner"—who seeks to bring each household's utility above a threshold—changes households' welfare weights discontinuously at those thresholds. We show that this subtle distinction has important implications, giving rise to a further requirement that taxes need to satisfy, specifically, a new second-order condition.

Interestingly, we show that this condition is not only necessary but—when combined with a standard first-order condition—also locally sufficient for tax rationalizability. This is surprising because the condition is constructed by looking at only a small subset of possible deviations from a given tax schedule. However, we demonstrate that all other tax deviations have significant welfare effects on a finite mass of households, which implies that they are sub-optimal for a planner who is sufficiently biased toward the status quo. Our test therefore provides a complete characterization of local tax rationalizability.

Our analysis benefits in several ways from the approach of "testing" tax schedules by asking whether some preferences rationalize them instead of fixing preferences and "solving" for the optimal schedule. First, we can avoid embedding strong normative judgments into the positive analysis of tax systems. Second, we need not extrapolate locally observed sufficient statistics to non-local tax schedules where one does not know their true values. Third, in our particular setting, the version of the second-order condition that guarantees the existence of preferences that rationalize the tax schedule is dramatically simpler than the second-order condition for any specific utilitarian objective function.

In contrast to the existing literature (e.g. the first-order Pareto efficiency test of Werning [2007]), our second-order test depends not only on the mean ETI at each income level, but also on its variance, reflecting the scope for a planner to "sort and extort" households by elasticity. In light of the well-documented empirical variation in ETIs, this raises the question: "Is the heterogeneity in ETIs large enough that real world tax schedules can be improved?" The second part of the paper seeks to answer this question by evaluating our rationalizability test in a panel of US tax returns from 1979 to 1990. Doing so requires us to move beyond the existing literature on ETI estimation, which does not study the variance of elasticities conditional on income.

We estimate ETI variances with a number of empirical strategies: First, we provide a lower bound on these variances by estimating mean elasticities and group size as non-parametric functions of income for two different groups of taxpayers—those who claim a high total deductions amount and those who claim a low total deductions amount — and computing the variance of elasticities across those two groups at each income level.

Second, we take a less conservative structural approach, postulating and estimating a flexible distribution for elasticities, and then using it to obtain non-parametric point estimates for the variance of elasticities at each income level. Third, we validate these explicit estimates by providing reduced-form evidence for the sorting mechanism that underlies failures of the second-order condition. Specifically, we document that income-conditional ETIs change across years as predicted by the differential reshuffling of high- and low-elasticity households in response to observed tax changes.

Our preferred estimates have stark implications for income taxation. Although the first-order test passes, we find that our novel second-order test fails in every year of our sample. This implies it is impossible to rationalize the income tax schedule within the set of social planners we consider. Said differently, any social planner willing to make at least some minimal welfare trade-offs between households would prefer a different tax schedule. In this sense, a "free lunch" is available through tax reform. A conservative quantification exercise suggests that raising (lowering) top taxes by 20 percentage points results in yearly welfare gains equivalent to approximately \$3700 (\$600) per top earner.

Related literature Theoretically, we contribute to a long literature on the design of non-linear income taxation schedules when labor is supplied on the intensive margin and production is linear. Recent work by Werning [2007], Jacquet and Lehmann [2015], and Scheuer and Werning [2018] has shown that the so-called "ABC formula" of Diamond [1998]—a necessary first-order condition for optimality—extends to the case of multi-dimensional heterogeneity if one simply uses the average elasticities at each income level.[1] We complement this result by providing an additional necessary condition—a second-order condition—which, when combined with the first-order condition, is also sufficient. To our knowledge, we are the first to explicitly study the planner's second-order condition with respect to non-linear income taxes.[2]

The approach of "testing" a tax schedule rather than "solving" for optimal taxes is most similar to that of Werning [2007], Hosseini and Shourideh [2019], and Bierbrauer et al. [2020]. However, our notion of "rationalizability" departs slightly from the Pareto efficiency criterion these papers study by requiring the planner's utilitarian objective to be smooth with respect to each household's utility, as well as continuous across households of similar income levels. This small distinction leads to large differences in our conclusions. Namely, the main result of Bierbrauer et al. [2020] implies that, in our setting, the first-order condition of Werning [2007] is not only necessary but also locally sufficient for taxes to be Pareto efficient, even when one allows for multi-dimensional heterogeneity. By contrast, we show that a distinct second-order condition is necessary—and

[1] Also see a much earlier informal derivation by Saez [2001].

[2] Werning [2007] shows that when households vary along one dimension and their preferences have a certain functional form, the planner's problem is concave. This implies that the second-order condition *would hold*, were one to compute it.

together with the first-order condition, locally sufficient—in order for taxes to be rationalizable by a planner who values household utility in a smooth way. While our result does not apply to a Pareto planner—since her desire to make transfers to a given household changes discontinuously when that household's utility crosses a threshold—it does apply to arbitrarily accurate smooth approximations of any Pareto objective. Under any such smooth approximation, the planner becomes willing to tolerate a small welfare loss to any household if it is accompanied by a large enough gain for other households. Our results therefore illustrate a sense in which the Pareto efficiency criterion is knife-edge, falling apart once one allows arbitrarily small trade-offs to be made across households.

Closely related to our paper is Jacquet and Lehmann [2020], who highlight shifts in the elasticity-composition of the income distribution as a source of endogeneity in the average elasticity at any income. They provide analytical cases and numerical examples where accounting for heterogeneity leads to a larger gap between optimal and actual income taxes at high incomes.[3] We also emphasize these compositional shifts, but take a different approach. First, by testing existing tax schedules rather than solving for new optimal schedules, we emphasize a novel and distinct aspect of elasticity heterogeneity—that it can be so strong to prevent existing tax schedules from being rationalized by any (smooth) social preferences. Second, our results accommodate arbitrary multidimensional heterogeneity in preferences. The generality of our analytical results allows us to decompose the contributions of many empirically-relevant quantities—such as the shape of the income distribution and the curvature of the tax schedule—in mediating compositional effects. This generality also allows us to connect our test to the data without restrictive assumptions on the joint distribution of elasticities and productivities.

Our findings also complement other existing theoretical results in public finance and mechanism design. First, we share with the random taxation literature an emphasis on the non-convexity of certain taxation problems [Stiglitz, 1982b,a, Hines Jr and Keen, 2021]. While this literature shows that stochastic tax schedules can be welfare-enhancing in some circumstances,[4] we demonstrate that under similar—and testable—conditions, targeted deterministic deviations can provide an alternative way to improve outcomes. Second, our interest in targeting households according to their elasticities relates to a strand of the literature that proposes using multi-dimensional policy instruments to improve efficiency in the presence of multi-dimensional heterogeneity. Concretely, Moser and Olea de Souza e Silva [2019] argue that income-contingent retirement savings policies can be used to more efficiently screen on productivity when households differ in terms of both

[3]In a similar vein, Lockwood et al. [2020] discuss how accounting for a planner's uncertainty over households' taxable income elasticities affects optimal taxes. Although different in its motivation, we view Lockwood et al. [2020] as closely related to Jacquet and Lehmann [2020] since—when a planner faces an expected budget constraint—uncertainty over states of the world and heterogeneity across households are interchangeable for the purposes of income taxation.

[4]This possibility appears more generally in other mechanism design problems, see for example Maskin and Riley [1984a,b].

productivity and present bias. We show how a social planner can use a *single-dimensional* instrument to target households "as if" they could condition on untaxed characteristics.[5]

On the empirical side, we build on the approach of Gruber and Saez [2002] to estimate elasticities of taxable income (ETIs). We use the same NBER panel of tax returns as these authors and a similar identification strategy, identifying ETIs from households' changes in income following changes in taxes. Our work extends beyond Gruber and Saez [2002] in two ways. First, we estimate elasticities *conditional on income level* nonparametrically using local polynomial methods. Second, we place a greater emphasis on the *variance* of ETI across households. Of course, our interest in ETI heterogeneity has significant precedent in the literature: Gruber and Saez [2002] and Kopczuk [2005] show that itemizers are more elastic, and related work on labor supply elasticities suggests that ETIs may be higher for second-earners, single mothers, and those near retirement [Blau and Kahn, 2007, Eissa and Liebman, 1996, Vere, 2011]. More recent work by Kumar and Liang [2020] emphasizes heterogeneity in average ETIs *between income brackets* as a key source of variation across ETI estimates in the literature.[6] What distinguishes our approach to the variation in ETIs is an interest in a new, model-implied, sufficient statistic: the variance in ETI *conditional on income level*. We employ a number of identification strategies tailored to this statistic. Among these strategies is a structural procedure that relies on flexible parametric assumptions on the distribution of elasticities and on the linearity of tax responses. We combine this identification strategy with methods from the recent literature on discrete-approximation algorithms in econometrics [Bonhomme and Manresa, 2015, Bonhomme et al., 2017, Lewis et al., 2019].

Paper outline Section 2 presents a simple example that introduces the key ideas of the paper. Section 3 lays out the formal model. Section 4 presents and discusses our theoretical test for rationalizability. Section 5 presents our empirical estimates and evaluates our test using US tax return data. Section 6 presents a simple quantification of the welfare losses from un-rationalizable tax policy. Section 7 concludes.

[5] More generally, our work relates to the multidimensional screening and nonlinear pricing literatures. Whereas the multidimensional screening literature has emphasized direct mechanisms and revelation principle (e.g. Rochet and Stole [2003]), we apply the taxation principle and focus on the indirect mechanism [Hammond, 1979, Rochet, 1985, Guesnerie, 1998]. This approach allows us to sidestep challenging technical issues and derive new conditions for optimality. Our approach can easily be applied to nonlinear pricing problems, where our second-order condition would translate to a condition for the local convexity of revenues.

[6] Although we do not focus on inattention to tax changes, our estimation is related in spirit to Taubinsky and Rees-Jones [2018], who estimate heterogeneity in responses to sales taxes.

2 Motivating example

Before proceeding to the general model, we present a simple example that introduces the key ideas featured in our later formal results.[7]

A unit measure μ of households $h \in \mathcal{H}$ supply labor and consume a good—produced one-for-one with labor—in a static economy. Households face a non-linear tax schedule T. Each household h's preferences are additively separable between consumption and labor, feature a constant elasticity of labor dis-utility, and have no income effects:

$$V^h(T) \quad \equiv \quad \max_z \quad z - T(z) \quad - \quad \frac{z^{1+\frac{1}{\beta^h}}}{1+\frac{1}{\beta^h}} \Big/ (\theta^h)^{\frac{1}{\beta^h}}, \tag{1}$$

where $\theta^h > 0$ and $\beta^h > 0$ are productivity and elasticity parameters. We let $z^h(T)$ denote the maximizer. Within elasticity groups, productivity is Pareto with a common shape, i.e. $\theta^h \mid \beta^h \sim \text{Pareto}(\alpha > 1)$.

Now, consider the problem of a planner reforming a constant top tax rate $\bar{\tau}$ that applies to all income earned above income level $\bar{z}$.[8] Let $T_{\bar{\tau}}$ denote the entire schedule as a function of $\bar{\tau}$. We assume the planner places a constant weight $\widetilde{\lambda}$ on transfers to each household in the top bracket, relative to tax revenue, and therefore chooses $\bar{\tau}$ to maximize a weighted sum of welfare and tax revenue in the top bracket, or

$$\mathcal{L}(\bar{\tau}) \quad \equiv \quad \widetilde{\lambda} \cdot W_{\text{top}}(\bar{\tau}) + \text{Rev}_{\text{top}}(\bar{\tau}) \quad = \quad \mathbb{E}_h \left[\widetilde{\lambda} \cdot V^h(T_{\bar{\tau}}) \; + \; \bar{\tau} \cdot z^h(T_{\bar{\tau}}) \;\middle|\; z^h(T_{\bar{\tau}_0}) \geq \bar{z} \right], \tag{2}$$

Our assumptions imply that, within each elasticity group, the average post-reform income of households with pre-reform incomes in the top bracket is a power function of the top tax rate: $\mathbb{E}_h \left[z^h(T_{\bar{\tau}}) \;\middle|\; z^h(T_{\bar{\tau}_0}) \geq \bar{z},\ \beta^h = \beta \right] \;\propto\; (1-\bar{\tau})^{\alpha\beta}$.

A natural starting place for this analysis is the standard first-order condition for the revenue effects of a small increase in top taxes around its initial level $\bar{\tau}_0$.

$$\left.\frac{d}{d\bar{\tau}}\right|_{\bar{\tau}=\bar{\tau}_0} \mathcal{L}(\bar{\tau}) \quad \propto \quad \underbrace{-\widetilde{\lambda}}_{\text{Welfare effect}} \quad + \quad \underbrace{1}_{\text{Mechanical effect}} \quad - \quad \underbrace{\frac{\bar{\tau}_0}{1-\bar{\tau}_0}\, \alpha\, \mathbb{E}_{\text{top}}[\beta]}_{\text{Behavioral effect}}, \tag{3}$$

where $\mathbb{E}_{\text{top}}[\cdot]$ weights each elasticity proportionally to the total top-bracket earnings of households with that elasticity. The welfare effect captures welfare losses that result—by the envelope theorem—from the increased tax burden at households' initial incomes. The mechanical effect captures revenue gains as taxes increase at households' initial incomes, whereas the behavioral effect captures negative fiscal externalities as households reduce their incomes in response to higher taxes.

[7] Appendix B.1 walks through the algebra, which is straightforward.

[8] We also assume that (a) the planner is considering a tax *increase*—to ensure no responses by earners outside of the top bracket—and (b) the tax schedule is convex—which ensures intensive margin responses.

