

NBER WORKING PAPER SERIES

SUFFICIENT STATISTICS FOR NONLINEAR TAX SYSTEMS WITH PREFERENCE
HETEROGENEITY

Antoine Ferey
Benjamin Lockwood
Dmitry Taubinsky

Working Paper 29582
<http://www.nber.org/papers/w29582>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2021

We are grateful to Afras Sial for excellent research assistance. This project was supported through a Quartet Pilot Research award and was funded by the Center for Health Incentives and Behavioral Economics at the University of Pennsylvania, by The Research Council of Norway 315765, and by the Alfred P. Sloan Foundation. Ferey gratefully acknowledges the financial support of Labex ECODEC at CREST, and of the Deutsche Forschungsgemeinschaft through CRC TRR 190 at LMU Munich. The content is solely the responsibility of the authors and does not necessarily represent the official views of the University of Pennsylvania. We thank Pierre Boyer, Philippe Choné, Ashley Craig, Nathan Hendren, Laurence Jacquet, Dirk Krueger, Etienne Lehmann, Jonas Loebbing, Jean-Baptiste Michau, Andreas Peichl, Dominik Sachs, Emmanuel Saez, Florian Scheuer, Stefanie Stantcheva, Joseph Stiglitz, Aleh Tsyvinski, Nicolas Werquin, Eric Zwick, participants at NBER Macro Public Finance, NBER Public Economics, IIPF, LAGV, NTA, and audiences at UC Berkeley, CREST - Ecole Polytechnique and LMU Munich for their valuable comments. Ferey: LMU Munich & CESIFO (antoine.ferey@econ.lmu.de). Lockwood: Wharton & NBER (ben.lockwood@wharton.upenn.edu). Taubinsky: UC Berkeley & NBER (dmitry.taubinsky@berkeley.edu). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Antoine Ferey, Benjamin Lockwood, and Dmitry Taubinsky. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Sufficient Statistics for Nonlinear Tax Systems with Preference Heterogeneity
Antoine Ferey, Benjamin Lockwood, and Dmitry Taubinsky
NBER Working Paper No. 29582
December 2021
JEL No. D61,H21,H24

ABSTRACT

This paper provides general and empirically implementable sufficient statistics formulas for optimal nonlinear tax systems in the presence of preference heterogeneity. We study unrestricted tax systems on income and savings (or other commodities) that implement the optimal direct-revelation mechanism, as well as simpler tax systems that impose common restrictions like separability between earnings and savings taxes. We characterize the optimum using familiar elasticity concepts and a sufficient statistic for across-income preference heterogeneity: the difference between the cross-sectional variation of savings with income, and the causal effect of income on savings. The Atkinson-Stiglitz Theorem is a knife-edge case corresponding to zero difference, and a number of other key results in optimal tax theory are subsumed as special cases. Our formulas also apply to other sources of across-income heterogeneity, including heterogeneity in rates of return on savings, inheritances, and the ability to shift income between tax bases. We provide tractable extensions of these results that include multidimensional heterogeneity, additional efficiency rationales for taxing heterogeneous returns, and corrective motives to encourage more saving. Applying these formulas in a calibrated model of the U.S. economy, we find that the optimal savings tax is positive and progressive.

Antoine Ferey
University of Munich
Akademiestr. 1 / II
Wirtschaftspolitikseminar Munich
80799
Germany
antoine.ferey@econ.lmu.de

Dmitry Taubinsky
University of California, Berkeley
Department of Economics
530 Evans Hall #3880
Berkeley, CA 94720-3880
and NBER
dmitry.taubinsky@berkeley.edu

Benjamin Lockwood
The Wharton School
University of Pennsylvania
1400 Steinberg-Dietrich Hall 3620
Locust Walk
Philadelphia, PA 19104
and NBER
ben.lockwood@wharton.upenn.edu

A data appendix is available at <http://www.nber.org/data-appendix/w29582>

1 Introduction

Taxes on capital income, estates, inheritances, and certain categories of consumption are a widespread feature of modern tax systems. Yet there is considerable debate, both among economists and in policy circles, about their optimal design. The celebrated theorem of Atkinson and Stiglitz (1976) is sometimes interpreted to suggest that such taxes should be eliminated. The theorem states that if preferences are homogeneous and weakly separable in consumption and labor, then differential taxes on commodities—including on future consumption in the form of savings—are suboptimal, and welfare is maximized when redistribution is carried out solely through an income tax. However, as was appreciated by contemporaneous work (Mirrlees, 1976) and emphasized by the authors themselves (Stiglitz, 2018), the assumptions underpinning the Atkinson-Stiglitz Theorem are strong, and the theorem does not apply in settings where earnings ability co-varies with commodity preferences, or with other attributes that affect saving levels, such as heterogeneous inheritances, rates of return, or income-shifting abilities.

As a result, an active literature has developed to demonstrate that non-zero commodity and savings taxes may be optimal when the Atkinson-Stiglitz assumptions are relaxed. Yet general, elasticity-based “sufficient statistics” formulas for optimal nonlinear commodity and savings taxes, of the kind common in the optimal income tax literature (e.g., Saez, 2001), have remained elusive. Existing results have instead studied settings with restrictions to a small number of discrete “types,” restrictions on the form of utility functions or tax functions, or they have focused on qualitative insights.¹

In this paper, we derive generally applicable sufficient statistics formulas for optimal linear and nonlinear commodity taxes in a setting where preferences or other consumer attributes, such as inheritances or rates of return, vary with income-earning ability. We study a general version of standard models where consumers with heterogeneous earning abilities and tastes choose labor supply and a consumption and savings bundle that exhausts their after-tax income.² Our formulas nest prior results in this setting, as well as the Atkinson-Stiglitz Theorem itself, as special cases. For concreteness in what follows, we describe results in terms of taxes on savings, although they also apply to other commodities.

We organize the paper around the following key contributions.

The first is a set of results about the optimal unrestricted, nonlinear tax system on earnings and savings. We begin with the question of implementation: Can the optimal allocation be implemented by a smooth (i.e., differentiable) tax on earnings and savings? A smooth tax system allows for *double deviations*, where individuals can jointly alter their earnings and savings to reach bundles not chosen

¹Of particular note, Saez (2002) used a model like the one in this paper to answer the qualitative question of when a “small” *linear* commodity (savings) tax can increase welfare in the presence of preference heterogeneity, but left to future work the task of deriving an expression for the optimal tax, writing “It would of course be extremely useful to obtain optimal commodity tax formulas” in such a framework.

²See, e.g., Atkinson and Stiglitz (1976); Saez (2002); Farhi and Werning (2010); Diamond and Spinnewijn (2011); Golosov et al. (2013); Piketty and Saez (2013); Scheuer and Wolitzky (2016); Saez and Stantcheva (2018); Allcott et al. (2019); Gaubert et al. (2021)

by any other type; such deviations can simply be disallowed under the optimal mechanism. This introduces a complication not present in the standard income taxation model of Mirrlees (1971), nor in the Atkinson-Stiglitz setting: in the presence of preference heterogeneity, double deviations are generally the most attractive direction of adjustment. Nevertheless, we show that under modest regularity conditions, it is possible to construct a smooth tax system, dependent only on earnings and savings, that implements the optimal direct-revelation mechanism.

We then present new elasticity-based formulas for the optimal nonlinear tax on savings and earnings. We show that these formulas can be written entirely in terms of welfare weights and empirically measurable statistics, including a key sufficient statistic for preference heterogeneity: the difference between the cross-sectional variation of savings s with earnings z , denoted $s'(z)$, and the causal effect of income changes on savings, which we denote $s'_{inc}(z)$. The residual, $s'_{pref}(z) := s'(z) - s'_{inc}(z)$, is a sufficient statistic for (local) preference heterogeneity.³ Intuitively, the total derivative of s with respect to z is the sum of two partial derivatives: (i) the causal income effect s'_{inc} , holding preferences constant and (ii) the degree to which higher-ability types prefer more s , holding earnings constant. The second component is captured by s'_{pref} , which we show can be estimated from existing data on the correlational and causal associations with earnings, avoiding the need to explicitly measure or model the relationship between unobserved preferences and ability.

The formula for optimal savings tax rates is a product of s'_{pref} and a term that resembles the optimal income tax formula in Saez (2001), with the elasticity of earnings replaced by the elasticity of savings with respect to the savings tax rate. This result provides an immediate generalization of the Atkinson-Stiglitz Theorem, as it implies that the optimal savings tax rate is everywhere zero when $s'_{pref}(z) = 0$ for all earnings levels z . We also present Pareto-efficiency conditions that use the same sufficient statistics and that can be used to test for (and address) inefficiencies in existing tax systems, without additional assumptions about social marginal welfare weights.