We later show that in the data, the sum of the two revenue effects is always positive on net: raising taxes increases tax revenues, i.e. we are on the correct side of the "Laffer curve." Importantly, this implies that there exists a welfare weight $\widetilde{\lambda} > 0$ that rationalizes the planner's first-order condition. To verify that we are not a local minimum, or—as we will say throughout—in a "Laffer valley," we check the second-order condition:

$$0 \quad \geq \quad \left.\frac{d^2}{d\bar{\tau}^2}\right|_{\bar{\tau}=\bar{\tau}_0} \mathcal{L}(\bar{\tau}) \quad \geq \quad \left.\frac{d^2}{d\bar{\tau}^2}\right|_{\bar{\tau}=\bar{\tau}_0} \text{Rev}_{\text{top}}(\bar{\tau}) \quad \propto \quad -(2-\bar{\tau}_0)\,\mathbb{E}_{\text{top}}[\beta] \;+\; \bar{\tau}_0\,\alpha\,\mathbb{E}_{\text{top}}[\beta^2] \tag{4}$$

Interestingly, this second-order test is sensitive to not only the mean elasticity of top earners, but also its variance $\mathbb{V}\text{ar}_{\text{top}}[\beta] = \mathbb{E}_{\text{top}}[\beta^2] - \mathbb{E}_{\text{top}}[\beta]^2$. And *we are in a "Laffer valley" (i.e.* (4) *fails) if—fixing the mean elasticity—elasticities vary enough.* This conclusion is normatively neutral, as the test does not depend on the welfare weight $\widetilde{\lambda}$.[9]

What economic mechanism underlies the planner ability to improve taxes when faced with heterogeneous elasticities? This is best illustrated by conceptualizing the second-order effects of a tax increase as the difference between the effects of two successive, infinitesimal tax increases. How does the first tax change shape the effects of the second? Here, the key idea is that *the behavioral effects of the first tax change determine the income level at which each household experiences the second tax change*, and therefore determine the marginal tax change it faces. For low-elasticity households—who by definition do not respond much to the first tax change—this has little bite. However, many high elasticity households exit the top bracket in response to the first tax change, and so are unaffected by the second one. Due to this differential sorting, *the first tax change decreases the effective elasticity of top earners; the planner's second increase in taxes exploits this.*

Is this theoretical consideration of any practical importance? Consider a back-of-the-envelope calibration: Suppose top marginal tax rates $\bar{\tau}_0$ are 50%, the Pareto shape parameter α is 2.5, and the average elasticity of the top earners is 0.3, consistent with the literature. Then the first-order condition holds if the planner places a weight $\widetilde{\lambda} = 0.25$ on transfers to top earners. The second order test (4) holds—so that taxes locally maximize rather than minimize revenue—if and only if $\mathbb{V}\text{ar}_{\text{top}}[\beta] \leq 0.27$. This is roughly the variance if, for example, three quarters of households have elasticity zero and one quarter have elasticity ≈ 1.2. This value for the variance is only slightly above a conservative lower bound we estimate in Section 5 based on differences in elasticities between light- and heavy-itemizers. This suggests that the planner's second-order condition may well be violated in practice.

The rest of the paper develops the ideas illustrated above in order to (a) accommodate

[9] The second inequality of (4) requires that second-order welfare effects are weakly positive. For the purpose of this example, this relies on our assumption that the planner's desire to make transfers to each top-earner is constant, not changing as their consumption or income. However, our later results show that this assumption is unnecessary when the planner uses more finely-tuned tax variations.

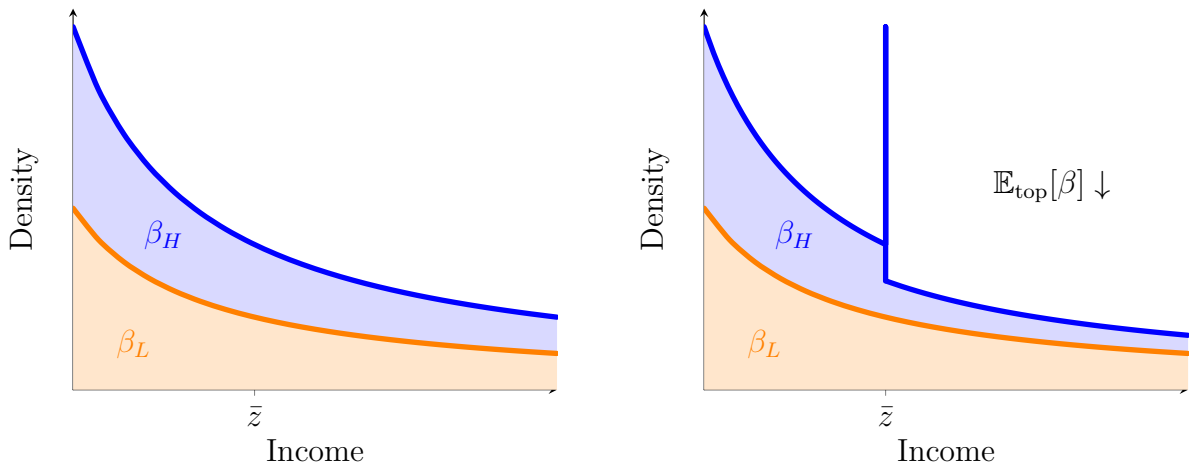

Figure 1: Income density of high- (blue) and low-elasticity (orange) households before (left panel) and after (right panel) taxes are increased. The compositional shift induced by an increase in taxes reduces the aggregate elasticity of top earners. (The vertical line in the right panel is bunching at a kink.)

general planner and household preferences, (b) provide an efficiency test that applies to the whole tax schedule and is both necessary and sufficient, and (c) formulate this test in terms of sufficient statistics that we evaluate in the data. The same key ideas—elasticity variance in the planner's second-order condition and the "sort and extort" motive—continue to play central roles in a general environment with non-linear income taxation.

3 Model

We study a standard, static Mirrlees model of income taxation, but allow for arbitrary preference heterogeneity. After laying out the model, we describe a number of mild regularity conditions used for our main results.

3.1 Environment

Time is static. There is a single consumption good and a single labor factor. Production is linear, and we normalize the price and the wage to one.

A unit measure μ of households $h \in \mathcal{H}$ supply labor z and consume subject to a non-linear income retention schedule to maximize utility u^h. For any income retention schedule $\widetilde{R} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, we define

$$z^h(\widetilde{R}) \equiv \arg\max_{z \geq 0} u^h(\widetilde{R}(z), z) \tag{5}$$

if the arg max exists, and let $V^h(\widetilde{R})$ denote the associated max.

We take as a primitive a particular income retention schedule

$$R : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \tag{6}$$

that maps each household's pre-tax income to its post-tax income, i.e. consumption. We denote labor supply, consumption, and utility at R by z_0^h, c_0^h, and V_0^h, respectively. Throughout, we will use "income retention schedule" and "tax schedule" interchangeably.

Throughout the paper, we study tax changes $\Delta : \mathbb{R}_{\geq 0} \to \mathbb{R}$ relative to R. Our main results focus in particular on very "small"—i.e. local—tax changes. Put simply, we call a tax change small if its level and derivatives are uniformly bounded by those of R, with a small bound. More formally, we define a norm on the space of tax changes[10] by

$$||\Delta|| = \sup \left\{ B \in \mathbb{R}_{>0} \;\middle|\; \forall z \in \mathbb{R}_{>0},\ |\Delta(z)| \leq B|R(z)| \text{ and } |\Delta'(z)|, \left|\frac{d\Delta'(z)}{d\log z}\right|, \left|\frac{d^2\Delta'(z)}{d\log z^2}\right| \leq B|R'(z)| \right\}. \quad (7)$$

We denote by $\mathbf{0}$ the zero tax change and by $B_\delta(\mathbf{0})$ the open ball of radius δ around $\mathbf{0}$.

3.2 Class of social planners

Throughout the paper, we study the problem of hypothetical social planners who assess welfare according to "generalized utilitarian" criteria and face tax revenue constraints.

Definition 1. *We call a level of government expenditure G and a collection of welfare-weighting functions $(w^h)_{h\in\mathcal{H}}$ with $w^h : \mathrm{Im}(u^h) \to \mathbb{R} \cup \{-\infty\}$ a* ***standard social objective*** *if there exists $\delta > 0$ such that, for all $h \in \mathcal{H}$, w^h satisfies the following conditions on the domain $V^h(R + B_\delta(\mathbf{0}))$:*

- *w^h is finite,*
- *w^h is twice-continuously differentiable,*
- *w^h is weakly increasing, and strictly increasing for a positive measure of $h \in \mathcal{H}$;*

and $w^h \circ V^h(\widetilde{R})$ and its first two Frechet derivatives in $\widetilde{R}$ are h-measurable within $R + B_\delta(\mathbf{0})$.[11]

Our main results rely on the notion that a social objective may "rationalize" the tax schedule R by providing a metric according to which it is optimal among all other (nearby) tax schedules.

Definition 2. *A standard social objective $((w^h)_{h\in\mathcal{H}}, G)$* ***locally rationalizes*** *R if there exists $\delta > 0$ such that*

$$R \in \underset{\widetilde{R}\in R+B_\delta(\mathbf{0})}{\arg\max} \int w^h \circ V^h(\widetilde{R}) \qquad s.t. \qquad \int \left[z^h(\widetilde{R}) - \widetilde{R}(z^h(\widetilde{R}))\right] d\mu \geq G. \quad (8)$$

[10] Our regularity assumptions ensure the derivatives used in this definition exist. In Appendix A.1 we define a Banach space of tax deviations on which $||\cdot||$ is defined and is a norm in the formal sense.

[11] These derivatives exist by w^h's differentiability, Assumption 2, and (for small enough δ) Lemma 2; see Appendix A.

The regularity conditions we impose for our formal results ensure that aggregate welfare and tax revenue always exist locally to R, so (8) is well-posed.

By working within the class of standard social objectives, we restrict our focus in a few important ways. First, the importance of any set of households to welfare is proportional to their measure. Second, welfare is additively separable across households; this rules out certain explicit forms of fairness concerns. Third, welfare is increasing in each household's utility. Finally, welfare is sufficiently smooth in each household's utility.

The class of standard social objectives nests most, but not all planners considered in the income taxation literature. One notable exception is a "Pareto planner" who prioritizes getting each household's utility above a (household-specific) "target utility." A Pareto planner's marginal desire to allocate additional utility to any particular household is discontinuous at its target utility for that household, whereas a planner with a standard social objective values this additional utility in a continuous way. Another exception is a "Rawlsian planner" who maximizes the utility of the worst-off household. Although such a planner lacks a standard social objective due to non-separability, one may think of the Rawlsian motive as a limit case of standard social objectives with increasingly concave weighting functions w^h.

3.3 Regularity conditions

Whereas many results in the optimal income taxation literature rely on specific function form assumptions, our theoretical results instead hold under much weaker regularity conditions. For readability, we summarize the main content of these assumptions below and state them formally in Appendix A.

Our regularity conditions serve four main purposes. First they ensure that welfare and tax revenue are well-defined and well-behaved (i.e. twice-continuously differentiable) in a neighborhood around the initial tax schedule R. The main conditions we impose to this end are that the initial tax schedule R is sufficiently smooth and has bounded elasticities, that—in a similar spirit—household utility is sufficiently smooth and generates income and compensated elasticities of labor supply that are bounded across households.

Second, our assumptions ensure that welfare and tax revenue can be expressed as integrals over household income z, facilitating integration by parts. To this end we assume the existence of sufficiently smooth functions for the income density $g(z)$ and income-conditional elasticity moments (e.g. the mean compensated elasticity of labor supply conditional on income z).[12]

Third—and of somewhat more qualitative interest—we assume that, local to the initial tax schedule R, all labor is supplied on the intensive margin. In doing so, we abstract

[12]While the necessity of our rationalizability test only relies on weak assumptions of this form, the sufficiency of our test additionally requires some mild conditions on the relationship between taxes, incomes, and conditional elasticity moments in the limits of zero and infinite incomes.

away from extensive margin decisions, such as labor force participation and migration. We view the extensive margin of labor supply as an interesting area for future work.

Fourth, we adopt a slightly stronger notion of a standard social objective, which we call a **standard, regular social objective**. This ensures the existence of income-conditional moments for several welfare statistics—such as the planner's desire to transfer income to households at income z—and imposes that these moments are sufficiently smooth in income. Our results use the assumption that average income-condition welfare weights are continuous in income in order to argue that a small decrease in the welfare of households earning z can be fully compensated (from the perspective of aggregate welfare) by a proportional increase in the welfare of $z + dz$ earners.

4 Rationalizability test for tax schedules

In this section, we present our main theoretical results. We provide a set of simple conditions on locally observable sufficient statistics—i.e. a "test"—that holds if and only if the tax schedule is locally rationalizable. Our test augments the first-order test of Werning [2007] with a novel, second-order condition.

After showing these conditions are necessary for rationalizability, we offer two interpretations: First, we illustrate variational arguments about the first- and second-order changes in welfare and tax revenue that tax change. Second, we explain the economic mechanisms at work, drawing particular attention to a novel "sort and extort" motive for taxation with multi-dimensional heterogeneity. Finally, we show that the conditions are also locally sufficient for rationalizability.

4.1 Necessary conditions

In order to state our first result, we will now introduce notation used throughout the paper. For all $z \in \operatorname{supp} g$, we denote by $\eta(z)$, $\varepsilon(z)$, and $\alpha(z)$ the average income and compensated elasticities[13] and the local shape parameter of the income distribution at income z.[14]

$$\eta(z) \equiv \mathbb{E}[\eta^h(R)|z_0^h = z] \qquad \varepsilon(z) \equiv \mathbb{E}[\varepsilon^h(R)|z_0^h = z], \qquad \alpha(z) \equiv -\frac{d\log(zg(z))}{d\log z} \tag{9}$$

Income elasticities capture behavioral responses to changes in the *level* of income retention, whereas compensated elasticities capture behavioral responses to changes in the *slope* of income retention. In addition to these familiar statistics, our results also depend

[13] We define these elasticities more formally in Appendix A.2.

[14] This is related to the well-known concept of the income distribution's "Pareto tail." The local shape parameter is a local version of this concept that coincides with the Pareto shape parameter if g is the density of a Pareto distribution.

on two less well-studied statistics. These are the income-conditional average compensated elasticity *squared* and an income-conditional average compensated *super*-elasticity $\varepsilon^{+h}(R)$ that captures how households' compensated elasticity change due to their labor supply responses to marginal tax changes.[15] We use the notation

$$\varepsilon^2(z) \equiv \mathbb{E}[\varepsilon^h(R)^2 | z_0^h = z], \qquad \varepsilon^+(z) \equiv \mathbb{E}[\varepsilon^{+h}(R) | z_0^h = z]. \tag{10}$$

Our first main result provides necessary conditions that constrain the relationships between these parameters and taxes at any locally rationalizable tax schedule. This test consists of a first order condition, which is standard in the literature, and a complementary second order condition, which is novel.

Theorem 1. *Suppose R is locally rationalized by some standard, regular social objective. Then for all $z \in \operatorname{supp} g$,*

$$0 \;\geq\; -1 + (1-R'(z))z\left[\frac{\eta(z)}{R(z)} \;+\; \left(\alpha(z) - \frac{d\log}{d\log z}\left(\frac{1-R'(z)}{R'(z)}\right) - \frac{d\log\varepsilon(z)}{d\log z}\right)\frac{\varepsilon(z)}{R'(z)z}\right]. \tag{ABC}$$

Moreover if (ABC) *does not everywhere hold with equality,*[16] *then for all $z \in \operatorname{supp} g$,*

$$0 \;\geq\; -\left(1+R'(z)\right)z\varepsilon(z) \;+\; \left(1-R'(z)\right)z\left[\alpha(z)\varepsilon^2(z) \;-\; \frac{d\varepsilon^2(z)}{d\log z} \;+\; \varepsilon^+(z)\right]. \tag{DEFG}$$

We label the first-order condition "ABC" in analogy to Diamond [1998]'s well-known "ABC formula", because they both derive from the planner's first-order condition. More precisely, (ABC) can be understood as a differentiated version of Diamond's formula. Analogously, we dub our second-order test "DEFG."

The first-order condition (ABC) and the second-order condition (DEFG) reflect two very different senses in which R is robust to changes in taxes. First, (ABC) says that the planner cannot lower taxes locally to z without reducing tax revenue, i.e. we are on the correct side of the Laffer curve. If this condition fails then—since it is possible to lower taxes at the same time as raising tax revenue—taxes are *Pareto inefficient* [Werning, 2007]. By contrast, a failure of the second-order test does not imply Pareto inefficiency but rather a slightly weaker condition: *non-rationalizability* within the class of planners we consider. The distinction is that, whereas the planner responds to a failure of (ABC) by only lowering taxes, the planner responds to a failure of (DEFG) at z with *a small decrease in taxes just above z and an equal and opposite increase in taxes just below z.*[17]

[15] See Appendix A.4.1 for a more formal definition.

[16] (ABC) does holds everywhere with equality if and only if R is a stationary point for tax revenue; this is generically false.

[17] The reason we use this particular deviation is that a planner's second-order condition is only informative in directions in which their Lagrangian is constant to first order (other directions are either infeasible or welfare-reducing). Our assumptions guarantee that the Lagrangian is constant to first order under this deviation because they guarantee that the marginal effects of tax changes around some income

Although this makes some households just below z worse off, we show that their decrease in welfare can be made arbitrarily small relative to the planner's revenue gains. This willingness to make welfare trade-offs—so long as the downside is small enough—is a key feature that distinguishes our notion of rationalizability from Pareto efficiency.

Another important contrast between the standard first-order condition and our new second-order condition is that the former only depends on *average* elasticities within each income level, whereas the latter depends on a higher moment of the distribution of elasticities. In order to evaluate (DEFG), one must—since $\varepsilon^2(z) = \varepsilon(z)^2 + \mathbb{V}\mathrm{ar}[\varepsilon^h(R)|z_0^h = z]$— consider the *variance* of compensated elasticities within each tax bracket. For this reason, the inclusion of multi-dimensional heterogeneity can have profound effects on the planner's problem. Namely, when the income density is falling quickly enough ($\alpha(z) > 0$)—as at high income levels—the planner's second-order condition fails whenever the variance in elasticities is sufficiently large. In this case, it is impossible to rationalize the tax schedule R within the broad class of social objectives we consider, even though the first-order (ABC) test known to the existing literature may pass.

4.2 Variational interpretation

In this section, we explain the first- and second-order conditions of Theorem 1 by illustrating the variations in taxes that they reflect. Intuitively, if R is rationalizable—i.e. optimal for some social planner—then, for any small variation in taxes, that planner's Lagrangian must be constant to first order in that variation and concave to second order.

The necessary conditions of Theorem 1 reflect these first- and second- order requirements for a *particular* type of tax change: a smooth increase in retention, narrowly concentrated between $z - dz$ and $z + dz$.[18] By "narrow," we mean that the amount this variation changes retention at its peak is much greater than the range of incomes over which it changes retention. Crucially, a typical household affected by a narrow tax change experiences a much greater change in marginal taxes than in the level of taxes. This simplifies our analysis by allowing us to ignore (a) curvature in the planner's desire to transfer income to any given household and (b) income effects, which become small compared to compensated effects.[19] Although our necessary conditions focus on narrow tax variations, we later show all other tax changes are undesirable for some planner, so that these necessary conditions are also sufficient for rationalizability.

level—both on household behavior and on welfare—are continuous both in (a) what that income level is and (b) the sign of the tax changes.

[18] As discussed below Theorem 1, our second-order condition couples this narrow increase in retention with a nearby, narrow decrease in retention, which ensures that the planner's Lagrangian is constant to first order. Below, we keep things simple by discussing the second-order effects of only the increase in retention. This is without loss because although the decrease in retention has *opposite first-order effects*, it has *identical second-order effects*.

[19] Formally, we do not "ignore" these effects. Rather, our proofs explicitly model them but show that, for sufficiently "narrow" tax variations, they are dominated by other considerations.

To begin the analysis, consider the first-order effects of a tax change concentrated at some income z, as in Figure 2. These effects can be decomposed, à la Saez [2001], into welfare effects, mechanical revenue effects, and behavioral revenue effects. Welfare effects simply reflect that—since we have weakly raised retention everywhere—households are better off as a result of the tax change. Mechanical revenue effects capture the losses in tax revenue that occur because, ignoring any change in household labor supply behavior, taxes are lower at each income level. Finally, behavioral revenue effects reflect that tax-change-induced changes in household behavior—through both income and compensate effects—have fiscal externalities on the planner whenever marginal taxes are non-zero.

Since the welfare effects are positive, revenue effects must be negative at any rationalizable schedule (otherwise the tax schedule can be improved). The RHS of (ABC) captures these revenue effects: The -1 corresponds to mechanical revenue losses and the remaining terms correspond to behavioral effects, which can increase or decrease revenue.

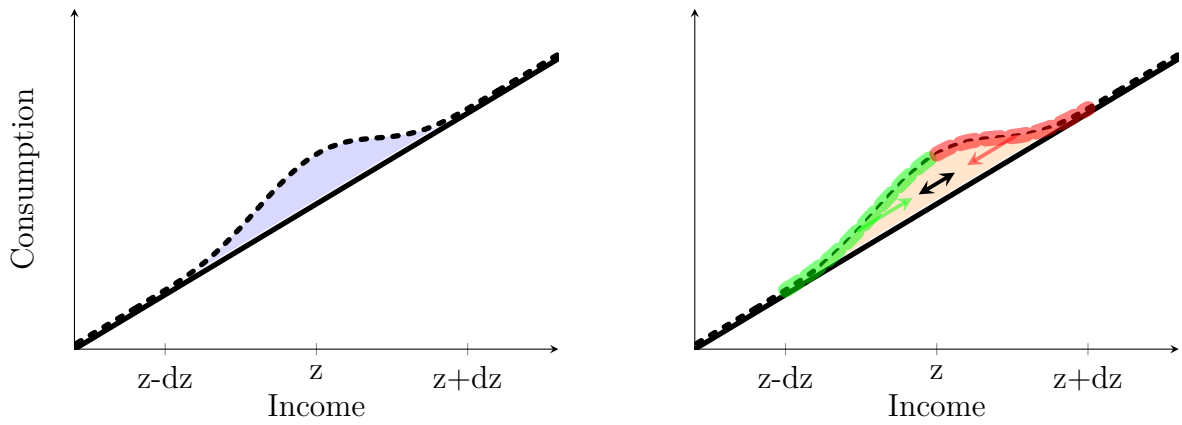

Figure 2: Left panel: An initial income retention schedule, before (solid) and after (dashed) a small variation in taxes. The shaded blue area represents the size of the mechanical effect and (if appropriately weighted) the welfare effect. Right panel: Behavioral effects due to a the same variation in taxes. Behavioral income effects (black arrow) act on households whose level of income is affected (shaded orange area), and can have positive or negative sign. Behavioral compensated effects increase the incomes of households whose marginal taxes decrease (green), and vice-versa (red).

We now consider the *second-order effects* of the same variation in taxes we have already studied to first order. The second derivative of the planner's Lagrangian captures how its first derivative *changes* as the same variation is done more or less. In other words, we ask, "how do the effects of an infinitesimal variation in taxes differ when we do it for a second time?" If R is optimal for some planner, then these second-order effects must be negative on net, i.e. the Lagrangian must be locally concave. In order to compute the Lagrangian's local curvature, we now walk through how each component of its first derivative with respect to the narrow variation studied above—i.e. the welfare, mechanical, and welfare effects—change when we repeat that variation for a second time.

To begin, consider how the welfare effect changes when our tax variation is repeated for a second time. For a narrow tax variation, the only second-order effect on welfare is as

follows: When taxes change for a second time, the amount of post-tax income transferred to each household may differ the amount transferred during the first tax change, since—due to its behavioral response to the first tax change—that household's income is different than it was during the first tax change. Since these first-order behavioral effects shift all households with incomes in $[z - dz, z + dz]$ toward z, where the change in retention is higher, the second tax variation transfers more post-tax income to each household than did the first; see Figure 3. For this reason, the second-order effect on welfare is positive. Rather than quantifying the precise size of this convexity, our second-order test captures its sign with an inequality. This expresses that—as the change in welfare effects is always positive—the change in (mechanical and behavioral) revenue effects must be negative in order for the planner's Lagrangian to be concave on the whole. The test therefore simply states that tax revenues must be concave with respect to narrow tax variations.

We now turn to these revenue effects. We start with our tax variation's mechanical effects, asking how they change when the variation is repeated for a second time. Conveniently, this change is simply the flip side of the second-order welfare effect considered above: When taxes change for a second time, each household receives a greater transfer than it did the first time, since—due to its behavioral response to the first tax change—that household's income has moved toward z, where the variation increases retention the most. This effect contributes a term of $-R'(z)z\varepsilon(z)$ to (DEFG).

Just as the second tax variation's mechanical effect changes due to households' behavioral response to the first variation, so too does the second tax variation's behavioral effect change due to the mechanical effects of the first variation. Intuitively, when taxes change for a second time, each household's behavioral effect on revenue is smaller, since it moves to a new income at which—due to the first variation—retention is higher, i.e. taxes are lower. This effect contributes a second term of $-R'(z)z\varepsilon(z)$ to (DEFG), a manifestation of Hessian symmetry

The tax variation's second-order behavioral revenue effects are also shaped by its first-order effects on household behavior. We divide these "behavioral-affecting-behavioral" effects into three groups: (a) one effect proportional to elasticities, (b) one key effect proportional to elasticity-squared, and (c) two small "correction" terms. The elasticity-proportional effect captures the following mechanism, which leads to less revenue-advantageous changes in behavior: Where the first variation has increased marginal retention, households respond less to the second tax change since the change in log marginal retention is smaller,[20] and vice-versa. This dampens positive labor supply responses and amplifies negative labor supply responses, contributing a factor of $-(1 - R'(z))z\varepsilon(z)$.

We now turn our attention to a key effect of interest, one proportional to the average square of elasticities at z. The dependence of this effect on a higher moment of the elasticity distribution means that it—unlike the previous effects discussed—can have a

[20]This is true to the extent that household elasticities remain constant; see later discussion.

qualitatively different effect when elasticities vary conditional on income. The effect is as follows: Because households adjust their incomes in response to the first tax change, they experience the second tax change at a different income level, where it may be larger or smaller. Moreover, how many households relocate to each different income level—and so how many households experience greater or smaller increases in taxes during the second tax change—depends on how many households start at each income level, i.e. on the local shape of the income distribution. For example, suppose the income density is sharply decreasing locally to z. In this case, the amount of income earned by households who—due to their behavioral responses to the first tax change—face larger marginal tax decreases or smaller marginal tax increases during the second tax change than the first is greater than the income earned by households who face smaller tax decreases or larger tax increases during the second tax change. This is because the former set of households are those inhabiting $[z - dz, z - \frac{dz}{2}]$ and $[z, z + \frac{dz}{2}]$, which contains more mass than the complementary range of incomes when the income density is sharply decreasing; see, e.g. Figure 3. Since the larger first group responds with more advantageous behavioral effects to the second tax change than the first—whereas the second group does the opposite but is smaller—the second tax change has a more positive behavioral effect on revenue than the first. If, on the other hand, the income density is increasing around z, the effect flips. In total, this effect accounts for the term $(1 - R'(z))z\alpha(z)\varepsilon^2(z)$ in (DEFG). $\alpha(z)$ reflects the shape of the income density locally; $\varepsilon^2(z)$ reflects that households' elasticities determine not only how much they relocate different income levels (and so different tax changes) but also how strongly they respond to those changes.