We show that these formulas apply in a variety of other settings that depart from the Atkinson-Stiglitz assumptions, including heterogeneous endowments or inheritances, differential rates of return on investments, human capital investments that enhance productivity, and the ability to engage in income shifting (Slemrod, 1995). In each case, the difference between the cross-sectional profile of savings and the causal income effect on savings, $s'(z) - s'_{inc}(z)$, is the key sufficient statistic for across-income heterogeneity. Consequently, these formulas can be viewed both as a synthesis of prior work that qualitatively studied these extensions in isolation, and as a method for quantifying optimal tax rates when several of these forces are at play simultaneously. For simplicity, we still refer to $s'_{pref}(z)$ as a measure of preference heterogeneity, but we emphasize that it is a sufficient statistic for many other forms of heterogeneity across incomes, or for deviations from weak separability as in, e.g., Corlett and Hague (1953).

Our second contribution is a characterization of what we call “simple tax systems.” We document that across a large number of countries, the tax system consists of a nonlinear tax on income,

³To our knowledge, this statistic was first employed in Allcott et al. (2019), in a setting restricted to a separable linear commodity tax, which of course cannot implement the optimal mechanism.

accompanied by taxes on savings vehicles that can be classified as one of three types: (i) a separable linear (SL) savings tax; (ii) a separable nonlinear (SN) savings tax; or (iii) a system with a linear earnings-dependent (LED) savings tax, which allows, for example, lower-income people to have their savings taxed at a lower linear rate, as is the case for long-term capital gains in the U.S. We show that the optimal tax policy within each of these classes of simple systems can be expressed using the same sufficient statistics that appear in our formulas for the optimal smooth unrestricted tax system. Moreover, we provide sufficient conditions for the SN and LED systems to implement the optimal mechanism. Finally, an advantage of focusing on simple tax systems is that we can extend our results tractably to multidimensional heterogeneity and to a potentially suboptimal income tax. In this more general setting, the causal effect of income on savings, together with the joint distribution of savings and income, remain sufficient statistics for characterizing the optimal savings tax.

We provide further generality in three tractable extensions of our baseline results. First, we consider many dimensions of consumption and savings. For example, different categories of savings might be taxed differently. In this case, the additional necessary sufficient statistics are cross-price elasticities, which allow us to compute *tax diversion ratios*—the fiscal spillovers to taxes collected on goods $j \neq i$ relative to the reduction in taxes collected on good i , when the price of good i is increased. The optimal tax rate on good s_i is the sum of the formula in our baseline result and the tax diversion ratios.

In our second extension, we consider situations where the government wants to alter or correct individual behavior. Our model generalizes the setup of Farhi and Werning (2010), in which the government puts more weight on future generations than the parents, to allow for heterogeneous preferences. Our results also cover the case where individuals under-save due to behavioral biases such as myopia or lack of self control, as in Moser and Olea de Souza e Silva (2019).

In our third extension, we study settings in which there is an additional efficiency rationale for taxing savings, because the government can collect savings taxes either before or after returns are earned, and therefore can arbitrage heterogeneous private rates of return by shifting tax collections onto post-returns savings for high earners. This extension relates to independent work by Gerritsen et al. (2020), who study the special case where all across-income heterogeneity is from differences in rates of return, characterizing and quantifying the optimal separable nonlinear savings tax in terms of model primitives.

In the final part of the paper, we apply these sufficient statistics formulas to study the optimal tax treatment of savings in the U.S. We calibrate the distribution of savings across the income distribution using the Distributional National Accounts micro-files of Piketty et al. (2018). This evidence suggests that savings are approximately constant at low incomes but increase convexly at higher incomes, so that the cross-sectional slope $s'(z)$ is increasing with income. To calibrate the causal income effect on savings, we draw on two sources. The first is Fagereng et al. (2021), which estimates the medium-run marginal propensity to save out of windfall income using lottery prizes. The second is a new probability-based survey representing the U.S. adult population, conducted

on the AmeriSpeak panel, which asked respondents about their savings behavior in response to a possible raise. The two sources are consistent in suggesting similar magnitudes for $s'_{inc}(z)$, with little variation across incomes. Together, these findings yield a positive and increasing value of the residual $s'(z) - s'_{inc}(z) = s'_{pref}(z)$, our sufficient statistic for heterogeneity, across most of the income distribution. Incorporated into our formulas, this implies a (mostly) positive and progressive optimal tax on savings. Our baseline estimates of optimal savings tax rates are somewhat higher than those currently in place in the U.S. across much of the income distribution, although as in other work, these results are sensitive to the elasticity of savings with respect to tax rates, about which there is still substantial uncertainty.

Our paper contributes to a number of literatures. The first is the literature studying optimal commodity and savings taxation in the presence of preference heterogeneity. Saez (2002) considers the special case of a separable linear commodity tax and derives conditions under which its optimal value is non-zero, but does not provide a formula for the magnitude. Golosov et al. (2013) derive conditions characterizing the optimal mechanism in a model like the one we study, but formulate their results in terms of first-order conditions on structural primitives rather than empirically estimable sufficient statistics. Their empirical estimates suggest substantially less across-income heterogeneity than ours do, resulting in much lower optimal savings tax rates. This difference could be because they study heterogeneity in time discounting only, rather than the broader set of forces that can contribute to $s'_{pref}(z)$ and that we allow in our general characterization.⁴ Saez and Stantcheva (2018) study nonlinear capital taxation in a setting without income effects, which corresponds to the special case of our model where $s'_{inc}(z) = 0$ and $s'_{pref}(z) = s'(z)$. They consider multidimensional heterogeneity when tax systems are restricted to be either separable linear or separable nonlinear, so their results can be viewed as a special case of our extension characterizing optimal simple tax systems with multidimensional heterogeneity. Allcott et al. (2019) derive a sufficient statistics formula for the optimal separable linear commodity tax in the presence of preference heterogeneity across incomes.⁵ Our results build on these insights by developing methods to characterize and implement the optimal mechanism using an unrestricted smooth nonlinear tax system, by studying other more restricted but still nonlinear tax systems that are commonly used in practice, and by incorporating forms of across-income heterogeneity that are not just preference-based.

Second, we contribute to the literature studying structural models with various departures from the Atkinson-Stiglitz assumptions. As noted above, our sufficient statistics strategy for quantifying

⁴The lower measured heterogeneity in Golosov et al. (2013) could also be driven by attenuation bias. They measure preference heterogeneity by regressing a structural estimate of time preferences on a plausibly noisy proxy of earnings ability (performance on the Armed Forces Qualification Test), which may be biased toward zero due to a noisy right-hand-side variable.

⁵The application of separable *linear* savings taxes in the presence of multidimensional heterogeneity is also considered in Piketty and Saez (2013), Diamond and Spinnewijn (2011), and Gauthier and Henriet (2018). Piketty and Saez (2013) derive sufficient statistics formulas but make the additional restriction of a linear income tax. Diamond and Spinnewijn (2011) and Gauthier and Henriet (2018) allow for a nonlinear income tax but assume a finite number of possible earnings levels, and derive results in terms of model primitives. Jacquet and Lehmann (2021a) provide a generalization to a separable sum of many one-dimensional nonlinear tax schedules.

preference heterogeneity spans several other sources of across-income heterogeneity, which have been studied separately from one another. Gahvari and Micheletto (2016) and Gerritsen et al. (2020) study heterogeneous rates of return, Boadway et al. (2000) and Cremer et al. (2003) study heterogeneous endowments, Christiansen and Tuomala (2008) study income shifting, and Bovenberg and Jacobs (2005) and Bovenberg and Jacobs (2011) study human capital investments.⁶ Our methods can be viewed as providing a unified treatment of these different sources of across-income heterogeneity., but also a unified approach that can address—using the same set of sufficient statistics—both the growing literature on simpler tax systems with multidimensional heterogeneity and the smaller literature on optimal mechanisms with unidimensional heterogeneity.

Third, this paper complements the literature on dynamic taxation (see overviews by Golosov and Tsyvinski, 2006; Stantcheva, 2020), which typically assumes homogeneous preferences, but derives a theoretically robust role for capital taxation via the inverse Euler equation (e.g., Golosov et al., 2003; Farhi and Werning, 2013). Our work is complementary in relaxing the assumption of homogeneous and weakly separable preferences, but using a static (2-period) framework. Quantitatively, the dynamic taxation literature tends to find optimal savings “wedges” of only several percentage points (see, e.g., Golosov and Tsyvinski, 2015; Golosov et al., 2016; Farhi and Werning, 2013)—substantially lower than those suggested by our baseline calibrations at the same assumed values of elasticities. This suggests that across-income heterogeneity may play a quantitatively larger role in determining optimal savings tax policy than do the social insurance motives analyzed in the dynamic taxation literature, and it motivates future research incorporating our method of measuring and incorporating across-income heterogeneity into fully dynamic models.

The rest of this paper proceeds as follows. Section 2 presents our model and assumptions. Section 3 shows that smooth tax systems can implement the optimal mechanism, and provides sufficient statistics for optimal smooth tax systems. Section 4 studies simple tax systems. Section 5 presents extensions to our results. Section 6 applies our formulas to quantify optimal savings tax rates in the United States. Section 7 concludes. All proofs are gathered in the Appendix.