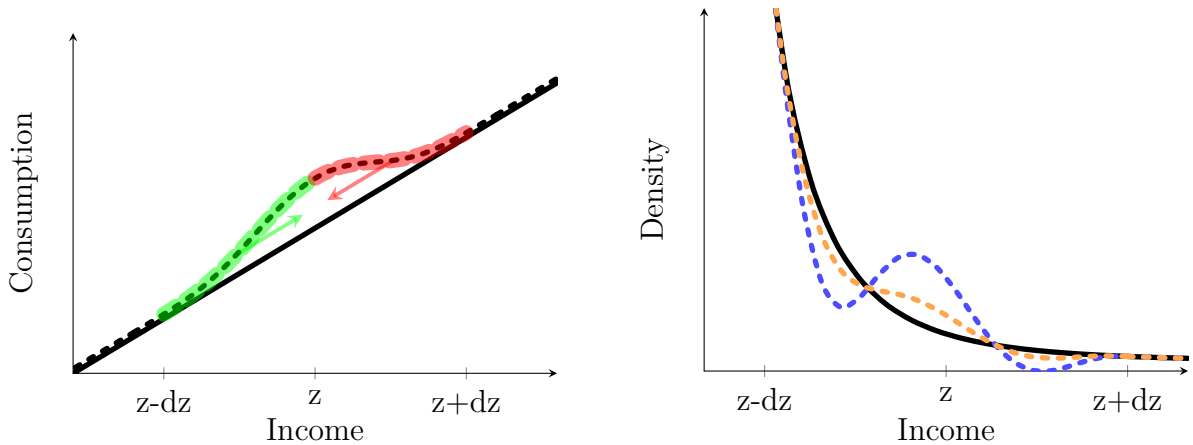

Figure 3: Left panel: Changes in behavior through compensated effects following the first tax change. Right panel: Initial income density (black) and income density following the first tax change for high- (blue) and low-elasticity (orange) households. Each household adjusts its income toward z, but the change in the income density depends on the shape of the initial income density and the distribution of elasticities.

Finally, two additional "corrections" to these second-order behavioral effects account for the final two terms of (DEFG). We emphasize these terms less because—to the extent

we can estimate them—they are quantitatively unimportant, and because they are zero in many parameterizations used in the literature. The first adjusts the effect described in the previous paragraph to account for the possibility that the distribution of elasticities conditional on income may vary with income locally to z. The second captures the fact that, due to changes in the curvature of household preferences and/or the tax schedule, households may have different elasticities during the second tax change than during the first. To the extent we can estimate these effects, they are small (see Section 5.3.2). However, we do not estimate the component of these "super-elasticities" stemming from changes in curvature of household preferences; this is an interesting area for future work.

4.3 Exploring the economic mechanism

While the analysis of Section (4.2) explains the various terms of our second-order test (DEFG) mathematically, we would also like to develop a more intuitive and economic understanding of it. Indeed, whether the test can fail in any reasonable description of the economy is an important question, considering that Werning [2007] shows the first-order condition (ABC) is both necessary and sufficient for global Pareto efficiency in a certain one-dimensional setting.

In this section, we show under what conditions (DEFG) fails and, when it does, provide an intuitive explanation of how the planner can improve the tax schedule. Our main focus is on the way that the test fails when—due to multi-dimensional household heterogeneity—elasticities vary widely within income levels. Here, we emphasize a novel "sort and extort" procedure through which the planner can increase tax revenue by using a first tax change to shift the elasticity-composition of the income distribution and a second tax change to exploit the separation of high- and low-elasticity households with differential taxes. Finally, we also explain how the test can fail even when households differ along single dimension of heterogeneity (i.e. productivity).

4.3.1 Multi-dimensional heterogeneity: "sort and extort"

The most natural and empirically relevant case in which (DEFG) fails is when there is significant variation in ETIs within income levels. Since one dimension of heterogeneity in household preferences is required to rationalize income differences, within-income differences are only possible with multi-dimensional heterogeneity. These within-income differences in household ETI point to an important limitation faced by the social planner. In general, a social planner might like to tax different households with the same income differently, either because she prefers some households to others or because she anticipates that they may respond to taxes differently. For example, a revenue-maximizing planner levies high taxes on low-ETI households and low taxes on high-ETI households, all else equal. However, when she cannot condition taxes explicitly on these elasticities,

the planner is forced set a single tax schedule that—at each income level z—must balance her desire to tax different z-earners differently. In this sense, the planner's ability to *target* taxes is limited by the fact that her tax instrument is lower-dimensional than the space of households she taxes with it.

This constraint—that the planner must set a single tax schedule for households who vary within income levels—is at the core of failures of the second-order condition. To see why, consider the simple case where households can be partitioned into a finite set of groups $i \in I$, each satisfying our regularity conditions in isolation. We show in Appendix C.3 that in this case the *aggregate* first- and second-order tests are simply the averages of the *within-group* first- and second-order tests. That is, if we let $\Pi^i_{ABC}(z)$ and $\Pi^i_{DEFG}(z)$ represent the right-hand sides of (ABC) and (DEFG) when expectations are taken over households within each group i, then our test can be expressed, for each $z \in \operatorname{supp} g$ as:

$$\mathbb{E}\left[\Pi^{i(h)}_{ABC}(z)\middle| z^h_0 = z\right] \;\leq\; 0 \qquad \text{and} \qquad \mathbb{E}\left[\Pi^{i(h)}_{DEFG}(z)\middle| z^h_0 = z\right] \;\leq\; 0. \tag{11}$$

Moreover, if preferences within each group i are concave and additively separable over consumption and labor and vary only in a labor dis-utility shifter, one may show that[21]

$$\Pi^i_{ABC}(z) \;\leq\; 0 \qquad \Longrightarrow \qquad \Pi^i_{DEFG}(z) \;\leq\; 0. \tag{12}$$

Together, these two observations imply that the second-order test can only fail if the first-order test fails for at least one group. That is, all failures of (DEFG) can be attributed to the fact that for at least one group, R is on the wrong side of the Laffer curve.

In a world with group-specific taxes, the planner could address the fact that one group is on the wrong side of the Laffer curve by simply lowering its taxes. However, this option is not available to the planner we consider: Any desirable decrease in taxes for one group requires a (potentially) undesirable decrease in taxes for another group. Indeed, if group-specific failures of the first-order condition are weak enough, it is possible for the second-order condition to hold still and for taxes to be rationalizable. In this context, the key insight of Theorem 1 is as follows: *If group-specific failures of first-order conditions are so severe that the aggregate second-order condition fails, then the planner can improve taxes with a special, "as-if" group-specific tax change.*

The rest of this section explains how the tax variation underlying (DEFG)—the one explored more mechanically in Section 4.2—approximates group-specific taxation. A first step is to recall that since this tax variation is arbitrarily "narrow", it has arbitrarily small welfare effects and arbitrarily small revenue effects through income elasticities. This implies that all of its impacts operate through compensated elasticity effects.[22] In

[21] See Appendix C.3. This fact is consistent with Werning [2007]'s observation that the dual of the planner's problem is convex when there is one such group i.

[22] These compensated elasticity effects are—as illustrated for the first order variation at the beginning

particular, the sense in which our tax variation approximates group-specific taxation is by effectively lowering taxes more for high-elasticity households and raising taxes more for low-elasticity households.

Our tax variation accomplishes this "as-if" elasticity-dependent taxation with a two-step change in taxes, captured by the first and second derivatives of the planner's Lagrangian. To first-order, changing taxes causes households to change their labor supply behavior as depicted in Figure 3. Low-elasticity households end up with roughly the same income density after an infinitesimal change in taxes, while high-elasticity households end up with a very different density, one which depends both on their initial density and the shape of the tax change itself. Importantly, we show that when the initial density is declining steeply, high-elasticity households are disproportionately drawn toward incomes in two regions—one where the variation decreases taxes by more and one where it increases taxes by less. That is, the variation's first-order effect is to "sort" households differentially, by elasticity. The second derivative of the planner's Lagrangian asks how the first-order effects change when the variation is performed a second time.[23] Critically, these effects improve as the variation is performed a second time, because they now operate on households who have been sorted in just the right way to generate positive behavioral effects on revenue: High elasticity-households are disproportionately represented where taxes decrease and low-elasticity households are disproportionately represented where taxes increase. That is, the planner exploits the sorting of households by "extorting" only the least responsive households with high taxes, *as if* she could tax them differentially.

Quantitatively, this sort-and-extort mechanism accounts for a factor of $(1-R'(z))z\alpha(z)$ $\mathbb{V}\mathrm{ar}[\varepsilon^h(R)^2|z^h(R)=z]$ in our second-order test, (DEFG). We show in Appendix C.3 that in the simplest natural case—when household preferences are linear in consumption and CES in labor supply, and when marginal taxes and the mean and variance of income-conditional elasticities are constant local to z—this term is the only factor that can cause (DEFG) to fail when (ABC) holds. That is, if the first-order test holds at some income z where marginal taxes and the income shape parameter $\alpha(z)$ are positive, then the second-order test fails there if and only if—holding average elasticities at z constant—there is enough heterogeneity in these elasticities. In other words, the second-order test binds if and only if there is enough scope for a sort-and-extort tax reform.

4.3.2 One-dimensional heterogeneity: "shift and exploit"

Although the most economically interesting and empirically plausible failures of (DEFG) derive from within-income household heterogeneity, failures are also possible in models

of Section 4.2—moderated by the local shape parameter $\alpha(z)$. Although our test accounts for cross-group differences in this shape, we keep this discussion simple by assuming that the shape parameter is constant across groups i.

[23] Because the variation is narrow, these effects can be simplified to compensated effects on revenue.

where households are homogeneous conditional on income. The main idea behind this possibility is similar to the "sort and extort" mechanism explored in the previous section: The planner can change elasticities at each income level with a first tax variation and then exploit this change by repeating the variation. The difference relative to the heterogeneity case is that elasticity changes come not through sorting but rather through "shifts" – either in the elasticity of individual households or in the identity of which households (each having fixed elasticities) are represented at each income level.

Notably, it is impossible for a planner to "shift and exploit" households when utility satisfies the functional form studied by Werning [2007].[24] This functional form implies that all variation in compensated elasticities across income levels is due to differences in income and consumption levels at which households' preferences are evaluated, rather than differences in the curvature of their preferences at a given level of income and consumption. It therefore prevents the planner from using a small tax variation to shift local elasticities, since—for any income level z—the households who each z after a tax change have the same post-tax change elasticity as the household who previous earned z had before the tax change. In Appendix B.2, we provide examples showing how—outside of this preference class—(DEFG) can fail with one-dimensional preference heterogeneity.

4.4 Sufficient conditions

So far, we have provided a novel set of necessary conditions that any locally rationalizable tax schedule must satisfy. However, we derived our test using only very "narrow" variations in taxes. So tax schedules which pass the (DEFG) test might still allow for other sorts of improvements. Our second main result says that this is not the case: The necessary conditions of Theorem 1 are also sufficient for local rationalizability.

This sufficiency result relies on a few mild assumptions beyond those required for Theorem 1. Aside from the additional regularity conditions in Assumption 6 and the condition—stated in the theorem itself—that some households at each income level have small enough elasticities, we also require that (ABC) and (DEFG) hold in a slightly stronger sense. Finally, our sufficiency result uses a slightly weaker notion of local optimality than the necessity result.[25]

We can now present Theorem 2, which shows that our characterization of locally rationalizable tax schedules in Theorem 1 is tight: If (ABC) or (DEFG) fail, the schedule sub-optimal for all planners; if they pass, the schedule is optimal for some planner.

Theorem 2. *Suppose that*

[24] This functional form nests additive CES preferences, which is perhaps the most commonly analyzed example in the literature.

[25] Even without this weaker notion, there exists—provided that R satisfies the test—a welfare function that, for any $\Delta \in \mathbf{\Delta}$, prefers R to $R + \epsilon\Delta$ for all $\epsilon > 0$ below some $\epsilon_\Delta > 0$. Our weaker notion of local optimality slightly restricts $\mathbf{\Delta}$ so as to ensure ϵ_Δ is uniform across all remaining tax changes Δ.

- (ABC) *and* (DEFG) *hold by amounts* $\Pi_{ABC}(z)$ *and* $\Pi_{DEFG}(z)$ *that satisfy:*
  - *For all* $z \in \operatorname{supp} g$, $\Pi_{ABC}(z) > 0$.
  - *There exist* $\overline{b}_c, \overline{b}_z > 0$ *such that* $\Pi_{DEFG}(z) \geq \overline{b}_c R(z) + \overline{b}_z z$ *for all* $z \in \operatorname{supp} g$.
  - *There exists* M *such that* $z|\Pi'_{ABC}(z)| \leq M|\Pi_{ABC}(z)|$ *for all* $z \in \operatorname{supp} g$.
- *For some sufficiently small* $\epsilon > 0$, $\mathbb{P}[\varepsilon^h(R) \leq \epsilon | z_0^h = z] > 0$ *for all* $z \in \operatorname{supp} g$.[26]

Then there exists a standard, regular social objective that—for any $M > 0$*—locally rationalizes* R *within the sub-space of deviations*

$$\mathbf{\Delta}_M^* \equiv \left\{ \Delta \in \mathbf{\Delta} \;\middle|\; ||\Delta|| \leq \; M \left[\int g(z) \left(R(z) + z \right) \left(\left(\frac{\Delta(z)}{R(z)} \right)^2 + \left(\frac{\Delta'(z)}{R'(z)} \right)^2 \right) dz \right]^{\frac{1}{2}} \right\}. \tag{13}$$

Considering that (DEFG) reflects only the robustness of R to very "narrow" deviations in taxes, this is a surprising result. Indeed, given a fixed social objective, we know that the planner must consider the robustness of R to all possible deviations. Why do we not need additional conditions in order to rule out these other second-order deviations?

The main insight is as follows: *For an arbitrary planner*, one must indeed verify additional second-order conditions corresponding to variations in taxes not spanned by "narrow" deviations. However, *there exists* a planner for whom all other deviations are undesirable. Intuitively, any non-narrow deviation has an appreciable effect on the level of income retention, instead of just its slope. While changes in the level of retention have several effects, one of them corresponds the curvature of the planner's re-distributional preferences: As she transfers more money to any one household, her desire to do so may increase or decrease. By taking the planner's preferences to be sufficiently risk-averse around each household's initial utility level, we can always make negative the second-order effects of tax variations that significantly vary the level of retention. Since (DEFG) ensures all variations that do *not* significantly vary the level of retention have negative second-order effects, this is sufficient to ensure local optimality.

While our focus on rationalizability facilitates this clean characterization, we also provide results relevant to the approach that fixes a welfare function and solves for optimal taxes: Lemmas 6 and 7 in Appendix C.1 provide the planner's full first- and second-order conditions.[27]

[26] More formally, we assume there exists a conditional expectation function $p_{\leq}(z;\epsilon) \equiv \mathbb{E}[\mathbb{1}_{\varepsilon^h(R)\leq\epsilon} | z_0^h = z]$ such that for all $z \in \operatorname{supp} g$, $p_{\leq}(z;\epsilon) > 0$. A conditional probability function exists because z_0^h is measurable by Assumption 3, because—since $\varepsilon^h(R)$ is measurable (see the second-to-last step in the proof of Lemma 2)—the indicator $\mathbb{1}_{\varepsilon^h(R)\leq\epsilon}$ is measurable, and because this indicator is also therefore integrable by dominated convergence.

[27] While we provide these conditions in full, they (a) require the estimation of many statistics in order to be evaluated and (b) do not lend themselves to simple verification of the second-order condition, even if all of the required sufficient statistics were known.

5 Empirical test of rationalizability

We now set out to empirically evaluate the (ABC) and (DEFG) tests. To do so, we estimate each of the sufficient statistics that comprise them in the NBER tax return datasets, which include both a panel of household tax returns and larger repeated cross sections. After evaluating the test, we check the robustness of our estimates using direct evidence that tax changes can indeed "sort" households by elasticities.

To set the stage, recall our second-order (DEFG) test for "Laffer valleys." In the simplest case where (a) the distribution of elasticities is locally constant in z, (b) taxes are locally linear in z, and (c) labor supply preferences are additive CES, the test becomes

$$(1 + R'(z))\varepsilon(z) \;\leq\; (1 - R'(z))\alpha(z)\left(\varepsilon(z)^2 + \mathbb{V}\mathrm{ar}[\varepsilon^h(R)|z_0^h = z]\right). \tag{14}$$

This expression serves two expository purposes. First, it highlights our main statistics of interest: The tax schedule ($R'(z)$), the shape of the income distribution ($\alpha(z)$), income-conditional mean compensated elasticities of taxable income ($\varepsilon(z)$), and the income-conditional variance of these elasticities. Of these statistics, income-conditional variance is the most novel to our analysis and poses the greatest estimation challenges.

Second, (14) facilitates simple, back-of-the-envelope calculations. In the US, top marginal taxes are roughly $R'(z) \approx 0.5$, the income distribution features a Pareto tail with shape $\alpha(z) \approx 2.5$, and existing elasticity estimates—while they vary widely—are in the vicinity of $\varepsilon(z) \approx 0.3$ [Gruber and Saez, 2002, Saez et al., 2012]. So, the second-order condition fails if and only if the variance of elasticities of top earners is ≥ 0.27. Indeed, this number is close to the lower bound on variance that is implied by the difference of elasticity between itemizers and non-itemizers estimated by Gruber and Saez [2002]: these groups are roughly evenly sized and their mean elasticities differ by about one, implying the variance of elasticities across only the two groups is about one quarter. We will estimate that within-income variances that are significantly higher, implying that taxes are in a Laffer valley.

5.1 Data

We use the NBER panel and repeated cross-section of tax returns from 1979 to 1990 [IRS-SOI, NBER, 1990]. This sample period includes major tax reforms such as the Economic Recovery Tax Act 1981, which decreased marginal rates in 3 years from 1982 to 1984; the Tax Reform Act 1986, which decreased the number of brackets and reduced the top marginal rate to 28%; the 1987 EITC expansion; and some state level tax reforms.

The data include limited demographic information, individual-specific federal and state income tax schedules,[28] and various measures of income. To construct a consis-

[28] These are computed using the NBER TAXSIM program, which calculates liabilities under US Federal

tent measure of taxable income for the whole period, we borrow from Gruber and Saez [2002]. The measure we use includes wage, business and capital income, and subtracts exemptions, standard and itemized deductions. The procedure of Gruber and Saez [2002] computes changes in income and marginal tax rates (at initial incomes) over 3-year windows. We drop individual-years with changes in marital status, initial (pre-deductions) income lower than $10 thousand dollars (in 1990 terms).

The panel sample consists of a random selection of four digit endings of social security numbers. There is purely random attrition; in some years a random subset of the sample social security numbers are excluded. The repeated cross-section sample is larger but not a panel; it over-represents higher-income individuals, which we account for with sample weights. Appendix Table A1 provides summary statistics.

This data set is appropriate for our exercise for three reasons. First, it contains tax changes that affect people throughout the income distribution, which allows us to estimate ETIs at each income level. Second, the data includes substantial tax variation at both the state and federal level, and as such workers' behavioral responses to these changes can be thought of as typical for real tax reforms. Finally, it contains tax changes of different sizes, which is part of the variation we leverage in our structural estimation.

5.2 Estimation of sufficient statistics

In order to evaluate our rationalizability test, we require estimates of the shape of income tax schedule, the shape of the income distribution, and several moments of households compensated, income, and super-ETIs conditional on income. Of these, the last is by far the most ambitious relative to the existing literature, which typically estimates only mean uncompensated elasticities within coarse income bins.

To make this task possible, we impose three structural assumptions:

1. Households have no income elasticity.

2. Households have CES disutility of labor supply.

3. Households respond fully to changes in taxes within three years.

These assumptions allow us to focus on estimating the income-conditional moments of medium-run compensated elasticities.[29] While these assumptions are strong, they are common in the literature and we feel that—despite them—our estimates provide a credible assessment of our rationalizability test.

and State income tax laws from individual data [Feenberg and Coutts, 1993].

[29] They imply that compensated elasticities are equal to uncompensated elasticities and that there households have no super-elasticities, except through changes in the curvature of the tax schedule.

5.2.1 Mean elasticities

We begin by estimating mean ETIs by income level. Our empirical strategy closely mirrors that of Gruber and Saez [2002], except that—instead of estimating a single ETI—we estimate ETIs locally in the space of year-demeaned log income $\tilde{z}$. Concretely, we estimate the following local regression:

$$\Delta \log z_t^h \;=\; a(\tilde{z}) \;+\; \varepsilon(\tilde{z}) \cdot \Delta \log R_t'(z_{t-1}^h) \;+\; c(\tilde{z}) \cdot mrs_{t-1}^h \;+\; d_{t-1} \;+\; \xi_t^h \tag{15}$$

where Δ represents time differences in t holding $t-1$ fixed, so that $\Delta \log R_t'(z_{t-1}^h) = \log R_t'(z_{t-1}^h) - \log R_{t-1}'(z_{t-1}^h)$ is the change in marginal retention at h's initial income. The year-demeaned-log-income-specific constant $a(\tilde{z})$ controls for differences in typical income changes in different brackets so that our estimates are not biased by, for example, mean reversion. We also control for a marital status dummy and time fixed effects.

Methodologically, we implement this regression by differencing out by year fixed-effects and then running local-polynomial regressions in year-demeaned-log-income space with a constant bandwidth and using the Epanechnikov kernel. We select this bandwidth using a leave-one-out cross validation procedure.[30] We compute confidence intervals by using the basic bootstrap method described in Chapter 5 of Davison and Hinkley [1997].

Under the assumption—which we maintain throughout—that, conditional on our controls, changes in marginal tax rates *within a (demeaned) income level* are as good as randomly assigned, this regression identifies the average ETI of households with each year-demeaned log-income $\tilde{z}$. This is exactly the statistic required by our theory.[31] Given the presence of year fixed effects, our identifying variation consists of (a) within-year, within-income variation in tax changes across individuals (especially those living in different states) and (b) across-year variation in the relative tax rates between different income levels.[32] Year fixed effects control for the fact that the government may adjust taxes in anticipation of changes in aggregate economic conditions, so long as these tax changes are uniform across tax brackets.

Figure 4 shows our estimates of ETI by income level.[33] Our estimates are consistent

[30] For more details, see Appendix D.1.

[31] Our estimation strategy differs from some earlier work, such as that of Gruber and Saez [2002], which identifies a hypothetical elasticity concept based on locally linear taxes. Our elasticity concept is inclusive of changes in income caused by "knock-on" changes in marginal taxes as a household adjusts its income.

[32] Decomposing these sources of variation helps to clarify in what cases our identifying assumptions are violated. For example, variation through (a) violates the identifying assumptions if incomes decline in a state in the same year that state has a tax change, but for unrelated—and yet not statistically independent—reasons. Variation through (b) violates the identifying assumptions if incomes at some income level decline differentially (relative to those at other income levels) in the same year that taxes change differentially across those brackets, but for unrelated—and yet not statistically independent—reasons.

[33] As discussed above, we estimate ETI locally in year-demeaned-log-income space. Figure 4 converts these estimates to income space in 1990 by combining log income levels that are the same distance

with—though somewhat on the lower end of—estimates in the literature. In line with Gruber and Saez [2002] and Kumar and Liang [2020], we find somewhat higher elasticities at the bottom and the top of the income distribution.

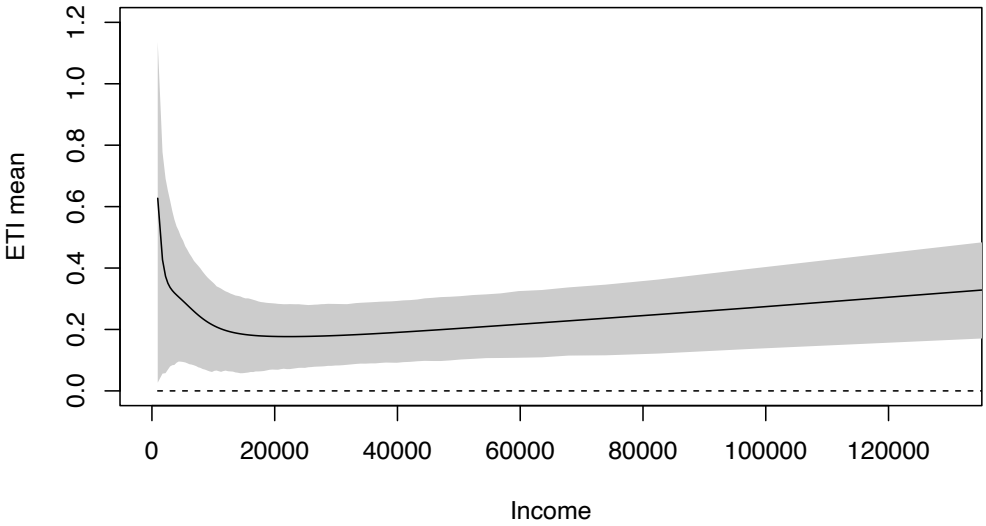

Figure 4: Mean ETI by income level in 1990 USD, and 95% confidence bands.

5.2.2 Variance of elasticities

Our second object of interest is the variance of ETI conditional on income level. While there is some precedent for studying differences in ETIs across groups, our approach varies in two main ways. First, the existing literature does not emphasize within-income variation and in some cases explicitly focuses on across-income variation [Eissa and Liebman, 1996, Gruber and Saez, 2002, Kopczuk, 2005, Kumar and Liang, 2020]. By contrast, our test explicitly calls for within-income variance estimates. Second, the literature has focused on differences in elasticity conditional on observed characteristics. However—insofar as our test does not apply when taxes can condition on observed heterogeneity—we are particularly interested in unobserved heterogeneity (see Section 4.3.1). This focus brings with it additional identification challenges, which we discuss.

We pursue two main strategies: (a) a conditioning on observables approach based on differences across itemization status and (b) a structural estimation approach that relies on the linearity of tax responses.