2 Model and Assumptions

Agents. There is a population of heterogeneous agents who differ in earnings ability and preferences for s , with their types denoted by θ . We begin with the common assumption that $\theta \in \Theta \subset \mathbb{R}$, where Θ is compact; Section 4.2 considers multidimensional heterogeneity. We assume that θ has a continuously differentiable cumulative distribution function $F(\theta)$.

Agents choose earnings z , and a consumption bundle (c, s) , and derive utility $U(c, s, z; \theta)$. One application is where c is period-1 consumption and s is the realized savings in period 2, as in Saez (2002), Golosov et al. (2013), and many others. A second application is where c is period-1 consumption by the parents, while s is the wealth bequeathed to their children and consumed in period 2, as in Farhi and Werning (2010). A third application is where c is numeraire consumption

⁶See Stantcheva (2017) for an analysis of human capital policies in a dynamic setting. Our framework spans the static models in Bovenberg and Jacobs (2005) and Bovenberg and Jacobs (2011) but not more dynamic models.

and s is another dimension of commodity consumption that could be taxed nonlinearly, such as energy or housing.

We assume a linear production technology with marginal rate of transformation p between s and c . In the savings and inheritance interpretations of the model, $p = 1/R$, where R is the gross rate of return in a linear savings technology between the two periods.

Throughout the paper, we assume that:

Assumption 1. $U(c, s, z; \theta)$ is twice continuously differentiable, increasing and weakly concave in c and s , and decreasing and strictly concave in z . The first derivatives U'_c and U'_s are bounded.

For example, a frequently used functional form (e.g. Saez, 2002; Golosov et al., 2013) involves additively separable utility and heterogeneity in agents' productivity w and discount factor δ :

$$U(c, s, z; \theta) = u(c) + \delta(\theta)u(s) - k(z/w(\theta)), \quad (1)$$

with $u(\cdot)$ the utility from consumption and $k(z/w)$ the disutility from work. There is preference heterogeneity across income-earning ability when the discount factor $\delta(\theta)$ covaries with productivity $w(\theta)$.

More generally, we say that there is across-ability preference heterogeneity when marginal rates of substitution between c and s vary with earnings ability. We denote the marginal rate of substitution by

$$\mathcal{S}(c, s, z; \theta) := \frac{U'_s(c, s, z; \theta)}{U'_c(c, s, z; \theta)}, \quad (2)$$

and we use the shorthand $S'_\theta(c, s, z; \theta_0) := \frac{\partial}{\partial \theta} \mathcal{S}(c, s, z; \theta)|_{\theta=\theta_0}$. We define across-ability preference heterogeneity as follows:

Definition 1. *There is across-ability preference heterogeneity for s if some agents prefer different (c, s) bundles conditional on having the same earnings level z ; i.e.,*

$$\exists \theta_0, \forall (c, s, z), S'_\theta(c, s, z; \theta_0) \neq 0. \quad (3)$$

For instance, in the formulation in (1), $S'_\theta(c, s, z; \theta) > 0$ whenever $\delta'(\theta) > 0$. Such across-ability preference heterogeneity is the focus of our results to follow, and for the rest of the paper we will refer to it simply as “preference heterogeneity.”

We similarly define \mathcal{Z} as the marginal rate of substitution between consumption c and earnings z ,

$$\mathcal{Z}(c, s, z; \theta) := \frac{U'_z(c, s, z; \theta)}{U'_c(c, s, z; \theta)}. \quad (4)$$

Government. An agent's type θ is private information and cannot be observed by the government; only the distribution of types, $F(\theta)$, is known. The government must design a tax and transfer system that depends only on the observable variables (c, s, z) .

Without any restrictions on the form of the optimal tax system, the resulting optimal allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$ must solve the following program:

$$\max \int_{\theta} \alpha(\theta) U(c(\theta), s(\theta), z(\theta); \theta) dF(\theta), \quad (5)$$

where $\alpha(\theta)$ represents some set of Pareto weights across types, subject to the resource constraint

$$\int_{\theta} [z(\theta) - ps(\theta) - c(\theta)] dF(\theta) \geq E, \quad (6)$$

where E is an exogenous revenue requirement, and incentive compatibility constraints

$$\forall (\theta, \theta') \in \Theta^2, U(c(\theta), s(\theta), z(\theta); \theta) \geq U(c(\theta'), s(\theta'), z(\theta'); \theta). \quad (7)$$

We refer to an allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$ that maximizes (5) subject to (6) and (7) as the *optimal incentive-compatible allocation*.

3 Optimal Smooth Tax Systems

In this section, we provide two key results about *smooth tax systems*, by which we mean twice continuously differentiable tax functions $\mathcal{T}(s, z)$.⁷ First, we show that the optimal incentive-compatible allocation is implementable by a smooth tax system, under intuitive regularity conditions. Second, we leverage our first result to derive a sufficient statistics characterization of optimal smooth tax systems.

We maintain the following assumptions throughout the rest of our analysis.

Assumption 2. *In the optimal incentive-compatible allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$, c , s , and z are smooth functions of θ . Any type θ strictly prefers its allocation $(c(\theta), s(\theta), z(\theta))$ to the allocation $(c(\theta'), s(\theta'), z(\theta'))$ of another type $\theta' \neq \theta$.*

Assumption 3. *Along the path of $\{c, s, z\}$ offered in the optimal incentive-compatible allocation \mathcal{A} , c and s are smooth functions of z , with c weakly increasing, and the following extended Spence-Mirrlees condition holds:*

$$\mathcal{S}'_{\theta}(c, s, z; \theta) \frac{ds}{dz} + \mathcal{Z}'_{\theta}(c, s, z; \theta) > 0. \quad (8)$$

Assumption 2 is a standard assumption required to apply optimal control methods to characterize the optimal allocation.

The main component of Assumption 3 is the extended Spence-Mirrlees condition, which generalizes the standard assumption, first stated in Mirrlees (1971), that $\mathcal{Z}'_{\theta}(c, s, z; \theta) > 0$. If $\mathcal{Z}'_{\theta}(c, s, z; \theta) > 0$ and s is *increasing* in z , this condition states that the relationship between earnings ability and

⁷Expressing the tax more generally as $\mathcal{T}(c, s, z)$ is redundant. Given such a tax function, any choice of s and z implies a consumption value given by $\mathcal{C}(s, z) := \max\{c | c = z - s - \mathcal{T}(c, s, z)\}$; thus, one can re-express the tax as a function of only savings and earnings: $\tilde{\mathcal{T}}(s, z) = \mathcal{T}(\mathcal{C}(s, z), s, z)$.

preferences for s isn't "too negative." If $\mathcal{Z}'_{\theta}(c, s, z; \theta) > 0$ and s is *decreasing* in z , this condition states that the relationship between earnings ability and preferences for s isn't "too positive." In the savings applications we consider, where evidence suggests that the preference for saving rises with income-earning ability ($\mathcal{S}'_{\theta}(c, s, z; \theta) > 0$) this assumption would be violated only if the optimal mechanism featured a savings allocation that is decreasing with earnings. We have not found examples of such counterintuitive mechanisms in our numerical applications.⁸

One consequence of Assumption 3 and the extended Spence-Mirrlees condition is that $z(\theta)$ is strictly increasing in θ (Appendix Lemma A.1). This allows us to define the function $\vartheta(z)$, which maps each earnings level z to the type to which it is assigned in the optimal incentive-compatible allocation.⁹

3.1 Implementability with Smooth Tax Systems

Definition 2. *We say that an allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_{\theta}$ is implementable with a tax system \mathcal{T} if*

1. \mathcal{T} satisfies type-specific feasibility: $c(\theta) + ps(\theta) + \mathcal{T}(s(\theta), z(\theta)) = z(\theta)$ for all $\theta \in \Theta$, and
2. \mathcal{T} satisfies individual optimization: $(c(\theta), s(\theta), z(\theta))$ maximizes $U(c, s, z; \theta)$ for all $\theta \in \Theta$, subject to the constraint $c + ps + \mathcal{T}(s, z) \leq z$.

Our first result shows that the optimal incentive-compatible allocation is implementable by some smooth tax system.

Theorem 1. *Under Assumption 1, 2, and 3, the optimal incentive-compatible allocation is implementable by a smooth tax system. In this smooth tax system, agents' choices are interior (first-order conditions hold), and their local optima are strict (strict second-order conditions).*

Although it is clear that the optimal incentive-compatible allocation $\{(c(\theta), s(\theta), z(\theta))\}_{\theta}$ can always be implemented by *some* two-dimensional tax system—for example, by defining $\mathcal{T}(s(\theta), z(\theta)) = z(\theta) - c(\theta) - s(\theta)$ for $\theta \in \Theta$ and letting $\mathcal{T}(s, z) \rightarrow \infty$ for $(c, s, z) \notin \{(c(\theta), s(\theta), z(\theta))\}_{\theta}$ —such a tax system is not guaranteed to be smooth. A smooth tax system allows agents to independently adjust s and z locally to points not chosen by any other type in the optimal allocation, and thus the set of possible deviations is much larger than when the optimal mechanism can simply disallow certain allocations.