Conditioning on observables: itemization status

Motivated by work that has documented higher elasticities for households who itemize deductions, we use total deductions to split our sample into heavy and light itemizers

from average log incomes in their respective years. Appendix Figure A3 shows the original estimates in year-demeaned-log-income space.

[Gruber and Saez, 2002, Kopczuk, 2005]. While we could simply compare households who do and do not itemize (at all), this would be problematic in our setting because almost all high-income households do itemize. So that we have variation in itemization status at each income level, we categorize households as light or heavy itemizers depending on whether they claim less or more than the mean value of deductions *at their income level*.[34] We define that status in year before the tax changes we consider, so that it does not implicitly control for changes in income.

After classifying individuals into these two groups, we estimate the local following regression at each year-demeaned income level $\tilde{z}$:

$$\begin{aligned}\Delta \log z_t^h \;=\; a(\tilde{z}) \;+\; \varepsilon_{L,t}(\tilde{z}) \cdot \Delta \log R_t'(z_{t-1}^h) \;+\; \delta(\tilde{z}) \cdot HI_{t-1}^h \cdot \Delta \log R_t'(z_{t-1}^h) \\ +\; c(\tilde{z}) \cdot mrs_{t-1}^h \;+\; d_{t-1} \;+\; \xi_t^h\end{aligned} \tag{16}$$

where HI_{t-1}^h is an indicator for heavy itemizers. This differs from our estimation of unconditional mean elasticities in two ways. First, we interact tax changes with a high income dummy in order to estimate the difference in elasticity by itemization status. Second, by allowing $\varepsilon_{L,t}(\tilde{z})$ to differ across years, we ensure that $\delta(\tilde{z})$ measures only *within-year* differences in elasticities by itemization status, despite the fact that during our sample period, a broadening of the tax base during our sample period may have both reduced elasticities and reduced itemizations [Kopczuk, 2005].[35] We estimate this regression locally in year-demeaned log-income space with the same methodology described in Section 5.2.1. To compute confidence intervals, we bootstrap both steps of this procedure, holding fixed the choices for the bandwidths, and use the basic bootstrap confidence interval.

The left panel of Figure 5 shows our estimates of differences in elasticity between heavy and light itemizers.[36] Next, we combine these estimates with estimates of the shares of heavy-itemizers by income bracket, in order to compute an implied lower bound on the variance of elasticities by income brackets.[37] Again, we compute confidence intervals by bootstrapping the whole procedure while holding bandwidth choices fixed. The right panel of Figure 5 shows the implied variance as a function of income. At high incomes, the lower bound on variance we estimate is already close to the level required to violate

[34] Concretely, we compute the mean total deductions by income bracket (*nitem*) non-parametrically by estimating the following local regression in the space of year-demeaned log income

$$nitem_{t-1}^h = a(\tilde{z}) + b(\tilde{z}) \log z_t^h + \xi_{t-1}^h.$$

While $a(\tilde{z})$ is our actual estimate, we also include the linear term as recommended by Fan and Gijbels [1996].

[35] Had we not allowed $\varepsilon_{L,t}(\tilde{z})$ to depend on year, we may have picked up *between-year* differences related to that there may be relatively more itemizers in years where both heavy and light itemizers have high elasticities—in which case the difference between average elasticities of heavy and light itemizers across years could exceed the average difference within years—or vice-versa.

[36] Appendix Figure A4 show the analogous plots of Figure 5 in the space of year-demeaned log income, where we perform the actual estimation before mapping to 1990.

[37] We estimate these shares by the same local-polynomial approach used for our other regressions.

the back-of-the-envelope second-order condition discussed above.

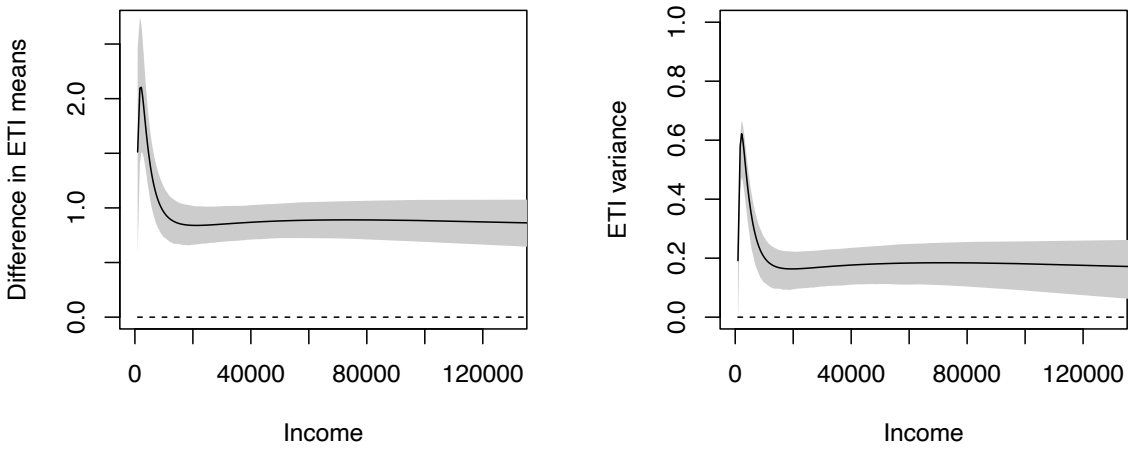

Figure 5: Left panel: Difference in elasticities between heavy and light itemizers, by income level in 1990. Right panel: Implied lower bound on variance in elasticity, by income level in 1990. In gray: 95% confidence bands.

This approach generates a conservative but robust estimate of the variance in ETI by income level. It is conservative because it leverages only differences in elasticities across two observable groups, ignoring observed and unobserved variation within these groups. At the same time, it is robust to the concern that (due to super-elasticities or tax salience) households' tax responses may be non-linear in the size of tax changes. Such non-linearities could cause us to mistake variances in the size of tax changes for variance in ETI at a given tax change; this is not an issue when estimating means, so long as each group faces similarly-sized tax changes.

Structural estimation

In order to estimate unobserved heterogeneity in elasticities, we complement the approach above with a second, more structural approach.

The key idea underlying this approach that if (a) each household's elasticities do not depend on the size of the tax changes they face and (b) one has access to continuous variation in the size of tax changes, then all moments of the distribution of elasticities—and in particular its variance—are non-parametrically identified. Intuitively, linearity implies that as the treatment size increases, the variance of treatment responses can be identified out of the change in the dispersion of the outcomes. This is because the heterogeneity in treatment effects should progressively magnify the dispersion of outcomes as the size of the treatment increases.[38]

While linearity is, admittedly, a strong assumption, it allows us to remain very flexible in modelling the structure of ETI heterogeneity. In particular, we assume that the

[38]For a proof, see Appendix D.2.

distribution of elasticities has finite support, but allow this support to have many points, each taking arbitrary values. This flexibility allows us to approximate many different distributions, while discreteness ensures that our resulting estimates of ETI variance cannot be driven by a fat tail.

More explicitly, we assume that

$$\widehat{\Delta \log z_t^h} \quad = \quad a_{g_t^h} \quad + \quad \varepsilon_{g_t^h} \cdot \Delta \widehat{\log R_t'(z_{t-1}^h)} \tag{17}$$

where g_t^h takes a finite number of values, and where $\widehat{\Delta \log z_t^h}$ and $\Delta \widehat{\log R_t'(z_{t-1}^h)}$ are changes in log income and change in log taxes at initial income, respectively, after partialling out by marital status, a 10-piece linear spline in log taxable income, and year dummies. We estimate this regression by running a k-means algorithm that minimizes the mean square error of (17) by assigning each (h, t) pair to a group, and iteratively estimating $\{a_g, \varepsilon_g\}$ within groups.[39]

The left panel of (6) displays our estimated distribution of ETIs. The estimated distribution has a very large mass around zero and some mass at higher values, up to a maximum ETI of 16. While this number is large, we believe it is plausible as an intensive margin elasticity at low incomes or as a proxy for extensive margin labor supply or tax avoidance decisions at any income. For robustness, we also present versions of our main results in which ETIs are capped at 5.

Given assignments $\hat{\varepsilon}_t^h$ of ETIs to each individual at each date t (i.e. between t and $t+1$), we compute the variance in ETIs by income level using local polynomial regressions.[40] In order to ensure that we capture only *within-year* variation in elasticities, we estimate the following regression locally in the space of year-demeaned log incomes $\tilde{z}$:

$$\left(\hat{\varepsilon}_t^h - \hat{m}_t(\tilde{z})\right)^2 = a(\tilde{z}) + b(\tilde{z}) \log z_t^h + \xi_{t-1}^h \tag{18}$$

where $\tilde{z}_t^h$ is year-demeaned log income, and $\hat{m}_t(\tilde{z}_t^h)$ is the year and income specific mean of elasticities, also estimated using first order local polynomials.

One important concern about this procedure is that it may be prone to small sample bias, as we estimate a large number of parameters. To address potential small sample bias and to obtain confidence intervals for our estimates of variance, we bootstrap the entire procedure, holding fixed the number of groups k and all bandwidths, and then subtract from our point estimates the bootstrap estimator of the bias.

The right panel of Figure 6 shows our structural estimates of income-conditional variance in ETI.[41] Notably, we estimate significantly higher variances than with the lower

[39] We do this while fixing the number of groups k, which we select with a Bayesian information criterion subject to a maximum possible number of groups $k_{max} = 100$ as in Bonhomme and Manresa [2015].

[40] We follow an analogous procedure to compute the third moment of ETIs.

[41] Here again we have combined incomes across years that share a common year-demeaned log income.

bound approach of the prior section. Appendix Figure A5 shows our structural estimates when ETIs are capped at 5.

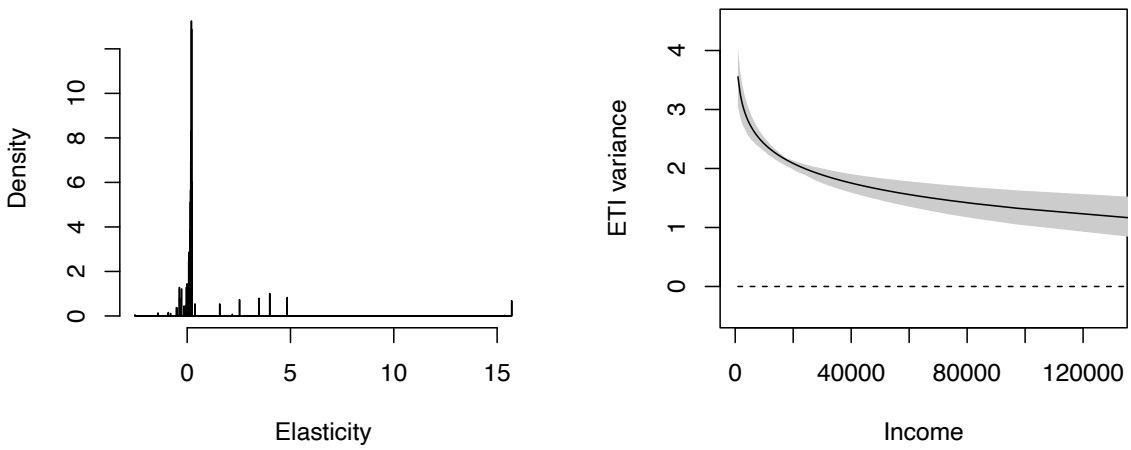

Figure 6: Left panel: Histogram of ETIs identified by k-means procedure. Right panel: Implied variance in elasticities by income level. In gray: 95% confidence bands.

5.2.3 Estimation of tax schedule, income density

Our rationalizability test also relies on estimates of the shape of the tax schedule and the income distribution. We estimate these moments in the NBER cross sectional files, which contain more observations and so allow for more precise estimation.

To compute $\alpha(z)$ in each year, we estimate a smooth functional form for the CDF of taxable income $G_t(z)$ in each year t with local polynomials regressions to third degree in log taxable income. Here we follow Fan and Gijbels [1996], who recommend including terms up to one order above the derivative of interest. We then translate our estimated coefficients for the first and second derivatives of $G_t(z)$ in log income into an estimate of $\alpha_t(z)$.[42] We obtain confidence intervals by bootstrapping the whole procedure while holding the estimated optimal bandwidth constant. Our estimates are broadly consistent across years and similar to other recent work such as those of Hendren [2020]. The left panel of Figure 7 shows our estimates in 1990.

We take a similar approach to estimate marginal tax rates and each of its derivatives required by our test. In these cases we estimate local-polynomial regressions of order d, where d is the nearest odd integer greater than the derivative of interest (in log income space). Again, we boostrap confidence intervals holding optimal bandwidths fixed. The right panel of Figure 7 shows our estimates in 1990.[43]

[42]Namely our estimate of $\alpha_t(z)$ is equal to the negative of our local estimate of $\frac{d^2 G_t(z)}{d \log z^2}$ divided by our local estimate of $\frac{dG_t(z)}{d \log z}$.

[43]For other years, see Appendix Figures A6 to A16.