⁸Note that the assumption that c is increasing could alternatively be characterized not as an assumption about the nature of the optimal mechanism, but rather as a modest assumption on the space of allowable mechanisms. Since the space of weakly increasing functions is compact, an optimal mechanism within this space is guaranteed to exist.

⁹Incentive compatibility implies $\mathcal{S}'_{\theta}(c(\theta), s(\theta), z(\theta); \theta)s'(\theta) + \mathcal{Z}'_{\theta}(c(\theta), s(\theta), z(\theta); \theta)z'(\theta) \geq 0$ for any type θ . Absent preference heterogeneity, $\mathcal{S}'_{\theta} = 0$, and the standard Spence-Mirrlees condition $\mathcal{Z}'_{\theta} > 0$ implies that earnings z increase with type θ in any incentive compatible allocation. With preference heterogeneity, this no longer holds.

Starting from any given allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$, a smooth tax system can implement the allocation only by satisfying each type's θ first-order conditions:

$$\mathcal{T}'_s(s(\theta), z(\theta)) = \mathcal{S}(c(\theta), s(\theta), z(\theta); \theta) - 1 \quad (9)$$

$$\mathcal{T}'_z(s(\theta), z(\theta)) = \mathcal{Z}(c(\theta), s(\theta), z(\theta); \theta) + 1. \quad (10)$$

In the presence of preference heterogeneity, individuals' incentives to deviate from their assigned allocation $(c(\theta), s(\theta), z(\theta))$ are higher under a smooth tax system than under an optimal incentive-compatible mechanism. For example, suppose that higher types θ have a stronger relative preference for s . If they deviate downward to some other earnings level $z(\theta') < z(\theta)$, then under the optimal mechanism they will be forced to choose $s(\theta')$. Under a smooth tax system, however, the deviating type θ will choose $s' > s(\theta')$ at earnings level $z(\theta')$, making this *double deviation* more appealing.

Tax implementation results that involve multidimensional consumption bundles and multidimensional tax systems typically avoid the difficulties associated with double deviations by ruling out the type of heterogeneity that we consider here. Thus, to our knowledge, our proof of Theorem 1 is different from typical implementation proofs in the optimal tax literature. The proof, contained in Appendix B.2, proceeds in three steps. The first step is to construct a sequence of tax systems \mathcal{T}_k such that each element in the sequence satisfies type-specific feasibility and the first-order conditions above. The sequence is ordered such that successive elements are increasingly convex around the bundles $(s(\theta), z(\theta))$ offered in the optimal mechanism.

In the second step of the proof, we show that for each type θ there exists a k sufficiently large such that this type's second-order conditions hold at the point $(c(\theta), s(\theta), z(\theta))$. In other words, for each type there is a sufficiently large k such that $(c(\theta), s(\theta), z(\theta))$ is a *local* optimum under the tax system \mathcal{T}_k . This step requires auxiliary Lemmas B.1 and B.2, which characterize individuals' budget constraints and second derivatives of indirect utility functions for any tax system \mathcal{T} that preserves only the first-order conditions of the optimal mechanism.

In the third step, we show that there exists a sufficiently large k such that $(c(\theta), s(\theta), z(\theta))$ is a *global* optimum for *all* types θ under \mathcal{T}_k . We complete this step via a proof by contradiction. Under the assumption that such a k does not exist, there exists an infinite sequence of values k and types θ_k such that type θ_k prefers to deviate from $(c(\theta_k), s(\theta_k), z(\theta_k))$ under \mathcal{T}_k . Because the type space is compact, the Bolzano-Weierstrass Theorem allows us to extract a convergent subsequence of types θ_j who all prefer to deviate from the allocation assigned to them under the optimal mechanism. We show that this implies a contradiction because, roughly speaking, the limit type of this sequence, $\hat{\theta}$, must then prefer to deviate from $(c(\hat{\theta}), s(\hat{\theta}), z(\hat{\theta}))$ to some other allocation $(c(\theta'), s(\theta'), z(\theta'))$ offered in the optimal mechanism.

Theorem 1 is an existence result, and our proof of the theorem does not offer insight into the structure of an optimal tax system. However, because agents' choices are shown to satisfy first-order and second-order conditions in a smooth tax system, we can use variational methods to characterize optimal tax systems. We now proceed by deriving optimal tax formulas expressed

in terms of empirically estimable sufficient statistics that transparently highlight the key economic forces governing the optimal tax system.

3.2 Sufficient Statistics for Smooth Tax Systems

3.2.1 Definitions

To define the sufficient statistics we use to characterize the optimal tax system, it is helpful to write agents' optimization problem under a tax system $\mathcal{T}(s, z)$ as

$$\max_z \left\{ \max_{c,s} U(c, s, z; \theta) \text{ s.t. } c \leq z - s - \mathcal{T}(s, z) \right\}, \quad (11)$$

where the inner problem represents the optimal choices of $c(z; \theta)$ and $s(z; \theta)$ for a given earnings level z , and the outer problem represents the optimal choice of earnings $z(\theta)$ taking into account endogenous choices of c and s .

Earnings responses to tax reforms are captured through ζ_z^c , the compensated elasticity of labor income with respect to the marginal labor income tax rate, and η_z , the income effect parameter. Formally, for each level of earnings $z(\theta)$ chosen by a type θ , we define

$$\zeta_z^c(z(\theta)) := - \frac{1 - \mathcal{T}'_z(s(\theta), z(\theta))}{z(\theta)} \frac{\partial z(\theta)}{\partial \mathcal{T}'_z(s(\theta), z(\theta))} \quad (12)$$

$$\eta_z(z(\theta)) := -(1 - \mathcal{T}'_z(s(\theta), z(\theta))) \frac{\partial z(\theta)}{\partial \mathcal{T}(s(\theta), z(\theta))} \quad (13)$$

where $\mathcal{T}(s(\theta), z(\theta))$ is the tax liability and $\mathcal{T}'_z(s(\theta), z(\theta))$ is the marginal labor income tax rate. Since the earnings choice takes into account endogenous choices of c and s , these elasticity concepts take into account the full sequence of adjustments due to changes in choices of c and s , as well as those due to any nonlinearities in the tax system.¹⁰

Changes in s in response to tax reforms are captured through $\zeta_{s|z}^c$, the compensated elasticity of s with respect to the marginal tax rate on s , $\eta_{s|z}$, the income effect parameter, and s'_{inc} , the causal effect on consumption of s from a marginal change in gross pre-tax income z . These are formally defined as follows:

$$\zeta_{s|z}^c(z(\theta)) := - \frac{1 + \mathcal{T}'_s(s(z; \theta), z)}{s(z; \theta)} \frac{\partial s(z; \theta)}{\partial \mathcal{T}'_s(s(z; \theta), z)} \Big|_{z=z(\theta)} \quad (14)$$

$$\eta_{s|z}(z(\theta)) := - (1 + \mathcal{T}'_s(s(z; \theta), z)) \frac{\partial s(z; \theta)}{\partial \mathcal{T}(s(z; \theta), z)} \Big|_{z=z(\theta)} \quad (15)$$

$$s'_{inc}(z(\theta)) := \frac{\partial s(z; \theta)}{\partial z} \Big|_{z=z(\theta)} \quad (16)$$

where $\mathcal{T}'_s(s(z; \theta), z)$ is the marginal tax rate on s of a type θ who earns labor income z . Elasticity concepts $\zeta_{s|z}^c$ and $\eta_{s|z}$ are conditional on z . They measure responses to tax reforms and nonlinearities

¹⁰This corresponds to the type of circular adjustment process described in e.g. Jacquet and Lehmann (2021b).

in the tax system, holding labor income z fixed fixed at $z(\theta)$. Note that we define the elasticity of s with respect to one plus the marginal tax rate, rather than with respect to $p + \mathcal{T}'_s$. This choice is natural in applications where s represents savings. However, defining the elasticity with respect to $p + \mathcal{T}'_s$ may be more natural in applications where s is a commodity sold at after-tax price of $q = p + \mathcal{T}'_s$. It is straightforward to convert our results between these elasticity definitions: the key results in Theorem 2 and Proposition 2 can be obtained by multiplying $\zeta_{s|z}^c$ by $(p + \mathcal{T}'_s)/(1 + \mathcal{T}'_s)$.¹¹ For all elasticity concepts, we use the “bar” notation, as in $\overline{\zeta_{s|z}^c}$, to denote a population elasticity.