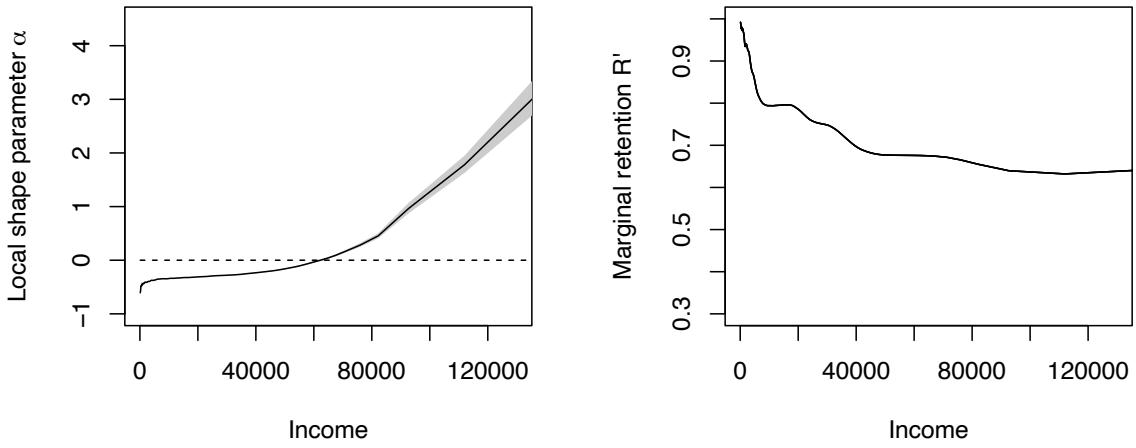

Figure 7: Left panel: Local shape parameter of the 1990 income distribution, by income level. Right panel: Average marginal taxes in 1990, by income level. 95% confidence bands.

By construction, our estimates of the tax schedule do not feature any kinks. This reflects that we run local regressions with a finite bandwidth, which smooths out the schedule. To the extent that kinks affect our rationalizability test, we therefore ignore their effects. However, we think that a smoothed tax schedule may serve as a reasonable proxy for households' imperfect perceptions of the schedule or their inability to perfectly adjust their incomes, so we do not dwell on this issue [Rees-Jones and Taubinsky, 2020].[44]

5.3 Evaluation of test

Having computed each of the elements in the (ABC) and (DEFG) formulas, we can evaluate whether the tax schedule is rationalizable.

5.3.1 First-order test

We start with the first-order test (ABC), which recall assesses whether taxes are above the top of the Laffer curve at each income level z. Concretely, we compute—for each year t in our data—the statistic

$$\widehat{ABC}_t(z) \;=\; -1 \;+\; \frac{1-R_t'(z)}{R_t'(z)}\left(\alpha_t(z) - \frac{d\log}{d\log z}\left(\frac{1-R_t'(z)}{R_t'(z)}\right) - \frac{d\log\varepsilon(z)}{d\log z}\right)\varepsilon(z) \qquad (19)$$

where recall R_t and α_t are estimated separately in each year. We compute confidence intervals for $\widehat{ABC}_t(z)$ by combining bootstrap replications from the two different data sets we use—the panel and the cross-sectional file for year t—assuming that observations

[44]To ensure our test is robust to this treatment of kinks, we also estimate our second-order test with an alternative methodology that excludes the term proportional to tax curvature. This has negligible effects, see Appendix Figure A29.

in separate datasets are drawn independently from each other.

Our estimates of $\widehat{ABC}_t(z)$ are similar across years and consistently negative, implying that taxes are below the top of the Laffer curve. The left panel of Figure 8 shows this for one representative year; Appendix Figures A17 to A27 show our estimates for all years.

5.3.2 Second-order test

We now evaluate our new, second-order test (DEFG), which recall assesses whether taxes are in a Laffer valley at each income level z. Concretely, we compute—for each year t in our data—the statistic[45]

$$\widehat{DEFG}_t(z) = -\left(1 + R_t'(z)\right)\varepsilon(z) + \left(1 - R_t'(z)\right)\left[\alpha_t(z)\varepsilon^2(z) - \frac{d\varepsilon^2(z)}{d\log z} + \frac{d^2 \log R_t'(z)}{d\log z^2}\varepsilon^3(z)\right] \quad (20)$$

where the elasticities are derived from our structural estimates. We compute confidence intervals in the same way described for $\widehat{ABC}_t(z)$.

Our estimates of $\widehat{DEFG}_t(z)$ are broadly consistent across years. Strikingly, they fail at incomes above around 90,000 in 1990 USD.[46] The right panel of Figure 8 shows this for one representative year; Appendix Figures A17 to A27 show our estimates for all years. In Appendix Figures A28 to A30, we compare $\widehat{DEFG}_t(z)$ with and without including the last two terms of (20); this has almost no effect, implying that our estimates are not driven by the third moment of elasticities or steep changes in the second moment.

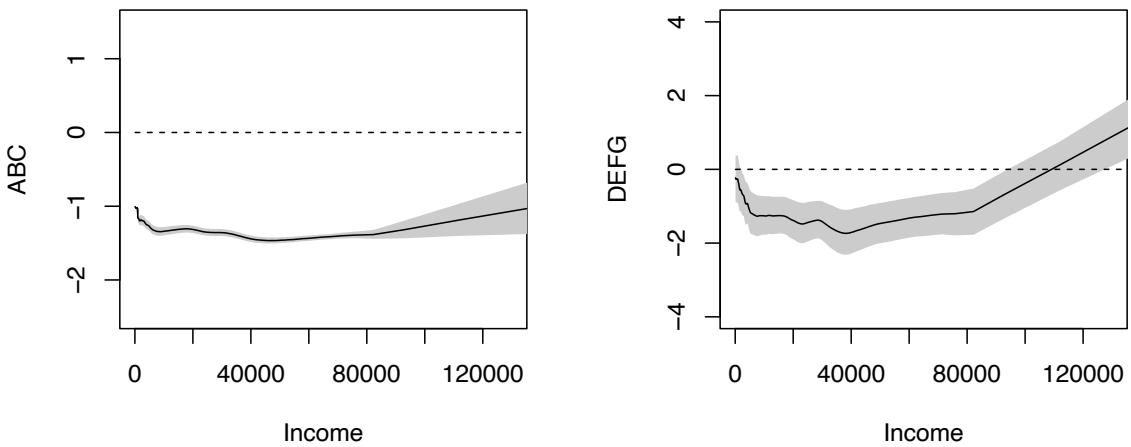

Figure 8: Left panel: ABC test evaluated in 1990. The test passes at all incomes. Right panel: DEFG test evaluated in 1990. The test fails at high incomes. 90% confidence bands. We use 90% confidence bands because the hypothesis we are testing is one-sided.

An alternative way to visualize (DEFG)—and to compare the implications of our dif-

[45] This is equal to the RHS of (DEFG) when, as we have assumed, income effects are zero and labor supply preferences are additive-CES; see Appendix A.4.1 for details regarding the super-elasticity term.

[46] The test continues to fail if we restrict our structural estimates so that the maximum ETI assigned to any k-means group is 5; see Appendix Figure A5.

ferent estimates of variance—is to reframe the second-order test in terms of the minimum level of variance at which it fails. We can then compare this level with our estimates of variance at each income level. Figure 9 show this comparison.[47] Blue and orange regions represent the levels of ETI variance at which the (DEFG) test passes and fails, respectively, at each income. On top of this background, we superimpose our estimates of variance based on the two strategies in Section (5.2.2). The test easily fails under our structural estimates of variance, while our lower bounds on variance based on itemization status come about half of the way to violating the test. Not shown below (see Appendix Figure A5) are our structural estimates of variance when elasticities are capped at 5; these violate the test on point estimates but cannot reject that the test passes.

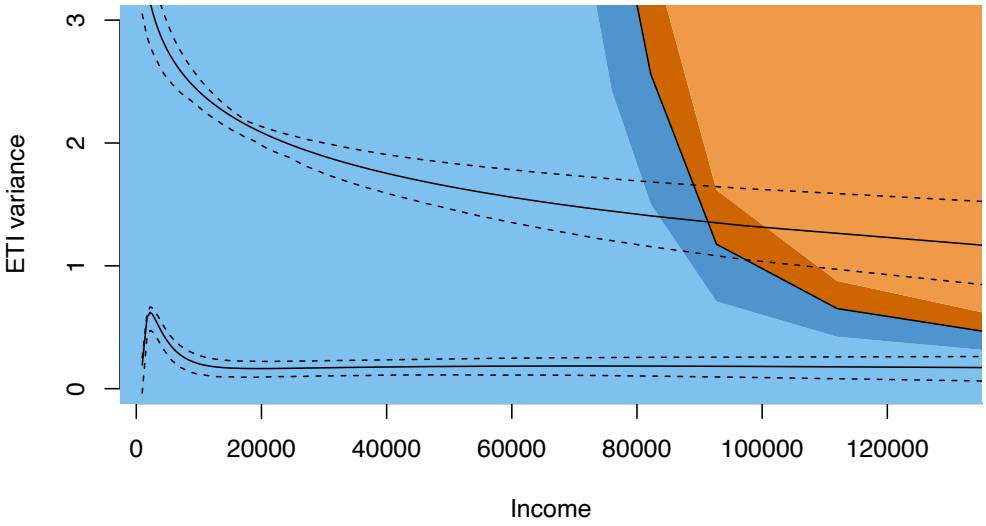

Figure 9: 1990. Background: ETI variance consistent (blue) and inconsistent (orange) with (DEFG). Foreground: Estimates of ETI variance. Darker colors and dashed lines colors are 95% confidence bands.

5.4 Inspecting the mechanism

A number of potential confounding factors could prevent the heterogeneity in elasticities we have estimated from translating into the type of elasticity-sorting that our theory predicts. For example, perhaps our estimates mistake differences across households in the *timing* of tax responses for differences in the *size* of tax responses.[48] Another possibility is that super-elasticities—which we do not estimate—could work to attenuate the level of

[47] It does so under the assumptions that, at each z, (a) the tax schedule is locally linear and (b) the variance of ETIs is locally constant. This allows the test to be framed as simple "variance test" without considering the other roles played by second and third moments of ETIs. Figure A29 shows that this simplification is without significant loss.

[48] Our conditioning-on-observables estimates are robust to this concern, but our structural estimates are not.

elasticity sorting, similar to the discussion in Section 4.3.2. Alternatively, heterogeneity in elasticities across income levels may reflect differences in the institutional arrangements of employment in different jobs, which may not respond to tax changes.

To address these concerns, we now attempt to directly validate the mechanism that underlies our theory, i.e. the idea that changes in taxes cause changes in the average elasticity within each income level. The main idea behind our strategy is as follows: In regions of the tax schedule where the income density is strongly decreasing ($\alpha(z) > 0$), a locally flat increase in marginal retention should increase the average elasticity conditional on income. This is because—when there is significant heterogeneity in elasticities and when all elasticity groups have proportional densities—the high elasticity types that end up at each income level z come from much lower incomes, where the density is higher. This pattern flips where $\alpha(z) < 0$, so that an increase in marginal retention should lower the average elasticity at z.

More concretely, one can show that that for small tax changes,

$$\Delta\varepsilon(z) \quad \approx \quad \alpha(z) \;\; \mathbb{V}\mathrm{ar}[\varepsilon^h(R)|z_0^h = z] \;\; \Delta R'(z). \tag{21}$$

In other words, there is a positive interaction between the ETI $\varepsilon(z)$ at a given income level and the product of the Pareto shape $\alpha(z)$ and the level of marginal retention $R'(z)$, and the size of this interaction term is equal to the local variance of ETIs. In principle, we could therefore identify ETI variance at each income level by including in the mean ETI regression (15) an interaction between tax changes and the product of $\alpha(z) \cdot R'(z)$. In practice, this approach has two limitations, and these limitations motivate our actual strategy. First, it may be under-powered, so we pool across income levels. Second, one source of variation in $R'(z)$—tax differences across states (rather than years)—may be correlated with elasticity differences for reasons other than our mechanism and therefore bias this regression. Namely, one should expect that states with more income-elastic or cross-state-migration-elastic populations (at any income z) respond by levying lower marginal taxes. We therefore focus on the component of variation in taxes that comes solely from time variation.

Our preferred specification is as follows:

$$\begin{aligned}\Delta \log\left(z_t^h\right) &= \varepsilon \cdot \Delta \log R_t'(z_{t-1}^h) \\ &\quad + \gamma \cdot \Delta \log R_t'(z_{t-1}^h) \cdot (\overline{\log R'}_{t-1}(z_{t-1}^h) - \overline{\overline{\log R'}}(z_{t-1}^h)) \cdot \alpha_{t-1}(z_{t-1}^h) \\ &\quad + c \cdot mrs_{t-1}^h + d_{t-1} + f(z_{t-1}^h) + \xi_t^h\end{aligned} \tag{22}$$

where $\overline{\log R'}_{t-1}(z_{t-1}^h)$ is the average log marginal retention rate at z_{t-1}^h in year $t-1$, $\overline{\overline{\log R'}}(z_t^h)$ is the average log marginal retention rate at z_t^h across years, and where $f(z_{t-1}^h)$ is a control for the level of income (either a single linear term or ten piece splines). As

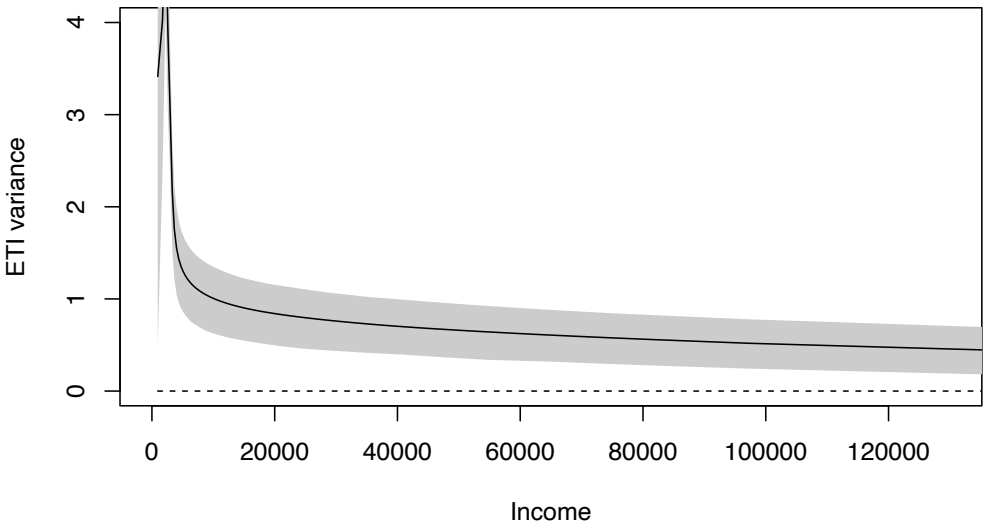

Figure 10: Income-conditional variance implied by elasticity differences across low- and high-tax years in 1990, 95% confidence bands.

discussed above, one may interpret the regression coefficient on this interaction term as an average of within-income ETI variances, taken across incomes.

Appendix Table A2 presents our results. We estimate $\hat{\gamma} = 1.05$ (standard error 0.19) when controlling linearly for incomes and $\hat{\gamma} = 0.11$ (standard error 0.12) when controlling for ten piece income splines, providing reduced form support for our hypothesis of tax-induced sorting.

Finally, we extract an implicit measure of variance by income bracket by running the regression above non-parametrically, allowing $\gamma(z), \epsilon(z)$ and $c(z)$ to depend on demeaned income. Figure 10 shows that we estimate variances similar to those in Section 5.2.2.

6 Quantitative application

So far, we have presented evidence that the taxes in our sample can—from the perspective of any planner in a broad class—be improved. But how large are these potential welfare gains, and what sort of tax changes achieve them? In this section, we speak to these questions using a simple calibrated model.

Model: Motivated by the fact that our efficiency test fails only at high incomes, we remain agnostic to the behavior of households with incomes below \$100,000 and to the planner's preferences regarding their welfare. For households with incomes above \$100,000, we impose both positive and normative assumptions. Positively, we assume the following: First, consistent with the year 1990, top incomes are distributed as Pareto($\alpha_{\text{top}} = 2.5$),

are subject to an initial top marginal tax rate of $\tau_{\text{top}} = 35\%$. Second, each top-earning household h has a constant compensated elasticity of labor supply β^h and no income elasticity. Third, β^h is distributed independently of productivity according to a truncated log-normal distribution with mean 0.3, variance 1.2, and maximum 10.[49]

Normatively, we must take a stance on how the planner values changes in household welfare relative to government revenue in order to quantify the gains from tax reform. To do so, we assume the following: First, the planner places a common and constant value λ_{top} on transfers to each top-earner, relative to government revenue.[50] Second, in setting the initial tax schedule, the planner has "followed the literature" by checking the first- but not the second-order condition of the tax schedule. Together with our positive assumptions, this allows us to back out the welfare weight:

$$\lambda_{\text{top}} \quad = \quad 1 \; - \; \frac{1 - R'_{\text{top}}}{R'_{\text{top}}} \; \alpha_{\text{top}} \; \varepsilon_{\text{top}} \quad \approx \quad 0.6. \tag{23}$$

Simple tax reforms: To motivate the particular tax changes we consider, we return to the planner's second-order condition. In our specialized model, this takes a particularly uncomplicated form for any tax change Δ that only affects top earners:[51]

$$\frac{d^2}{d\epsilon^2} \text{Welfare}(R + \epsilon\Delta) \;=\; \int \underbrace{\left[-(1 + R_{\text{top}})\varepsilon_{\text{top}} + (1 - R_{\text{top}})\alpha_{\text{top}}\varepsilon^2_{\text{top}}\right]}_{\text{LHS of (DEFG)}} \left(\frac{\Delta'(z)}{R'_{\text{top}}}\right)^2 zg(z)dz. \tag{24}$$

One may easily verify that the "LHS of (DEFG)" term is positive under our assumptions, implying that the planner's second-order condition fails, i.e. taxes are in the Laffer valley.

This expression says that a change in taxes on top earners improves welfare to the extent it changes the marginal taxes they face. In particular, there is no need to restrict to very "narrow" tax changes—as was done in the main text—in order to exploit the failure of the second-order condition; narrow tax changes are just a technical device used to handle income effects and non-constant welfare weights, both of which are absent here. Even simple tax changes, such as raising or lowering marginal taxes on all top earners, can "sort and extort," as illustrated in the motivating example of Section 2.

This in mind, we consider two very simple tax reforms: A 20 percentage point increase and a 20 percentage point decrease in the top marginal tax rate.

Welfare gains and discussion:

Consistent with the observation that taxes are in a Laffer valley, we document welfare

[49] We view this as a sensible approximation of our empirical estimates of elasticity heterogeneity.

[50] Here, we implicitly assume that the planner does not face a hard budget constraint, but instead has a constant marginal value of public funds.

[51] This follows from Lemmas 6 and 7; see Appendix C.1.

gains from either raising or lowering taxes. Concretely, the planner's gains from either raising (lowering) top taxes by 20pp are equivalent to the social value of transferring about \$3700 (\$600) to each household in the top bracket or $\lambda_{\text{top}} \cdot$ \$3700 $\approx$ \$2200 (\$400) to the planner for each household in the top bracket. Appendix Figure A31 shows welfare gains as a function of the top tax rate.[52]

The estimates above are a lower bound on welfare gains for several reasons. First, we take all taxes below \$100,000 as given, thus ignoring the potential gains from simultaneously changing them. Second, we restrict taxes above \$100,000 to be linear, ignoring the additional gains from using non-linear tax changes. Third, we limit ourselves to 20 percentage point changes in the top tax rate, ignoring the gains from using larger tax changes. Fourth, so that we need only make assumptions about behavior in the top bracket, we ignore any behavioral effects on households with lower initial incomes. This results in an underestimate since the only case in which such households respond to top rate changes is when they jump into the top bracket, which has positive fiscal externalities.

One topic not addressed above is the shape of the *optimal* tax reform, and in particular whether it is similar to the simple reforms we consider. For instance, one may worry that a schedule with many "squiggles" would be more effective in sorting households with different elasticities and therefore optimal. While we do not answer this question in general, Appendix B.3 presents a suggestive exercise in which, starting from a sub-optimal tax schedule, we update taxes in the direction suggested by the planner's first-order condition until converging to a new local optimum. In a simple case with two elasticity types, we find that taxes converge to a "two-part" schedule that smoothly transitions between a high tax rate used on low incomes and a low tax rate used on high incomes—rather than a "squiggly" schedule.

7 Conclusion

We reconsider the classical non-linear income taxation problem in an environment with arbitrary household heterogeneity. Our theoretical results show—without taking a stance on redistributive preferences—that sufficient ETI heterogeneity presents an opportunity for welfare-improving tax reforms that sort households according to their elasticities and then tax them differentially. Empirically, we use a panel of US tax returns to document large variances in ETI conditional on income, indicating scope for such "sort and extort" tax reforms. A conservative quantification exercise suggests that raising (lowering) top taxes by 20 percentage points results in yearly welfare gains equivalent to approximately \$3700 (\$600) per top earner.

[52] It also illustrates that whether increasing or decreasing taxes is better depends on non-local information on how the economy would behave away from the current tax system.

References

C Aliprantis and K Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag Berlin Heidelberg, 2006.

Lennard Bakker. Uniform convergence and differentiation. URL `https://math.byu.edu/~bakker/M341/Lectures/Lec29.pdf`.

Spencer Bastani and Daniel Waldenström. The ability gradient in tax responsiveness. *Journal of Public Economics Plus*, 2:100007, 2021. ISSN 2666-5514. doi: https://doi.org/10.1016/j.pubecp.2021.100007. URL `https://www.sciencedirect.com/science/article/pii/S2666551421000036`.

Felix Bierbrauer, Pierre Boyer, and Emanuel Hansen. Pareto-improving tax reforms and the earned income tax credit. 2020.

Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

Francine D Blau and Lawrence M Kahn. Changes in the labor supply behavior of married women: 1980–2000. *Journal of Labor economics*, 25(3):393–438, 2007.

Stéphane Bonhomme and Elena Manresa. Grouped patterns of heterogeneity in panel data. *Econometrica*, 83(3):1147–1184, 2015.

Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. Discretizing unobserved heterogeneity. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2019-16), 2017.

Anthony Christopher Davison and David Victor Hinkley. *Bootstrap methods and their application*. Number 1. Cambridge university press, 1997.

Peter A Diamond. Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95, 1998.

Nada Eissa and Jeffrey B Liebman. Labor supply response to the earned income tax credit. *The quarterly journal of economics*, 111(2):605–637, 1996.

Jianqing Fan and Irene Gijbels. *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, volume 66. CRC Press, 1996.

Daniel Feenberg and Elisabeth Coutts. An introduction to the taxsim model. *Journal of Policy Analysis and management*, 12(1):189–194, 1993.

Gerald B Folland. *Real analysis: modern techniques and their applications*, volume 40. John Wiley & Sons, 1999.

Jon Gruber and Emmanuel Saez. The elasticity of taxable income: evidence and implications. *Journal of public Economics*, 84(1):1–32, 2002.

Roger Guesnerie. *A contribution to the pure theory of taxation*. Number 25. Cambridge University Press, 1998.

Peter J Hammond. Straightforward individual incentive compatibility in large economies. *The Review of Economic Studies*, 46(2):263–282, 1979.

Nathaniel Hendren. Measuring economic efficiency using inverse-optimum weights. *Journal of Public Economics*, 187:104198, 2020.

CJ Himmelberg. Measurable relations. *Fundamenta Mathematicae*, (87):53–72, 1975.

James R Hines Jr and Michael J Keen. Certain effects of random taxes. *Journal of Public Economics*, 203:104412, 2021.

Roozbeh Hosseini and Ali Shourideh. Retirement financing: An optimal reform approach. *Econometrica*, 87(4):1205–1265, 2019.

IRS-SOI, NBER. Continuous work history sample, 1979-1990, 1990. The data is hosted at the NBER and can be accessed by contacting Dan Feenberg. Accessed: 03-08-20.

Laurence Jacquet and Etienne Lehmann. Optimal income taxation when skills and behavioral elasticites are heterogeneous. 2015.

Laurence Jacquet and Etienne Lehmann. Optimal income taxation with composition effects. *Journal of the European Economic Association*, 2020.

Ohad Kammar. A note on frechet diffrentiation under lebesgue integrals. 2016.

Wojciech Kopczuk. Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics*, 89(11-12):2093–2119, 2005.

Anil Kumar and Che-Yuan Liang. Estimating taxable income responses with elasticity heterogeneity. *Journal of Public Economics*, 188:104209, 2020.

Daniel J Lewis, Davide Melcangi, and Laura Pilossoph. Latent heterogeneity in the marginal propensity to consume. *FRB of New York Staff Report*, (902), 2019.

Qi Li and Jeffrey Scott Racine. *Nonparametric econometrics: theory and practice*. Princeton University Press, 2007.

Benjamin Lockwood, Afras Y Sial, and Matthew C Weinzierl. Designing, not checking, for policy robustness: An example with optimal taxation. Technical report, National Bureau of Economic Research, 2020.

Gunnar Martinsson. The implicit and inverse function theorems. URL `https://users.oden.utexas.edu/~pgm/Teaching/APPM5450_2006s/implicit_function_thm.pdf`.

Eric Maskin and John Riley. Optimal auctions with risk averse buyers. *Econometrica: Journal of the Econometric Society*, pages 1473–1518, 1984a.

Eric Maskin and John Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196, 1984b. ISSN 07416261. URL `http://www.jstor.org/stable/2555674`.

Helmut Maurer and Jochem Zowe. First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Mathematical programming*, 16(1):98–110, 1979.

Christian Moser and Pedro Olea de Souza e Silva. Optimal paternalistic savings policies. *Columbia Business School Research Paper*, (17-51), 2019.

Jeff Racine. An efficient cross-validation algorithm for window width selection for non-parametric kernel regression. *Communications in Statistics-Simulation and Computation*, 22(4):1107–1114, 1993.

Alex Rees-Jones and Dmitry Taubinsky. Measuring "schmeduling". *The Review of Economic Studies*, 87(5):2399–2438, 2020.

Jean-Charles Rochet. The taxation principle and multi-time hamilton-jacobi equations. *Journal of Mathematical Economics*, 14(2):113–128, 1985.

Jean-Charles Rochet and Lars A Stole. The economics of multidimensional screening. *Econometric Society Monographs*, 35:150–197, 2003.

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1):205–229, 2001.

Emmanuel Saez, Joel Slemrod, and Seth H Giertz. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of economic literature*, 50(1): 3–50, 2012.

Florian Scheuer and Iván Werning. Mirrlees meets diamond-mirrlees: Simplifying non-linear income taxation. 2018.

Galen R Shorack. *Probability for statisticians*. Number 04; QA273, S4. Springer, 2000.

Joseph Stiglitz. Utilitarianism and horizontal equity: The case for random taxation. *Journal of Public Economics*, 18(1):1–33, 1982a.

Joseph E Stiglitz. Self-selection and pareto efficient taxation. *Journal of public economics*, 17(2):213–240, 1982b.

Dmitry Taubinsky and Alex Rees-Jones. Attention variation and welfare: theory and evidence from a tax salience experiment. *The Review of Economic Studies*, 85(4): 2462–2496, 2018.

James P Vere. Social security and elderly labor supply: Evidence from the health and retirement study. *Labour Economics*, 18(5):676–686, 2011.

Iván Werning. Pareto efficient income taxation. Technical report, mimeo, MIT, 2007.