To quantify preference heterogeneity, we decompose the cross-sectional profile of s . Intuitively, $s'(z)$, the cross-sectional change in s with respect to z comprises both the causal income effect and the degree to which preferences are changing with earnings z . We thus define our measure of local across-income preference heterogeneity, $s'_{pref}(z)$, as the difference between the cross-sectional variation of s along the earnings distribution and the causal income effect $s'_{inc}(z)$:

$$s'_{pref}(z(\theta)) := s'(z(\theta)) - s'_{inc}(z(\theta)) \quad (17)$$

Formally, if we denote by $\vartheta(z)$ the type θ that earns z , $s'(z)$ is a total derivative equal to the sum of two partial derivatives:

$$\frac{ds(z, \vartheta(z))}{dz} = \underbrace{\frac{\partial s(z'; \vartheta(z))}{\partial z'} \Big|_{z'=z}}_{s'_{inc}(z)} + \underbrace{\frac{\partial s(z; \vartheta(z'))}{\partial z'} \Big|_{z'=z}}_{s'_{pref}(z)} \quad (18)$$

The first term, which is equivalent to the definition of $s'_{inc}(z)$ in equation (16), measures how a change in z affects s consumption, holding the type θ constant. The second term, $s'_{pref}(z)$, measures how a change in type affects s consumption, holding earnings z constant.

To illustrate how $s'_{pref}(z)$ relates to model parameters, suppose, as in example (1) above, that $U = \ln c + \delta(\theta) \ln s - k(z/w(\theta))$, where s is period-2 consumption, δ is the discount factor, and that $\mathcal{T}(s, z)$ is a *simple* tax system (see Section 4) that is separable in s and z . Then a z -earner chooses

$$s(z) = \frac{1}{p} \frac{\delta(z)}{1 + \delta(z)} \left(z - \mathcal{T}(s(z), z) \right) \quad (19)$$

where $\delta(z)$ is used to denote the discount factor of the z -earner. Cross-sectional heterogeneity in s is then given by

$$s'(z) = \underbrace{\frac{1}{p} \frac{\delta(z)}{1 + \delta(z)} \left(1 - \mathcal{T}'_z(s(z), z) \right)}_{s'_{inc}(z)} + \underbrace{\frac{1}{p} \left(z - \mathcal{T}(s(z), z) \right) \frac{d}{dz} \left(\frac{\delta(z)}{1 + \delta(z)} \right)}_{s'_{pref}(z)} \quad (20)$$

The causal income effect s'_{inc} is obtained by differentiating (19) with respect to z while holding the discount factor $\delta(z)$ constant. The local preference heterogeneity term $s'_{pref}(z)$ is obtained by

¹¹In this case, the only change in Theorem 2 is that the left-hand-side in equation (24) becomes $\frac{\mathcal{T}'_z(s(z), z)}{p + \mathcal{T}'_s(s(z), z)}$, and analogously for Proposition 2.

differentiating (19) with respect to z while holding after-tax income $z - \mathcal{T}$ constant.

The key insight that facilitates measurement is that $s'(z)$ and $s'_{inc}(z)$ are empirically measurable statistics that can be used to indirectly estimate $s'_{pref}(z)$, which is arguably more tractable than directly estimating how time preferences $\delta(z)$ vary with earnings.

Remarks. While we have referred to s'_{pref} as capturing across-income preference heterogeneity, we clarify that this term also captures failures of weak separability as in, e.g., Corlett and Hague (1953). For example, suppose that $z = w(\theta)l$, where l is labor and $w(\theta)$ is the wage, with $w'(\theta) > 0$. Suppose also that consumption of s and leisure are complements. Then, because higher types θ obtain more leisure for a fixed level of earnings z , higher types will have a stronger preference for s holding z constant. Thus, $s'_{pref} > 0$ in this example.

We also note that s'_{pref} does not capture the direct dependence of preferences for s on z . Because we leave the utility function $U(c, s, z; \theta)$ general, our framework accommodates preferences where, for example, the discount factor δ is directly a function of earnings z , but does not depend on earnings ability. In this case, preferences are heterogeneous and vary with earnings in the cross-section, but $s'_{pref} \equiv 0$. The direct dependence of δ on earnings z is fully captured by the s'_{inc} term.

3.2.2 Measurement

Because $s'(z)$ is the cross-sectional variation of s along the income distribution, it can be directly measured using standard data sources. The statistic $s'_{inc}(z)$ can be measured using a variety of strategies. Here we present three different methods of measuring $s'_{inc}(z)$, which rely on different sources of quasi-experimental variation.

Proposition 1. *Define $\xi_w^s(z)$ as the elasticity of s with respect to the wage rate w , $\xi_w^h(z)$ as the elasticity of hours with respect to the wage rate, and $\chi_s^c(z)$ as the elasticity of s with respect to the marginal net of tax rate on labor income. The sufficient statistic $s'_{inc}(z)$ can be measured as follows:*

- *If preferences are weakly separable and the tax system is separable in s and z , $s'_{inc}(z) = \frac{1-T'_z(z)}{1+T'_s(s(z))} \eta_{s|z}(z)$*
- *If wage rates w and hours h are observable, $s'_{inc}(z) = s(z) \frac{\xi_w^s(z)}{w(z)+h(z)\xi_w^h(z)}$*
- *If responses to tax reforms are measurable, $s'_{inc}(z) = \frac{s(z)}{z} \frac{\chi_s^c(z)}{\zeta_z^c(z)}$,*

If individuals' preferences are weakly separable as in example (1) above, and if the tax system is separable in s and z , then $s'_{inc}(z)$ is proportional to the income effect parameter for s . If individuals' preferences are not weakly separable but wage rates w and hours h are observable, $s'_{inc}(z)$ can be related to the elasticity of s with respect to the wage rate and to the elasticity of hours with respect to the wage rate. If the elasticities of both s and z with respect to the marginal tax rate on z are observable, $s'_{inc}(z)$ can be recovered from these elasticities.

A key question for empirical implementation is the time horizon over which the statistics must be measured. Interpreting our static model to represent a steady-state economy, $s'_{inc}(z)$ corresponds to the causal effect of a change in steady-state labor income on steady-state consumption of s .¹² Under the weak separability assumption, it is therefore necessary to measure the *long-run* marginal propensity to consume s . In the case of savings, this is the long-run marginal propensity to save, as estimated by Fagereng et al. (2021) for example, in response to a change in unearned income.¹³

3.2.3 Social marginal welfare weights

To encode the policymaker’s redistributive objective, we follow the literature in defining social marginal welfare weights as the marginal social welfare derived from an increase in consumption for an individual at a given point in the savings and earnings distribution:

$$g(s(\theta), z(\theta)) := \frac{\alpha(\theta)}{\lambda} U'_c(c(\theta), s(\theta), z(\theta); \theta). \quad (21)$$

We define $\hat{g}(s, z)$ as the social marginal welfare weights augmented with the fiscal impact of income effects. This represents the social value of marginally increasing the disposable income of individuals with savings s and earnings z . Formally,

$$\hat{g}(s, z) := g(s, z) + \mathcal{T}'_z(s, z) \frac{\eta_z(z)}{1 - \mathcal{T}'_z(s, z)} + \mathcal{T}'_s(s, z) \left(\frac{\eta_{s|z}(z)}{1 + \mathcal{T}'_s(s, z)} + s'_{inc}(z) \frac{\eta_z(z)}{1 - \mathcal{T}'_z(s, z)} \right), \quad (22)$$

where the last term comes from the fact that income effects on earnings, proportional to $\eta_z(z)$, induce changes in savings proportional to $s'_{inc}(z)$ affecting savings tax revenues.

Social marginal welfare weights embed judgments about interpersonal utility comparisons. These are usually treated as normative assumptions, although some research has utilized survey data to estimate these weights (see Saez and Stantcheva, 2016, Appendix C) or estimated them from existing policies via an “inverse optimum” procedure (e.g., Bourguignon and Spadaro 2012; Lockwood and Weinzierl 2016). Such normative assumptions are particularly strong when there is preference heterogeneity, because individuals prefer different bundles—and face different tax burdens—even when they have identical budget sets. Lockwood and Weinzierl (2015) show that this

¹²A natural question is whether the effect of income received earlier in life—e.g., family income in childhood—should be used to measure the long-run income effect $s'_{inc}(z)$. It should not. As shown by Lemma 1 below, the role of $s'_{inc}(z)$ is to quantify the distortion in work-life income induced by a change in the steady-state tax on s , and this distortion depends on the causal effect of earnings *during work-life* on s . To the extent that income earlier in life affects s consumption differently from income during work-life, the former behaves like a component of preference heterogeneity.

¹³Fagereng et al. (2021) use lottery winnings as a source of exogenous variation in unearned income. However, if individuals respond differently to a one-time change in unearned income than to a persistent change of equal present value, then $s'_{inc}(z)$ should be measured based on the latter. We discuss this issue, and an alternative measure of $s'_{inc}(z)$ based on survey data, in Section 6.

There is some evidence that mental accounting and other behavioral frictions affect people’s propensity to consume and save out of windfalls. For example, Thakral and To (2021) show that people save more out of long-anticipated windfalls. Since steady-state changes in earnings correspond to anticipated changes earnings, unanticipated windfalls could lead to an under-estimate of s'_{inc} when s corresponds to savings, and an over-estimate when s corresponds to immediate consumption.

difficulty arises even in the standard Mirrlees (1971) model, since it is observationally equivalent to a model with preference heterogeneity over leisure and consumption.

We write our theoretical results in terms of flexible welfare weights that span the degree of heterogeneity in individuals' types, so that optimal policy can be computed using whatever welfare weights the policymaker prefers.¹⁴ For results in the case of unidimensional heterogeneity, welfare weights are written as a function only of income, $g(z(\theta))$, without loss of generality. For results involving multidimensional heterogeneity, in which savings are heterogeneous conditional on income, we write social marginal welfare weights as a function of both savings and income, $g(s(\theta), z(\theta))$.

3.3 Sufficient Statistics Characterization of Optimal Smooth Tax Systems

A key result used to derive our sufficient statistics formula is an equivalence result for tax reforms affecting marginal tax rates on s versus z . The result is a generalization of Lemma 1 in Saez (2002) to arbitrarily nonlinear smooth tax systems.

Lemma 1. *A small increase $d\tau$ in the marginal tax rate on s faced by an individual earning z induces the same earnings change as a small increase $s'_{inc}(z) d\tau$ in the marginal tax rate on z .*

Lemma 1 relates the labor supply distortions induced by increasing taxes on s to the labor supply distortions induced by increasing taxes on earnings z . Intuitively, if the marginal tax rate on earnings z increases by $d\tau_z$, an individual realizes they must now pay an additional $d\tau_z$ on each marginal dollar of earnings, so they earn less in response. Alternatively, if the marginal tax rate on commodity s increases by $d\tau_s$, and the individual adjusts s by s'_{inc} for every dollar adjustment in earnings, then the individual realizes they must now effectively pay an additional $s'_{inc} d\tau_s$ more for each marginal dollar of earnings, accounting for the way in which they will also adjust s . If $d\tau_z = s'_{inc} d\tau_s$, then the induced earnings distortion will be the same from both reforms.

We are now in a position to write formulas characterizing necessary conditions for the optimal smooth tax system in terms of sufficient statistics. In the results that follow, we use $H(s, z)$ and $h(s, z)$ to denote the cumulative and density functions over (s, z) , with h_s and h_z denoting the marginal density over s and z , respectively.

Theorem 2. *Under the assumptions of Theorem 1, at each bundle (c, s, z) chosen by a type θ , an optimal smooth tax system satisfies the following conditions on marginal tax rates on z and s , respectively:*

$$\frac{\mathcal{T}'_z(s, z)}{1 - \mathcal{T}'_z(s, z)} = \frac{1}{z \zeta_z^c(z)} \frac{1}{h_z(z)} \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) - s'_{inc}(z) \frac{\mathcal{T}'_s(s, z)}{1 - \mathcal{T}'_z(s, z)} \quad (23)$$

$$\frac{\mathcal{T}'_s(s, z)}{1 + \mathcal{T}'_s(s, z)} = s'_{pref}(z) \frac{1}{s \zeta_{s|z}^c(z)} \frac{1}{h_z(z)} \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) \quad (24)$$

¹⁴Our empirical application in Section 6 employs a version of the inverse optimum approach, estimating optimal savings taxes consistent with the current U.S. taxes on labor income.

Any Pareto-efficient smooth tax system satisfies

$$\frac{\mathcal{T}'_s(s, z)}{1 + \mathcal{T}'_s(s, z)} = s'_{pref}(z) \frac{z \zeta_z^c(z)}{s \zeta_s^c(z)} \frac{\mathcal{T}'_z(s, z) + s'_{inc}(z) \mathcal{T}'_s(s, z)}{1 - \mathcal{T}'_z(s, z)} \quad (25)$$

Formula (23) constitutes a familiar “ABC” condition analogous to Saez (2001), with one modification: when tax rates on s are non-zero, the formula also accounts for how changes in earnings affect consumption of s , and therefore the revenue from taxes on s .

Formula (24) is one of our key results about optimal marginal tax rates on s . Optimal tax rates on s satisfy a condition that is remarkably similar to the standard “ABC” formula for optimal income tax rates, as presented in equation (23). When $s'_{pref} > 0$, the magnitude of the optimal tax rate at point (s, z) is decreasing in the elasticity of s with respect to the tax rate, increasing in the strength of redistributive motives, and decreasing in the density of individuals at point (s, z) .

Formula (24) also gives a transparent generalization of the Atkinson-Stiglitz Theorem. When the sufficient statistic for preference heterogeneity, s'_{pref} , is equal to zero, the condition implies that the optimal tax on s must equal zero as well. When $s'_{pref} > 0$, implying that higher earners have a stronger preference for s , the condition implies that the optimal tax rate on s must be positive.

We can combine conditions (23) and (24) to derive the Pareto-efficiency condition in (25).¹⁵ Because the condition in (25) does not feature social marginal welfare weights, it is an efficiency condition that must hold for any tax system that is not Pareto dominated. Intuitively, it quantifies the efficient balance between taxing s and taxing z , given the measure s'_{pref} of how tastes for s vary with earnings ability. The stronger the association between preferences for s and earnings ability, the more efficient it is to tax s instead of z .

An implication of the Pareto-efficiency condition in (25) is that in the absence of preference heterogeneity, positive tax rates on s are Pareto dominated, providing an extension of the Atkinson-Stiglitz Theorem to smooth tax systems $\mathcal{T}(s, z)$. On the other hand, any Pareto-efficient tax system must feature non-zero tax rates on s in the presence of preference heterogeneity.

3.4 Other Determinants of Taxes on s Captured by the Sufficient Statistics Formulas

In practice, the elasticities and the measure of across-income preference heterogeneity $s'_{pref}(z)$ in our formulas above may be affected by forces other than pure preferences for s . For example, across-income heterogeneity in prices of s (Gahvari and Micheletto, 2016; Gerritsen et al., 2020), income shifting (Slemrod, 1995; Christiansen and Tuomala, 2008), and heterogeneity in endowments (Boadway et al., 2000; Cremer et al., 2003) may all contribute to differences between the cross-sectional profile $s'(z)$ and the causal income effect $s'_{inc}(z)$. A key feature of our sufficient statistics approach is that the model can be reinterpreted so that $s'_{pref}(z)$ represents these alternative sources of heterogeneity—or a combination of them—and the characterization of optimal tax schedules and

¹⁵See Konishi (1995), Laroque (2005), Kaplow (2006) for Pareto-optimality conditions under the more restrictive assumptions of the Atkinson-Stiglitz theorem. These results are a special case of ours.

“simple” tax systems in Section 4 remain intact. The intuition for this generality stems from the logic of Feldstein (1999), which shows that the elasticity of taxable income is a sufficient statistic for efficiency losses irrespective of whether it is due to real labor supply responses or costly avoidance behavior.¹⁶

While these other sources of across-income differences have previously been studied in isolation to qualitatively assess the robustness of the Atkinson-Stiglitz Theorem, our sufficient statistics techniques can be applied to account for them in a quantitative and general manner. For each of these sources of heterogeneity, we show how it is possible to re-express a model with such heterogeneity in the form of our general model in Section 2, with an appropriately interpreted utility function \tilde{U} . Provided that \tilde{U} satisfies the regularity assumptions in Section 2, our results will carry through.

Heterogeneous Prices. Suppose that individuals face prices $p(e, z, s, \theta)$ that might depend on effort e to seek out lower prices, types θ , earnings z , and the level of s . For example, higher ability types may be better at finding lower prices or higher returns on investments; higher income z might generate beneficial network effects that expose individuals to better opportunities; and higher levels of s might allow individuals to lock in better prices or interest rates (so-called “scale effects”). Modifying our model to allow for these channels, the individual utility function becomes $U(c, s, z, e; \theta)$ and the budget constraint becomes

$$c + p(e, z, s, \theta)s \leq z - \mathcal{T}(s, z). \quad (26)$$

This economy is then equivalent to an economy where $p \equiv 1$ and where individuals maximize the utility function

$$\tilde{U}(c, s, z; \theta) = \max_e U(c + (1 - p(e, z, s, \theta))s, s, z, e; \theta) \quad (27)$$

subject to the budget constraint $c + s \leq z - \mathcal{T}(s, z)$. This is because with a price of $p = p'$ instead of $p = 1$, individuals receive $(1 - p')s$ more consumption c at a given choice s . The feasibility constraint $\int_{\theta} \mathcal{T}(s(\theta), z(\theta)) dF(\theta) \geq E$ is independent of the price p because the tax is deducted directly from an individual’s earnings z .¹⁷ This equivalence shows that it is without loss of generality to assume that $p \equiv 1$, which is a normalization we adopt in our proofs.

An insight from this reinterpretation is that some sources of across-income price heterogeneity justify taxing s , while others do not, and the decomposition of the cross-sectional profile $s'(z)$ into

¹⁶Chetty (2009) suggests limitations to Feldstein’s (1999) results due to some avoidance behaviors generating new types of fiscal externalities, or due to behavioral biases. Variations of these considerations are relevant in our setting as well, as explored in Sections 5.2 and 5.3, respectively.

¹⁷For example, suppose that s represents liters of soda purchased. In our baseline setup, the tax is measured in dollars and is paid out of earnings. If instead the tax were to have a two-part structure where the individual pays some amount T_1 in units of c (dollars) and T_2 in units of s (liters of soda), then the equivalence above would not apply. Although such a system might seem unnatural for a commodity like soda, it is common in the setting of savings, where taxes are often paid in units of “period 2” dollars, after returns have been realized. This creates an additional arbitrage motive to tax some individuals more heavily in units of c and others in units of s . We explore such arbitrage motives in Section 5.3.

$s'_{inc}(z)$ and $s'_{pref}(z)$ correctly distinguishes between them. *Type*-dependent heterogeneity in prices will generally lead to $s'_{pref}(z) \neq 0$ and thus to deviations from the Atkinson-Stiglitz result. For example, consider again the illustrative case from Section 3.2 with $U = \ln c + \delta \ln s - k(z/w(\theta))$ and additively separable \mathcal{T} , and suppose that types θ face heterogeneous prices $p(\theta)$. Then heterogeneity in prices functions like heterogeneity in discount rates, with $s'_{pref}(z) = \frac{\delta}{1+\delta} (z - \mathcal{T}) \frac{d}{dz} \left(\frac{1}{p(z)} \right)$, where $p(z)$ is the price faced by z -earners. In contrast, heterogeneity in prices that is due to *scale* effects does not contribute to s'_{pref} . For example, when $U = \ln c + \delta \ln s - k(z/w(\theta))$, allowing scale effects where p is an increasing function of s or z (but not θ) would affect s'_{inc} , while leaving $s'_{pref} \equiv 0$ when δ is homogeneous.

Income Shifting. Slemrod (1995) argues that some tax systems may provide incentives for individuals to “convert ordinary income into preferentially taxed capital gains,” “convert corporations into pass-through legal entities such as partnerships, or retain labor compensation within the corporation.” Generalizing the two-type model in Christiansen and Tuomala (2008), suppose that individuals can shift some of their labor income to capital income or savings.

Formally, let \tilde{z} and \tilde{s} be the individuals’ true labor income and savings, which are unobserved by the tax authority. Let χ denote the amount of labor income \tilde{z} that individuals shift to be realized as s , and let $m(\chi; \theta)$ denote any financial costs involved in income shifting. Individuals’ taxable labor income is thus $\tilde{z} - \chi$, their taxable realized savings are $\tilde{s} + \chi - m(\chi; \theta)$, and their utility function is $U(c, \tilde{s} + \chi - m(\chi; \theta), \tilde{z}, \chi; \theta)$, where χ might directly influence utility because of effort or psychic costs. This formulation allows for the possibility that individuals with higher earnings ability may be better able to engage in income shifting.

Individuals then choose c , \tilde{s} , \tilde{z} , and χ to maximize their utility subject to the budget constraint

$$c + \tilde{s} + \chi - m(\chi; \theta) \leq \tilde{z} - \mathcal{T}(\tilde{s} + \chi - m(\chi; \theta), \tilde{z} - \chi). \quad (28)$$

Setting $z = \tilde{z} - \chi$, $s = \tilde{s} + \chi - m(\chi; \theta)$, individuals equivalently choose c, s, z, χ to maximize $U(c, s, z + \chi, \chi; \theta)$ subject to the budget constraint $c + s \leq z + \chi - \mathcal{T}(s, z)$. Since individuals exhaust this budget constraint, we can alternatively rewrite this as an economy where individuals choose c, s, z to maximize

$$\tilde{U}(c, s, z; \theta) = \max_{\chi} U(c + \chi, s, z + \chi, \chi; \theta) \quad (29)$$

subject to the budget constraint $c + s \leq z - \mathcal{T}(s, z)$, in which χ no longer appears. Again, the feasibility constraint $\int_{\theta} \mathcal{T}(s(\theta), z(\theta)) dF(\theta) \geq E$ is unaffected, because taxes are always based on observed s and z . In this case, $s'_{pref}(z)$ corresponds to the heterogeneity in taxable savings across taxable labor income that arises from differences in types.

Heterogeneous Endowments. Suppose that individuals have endowments $y_0(\theta)$, from inheritance or other sources, such that their budget constraint is given by $c + s \leq y_0(\theta) + z - \mathcal{T}(s, z)$.

This economy is then equivalent to an economy where $y_0 \equiv 0$ and where individuals maximize the utility function

$$\tilde{U}(c, s, z; \theta) = U(c + y_0(\theta), s, z; \theta) \quad (30)$$

subject to the budget constraint $c + s \leq z - \mathcal{T}(s, z)$. This is because at a given choice s , individuals receive y_0 more consumption of c . Again, the feasibility constraint $\int_{\theta} \mathcal{T}(s(\theta), z(\theta)) dF(\theta) \geq E$ is unaffected, because it is independent of the endowments.

Human Capital Investments. Our framework can also be interpreted to a class of models studying human capital investments. Suppose that s is human capital investments such as education, which affect productivity during working life (Bovenberg and Jacobs, 2005; Stantcheva, 2017). Following the static framework in Bovenberg and Jacobs (2005), let $\phi(s)$ denote the productivity-enhancing effect of human capital and assume it scales down the disutility to generating earnings z . Thus, individuals maximize utility $U(c, s, \frac{z}{\phi(s)}; \theta)$ subject to the budget constraint $c + s \leq z - \mathcal{T}(s, z)$. We can remap this to our model by simply rewriting the utility function:

$$\tilde{U}(c, s, z; \theta) = U\left(c, s, \frac{z}{\phi(s)}; \theta\right) \quad (31)$$

subject to the budget constraint $c + s \leq z - \mathcal{T}(s, z)$. This reflects the fact that human capital investments can be viewed as a specific form of non-separability in preferences that our model accommodates. Again, the feasibility constraint $\int_{\theta} \mathcal{T}(s(\theta), z(\theta)) dF(\theta) \geq E$ is unaffected. The functional form assumptions in Bovenberg and Jacobs (2005) imply that $s'_{pref}(z) < 0$ and thus that human capital investments are subsidized at the optimum; Bovenberg and Jacobs (2011) consider more general settings in which these assumptions are relaxed.

4 Optimal “Simple” Tax Systems and Multidimensional Heterogeneity

In practice, tax systems must be defined by policymakers and implemented by institutions that may impose constraints on the degree of complexity in the tax function, such as separability or linearity. Although the details of these restrictions vary across settings, most taxes on savings and capital income can be classified into a few categories of functional restrictions that we call “simple” tax systems. In this section, we apply our sufficient statistics methods to characterize the optimal policy within each class of simple tax systems. These results can be used to provide policy guidance in settings with practical constraints on tax complexity.

By focusing on restricted tax systems, we can also characterize optimal policy in more general settings with multiple dimensions of heterogeneity. As is well-known, in settings with multidimensional heterogeneity, optimal fully flexible mechanisms tend to feature bunching, randomization, and other irregularities, which are sensitive to model assumptions. As a result, a common approach is to characterize the optimal policy within a restricted class of tax systems using conventional per-

turbation methods.¹⁸ This section extends this literature by considering a more varied set of simple tax systems, and by expressing all results in terms of the sufficient statistics discussed in the previous section. In all cases, we show that $s'(z)$ and $s'_{inc}(z)$, together with standard elasticity concepts, are sufficient to characterize optimal taxes on s .

4.1 A Taxonomy of Common Simple Savings Tax Systems

Many governments tax both labor income and savings (or capital interest income). While these tax systems take a variety of forms, the details of which depend on specifics such as timing, many of these tax policies can be interpreted as a function of earnings and savings, analogous to our function $\mathcal{T}(s, z)$. Table I presents three classes of simple tax systems: separable linear (SL), separable nonlinear (SN), and linear earnings-dependent (LED). Table II categorizes the tax policies on each of five classes of savings-related tax bases—wealth, capital gains, property taxes, pensions, and inheritances—for 21 countries, showing that most of these taxes can be understood as fitting into one of the three simple tax system classes from Table I. In cases where there is ambiguity—such as the distinction between short-term and long-term capital gains in the U.S.—we provide supplementary information in Appendix D.¹⁹

We can see examples of each of the three types of simple tax systems within the United States. Most property taxes, levied at the state and local level, take the form of a separable linear tax, with a flat tax rate, independent of one’s labor earnings, applied to the assessed value of the total property. The estate tax takes the form of a separable nonlinear tax: tax rates rise progressively with the value of the estate, but they do not vary with labor income of the donor or the recipient. And taxes on long-term capital gains and qualified dividends take the form of linear earnings-dependent taxes. In 2020, for example, an individual with \$50,000 in labor earnings faced a tax rate of 15% on long-term capital gains, whereas an individual earning \$500,000 faced a tax rate of 20%. Finally, although we focus on savings tax policies, these classes of simple tax systems are also relevant for other classes of commodities. Separable linear commodity taxes are ubiquitous (e.g., on lodging, airfare, and sales taxes that apply to specific classes of consumption); while separable nonlinear and linear income-dependent tax structures are often used for subsidies on goods like energy and education.²⁰

¹⁸See Piketty and Saez (2013), Diamond and Spinnewijn (2011), and Gauthier and Henriet (2018) for examples restricting to a linear tax on s , and Saez and Stantcheva (2018) for a nonlinear separable tax on s and z .

¹⁹We impose several simplifications for our interpretation of the tax codes. First, we treat ordinary income as consisting primarily of labor income (earnings), written as z in our notation. Second, we separately consider taxes on five broad categories of savings vehicles: wealth, capital gains, real property, private pensions, and inheritances. These categories may overlap—real property is a component of wealth, for example—but we use these groups to reflect the tax instruments that many governments use in practice.

²⁰One practical distinction between taxes on savings and on other commodities involves the measurement of the tax base. In our baseline model, the argument of the tax function s represents the amount of the commodity s consumed. This is natural for many commodities, but in the setting of savings, it is common for the tax system to be written as a function of gross savings before taxes, e.g., a tax $T_1(z)$ in period 1, and a tax $T_2(s_g, z)$ on gross pre-tax savings $s_g = (1 + r)(z - T_1(z) - c)$ in period 2, so that period-2 consumption is given by $s = s_g - T_s(s_g, z)$, where r is the compounded rate of return. In this formulation, a SL structure is one where $T_2(s_g, z) = \tau_s s_g$, a SN structure is one where $T_2(s_g, z)$ is a function of s_g only, and a LED structure is one where $T_2(s_g, z) = \tau_s(z) s_g$. Fortunately,

4.2 Optimal Simple Tax Systems

We now present optimality conditions for SL, SN and LED tax systems. We focus on marginal tax rates on s in the body of the paper, and present conditions for marginal tax rates on z in Appendix A.4. The preference heterogeneity statistic s'_{pref} remains the key sufficient statistic for the marginal tax rate on s . For SL systems, it is convenient to introduce the term $s_{pref}(z) = \int_{x=z_{min}}^z s'_{pref}(x)dx$, which integrates local preference heterogeneity across incomes to obtain a measure of total preference heterogeneity up to earnings level z .

Proposition 2. *Suppose that the optimal SL, SN, and LED systems are smooth and that at the optimum: (i) agents' optima are unique and their first-order and second-order conditions strictly hold, (ii) there is no bunching, (iii) c and s are smooth functions of z , and (iv) in the SN system s is strictly monotonic (increasing or decreasing) in z . Then, at each bundle (c, s, z) chosen by a type θ , these systems satisfy the following optimality conditions:*

$$SL : \frac{\tau_s}{1 + \tau_s} = \frac{1}{\bar{s}\zeta_{s|z}^c} \int_z s'_{pref}(z) \left[\int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) \right] dz \quad (32)$$

$$= -\frac{1}{\bar{s}\zeta_{s|z}^c} Cov[s_{pref}(z), \hat{g}(z)] \quad (33)$$

$$SN : \frac{T'_s(s)}{1 + T'_s(s)} = \frac{1}{s\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} s'_{pref}(z) \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) \quad (34)$$

$$LED : \frac{\tau_s(z)}{1 + \tau_s(z)} = \frac{1}{s\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} s'_{pref}(z) \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) \quad (35)$$

Moreover, if a SL/SN/LED tax system is not Pareto dominated by another SL/SN/LED system, then it must satisfy the following conditions:

$$SL : \frac{\tau_s}{1 + \tau_s} = \frac{1}{\bar{s}\zeta_{s|z}^c} \int_z s'_{pref}(z) z\zeta_z^c(z) \frac{T'_z(z) + s'_{inc}(z)\tau_s}{1 - T'_z(z)} dH_z(z) \quad (36)$$

$$SN : \frac{T'_s(s)}{1 + T'_s(s)} = s'_{pref}(z) \frac{z\zeta_z^c(z)}{s\zeta_{s|z}^c(z)} \frac{T'_z(z) + s'_{inc}(z)T'_s(s)}{1 - T'_z(z)} \quad (37)$$

$$LED : \frac{\tau_s(z)}{1 + \tau_s(z)} = s'_{pref}(z) \frac{z\zeta_z^c(z)}{s\zeta_{s|z}^c(z)} \frac{T'_z(z) + \tau'_s(z)s + s'_{inc}(z)\tau_s(z)}{1 - T'_z(z) - \tau'_s(z)s} \quad (38)$$

The optimal tax formulas and the Pareto-efficiency conditions for SN and LED systems are analogous to the conditions for \mathcal{T}'_s in the fully flexible smooth tax systems derived in Theorem 2.

there is an equivalence between these formulations of two-period tax systems and the type of “static” tax function \mathcal{T} considered in our baseline model, allowing us to translate results between them. Appendix A.6 shows the nature of this equivalence in two steps. First, the set of allocations implementable by these systems is identical, as there is a simple and natural translation between the “static” tax function \mathcal{T} we consider and the two-period function. Second, if $T_1(z) + T_2(s_g, z)$ implements the same allocation as $\mathcal{T}(s, z)$, then $T_1(z) + T_2(s_g, z)$ will be SL/SN/LED if and only if $\mathcal{T}(s, z)$ is SL/SN/LED.

Appendix A.2 derives sufficiency conditions under which SN and LED systems can implement the optimal allocation. In the case of SN systems, these conditions are relatively weak, although they do require that s is weakly increasing with z in the optimal allocation. In contrast, the sufficiency conditions for LED systems are more restrictive, loosely requiring that the local preference for s must not increase “too quickly” across incomes, or else the second-order condition for earnings may be violated. Yet LED systems do not require a non-decreasing schedule of $s(z)$ in the optimal allocation. Thus these results illustrate that SN and LED systems have distinct advantages in different settings.

The SL system is the most restrictive, and generically cannot implement the same allocation as the optimal smooth tax system. This is because the optimal smooth tax system does not generally feature constant marginal tax rates on s . Correspondingly, the optimal tax formula for the SL systems takes a different form. As shown in expression (32), the constant marginal tax rate τ_s for SL systems is in a certain sense an average of the z -earner specific marginal tax rates in expressions (34) and (35). Intuitively, the constant marginal tax rate is a population aggregate of the tax rates that would be optimal for individuals with different earnings levels. This is mirrored in the Pareto-efficiency condition (36). Expression (32) expresses this optimality condition in an alternative way, which was first derived by Allcott et al. (2019). This formulation has a familiar form resembling the Diamond (1975) “Many-person Ramsey tax rule.” The expression is identical to the Diamond (1975) formula when $s_{pref}(z) \equiv s(z)$; i.e., when all consumption differences are driven by preference heterogeneity. This illustrates the relevance of Ramsey taxation principles even in the presence of nonlinear income taxation, as well as their limitations. The SL formula when $s_{pref}(z) \equiv 0$ reduces to the original statement of the Atkinson-Stiglitz Theorem. More generally, even for arbitrarily nonlinear taxes on s , the optimal tax rate is always inversely proportional to the elasticity $\zeta_{s|z}^c(z)$, consistent with standard Ramsey principles, as long as $s'_{pref}(z) \neq 0$.

4.3 Multidimensional Heterogeneity and Suboptimal Earnings Taxes

Our next result generalizes Proposition 2 in two key ways. First, it allows for multidimensional heterogeneity, where types θ belong to a subset of \mathbb{R}^n for $n \geq 2$, so that the support of the distribution of (s, z) is two-dimensional. Second, it characterizes optimal taxes on s even when the earnings tax $T_z(z)$ is not necessarily optimal. Proposition A.5 in the Appendix characterizes optimal earnings tax schedules. The combination of Proposition A.5 and the result below provides a complete characterization of optimal simple tax schedules with multidimensional heterogeneity.

In settings with multidimensional heterogeneity, the relevant sufficient statistics may vary across the joint distribution of s and z . We use $\zeta_z^c(s, z)$ and $s'_{inc}(s, z)$ to denote the compensated elasticity of s and the causal income effect on s for an individual choosing the bundle (s, z) . The formulas below also allow social marginal welfare weights g to be functions of both s and z .

Proposition 3. *Consider smooth simple tax systems with (potentially suboptimal) earnings tax schedules $T_z(z)$ and optimally set marginal tax rates on s . Assume that agents’ first- and second-order conditions hold in these tax systems, and that there is no bunching. Then, at each bundle*

(c^0, s^0, z^0) chosen by some type θ , the marginal tax rates on s in SL/SN/LED systems must satisfy the following optimality conditions:

$$SL : \frac{\tau_s}{1 + \tau_s} \int_z \left\{ \mathbb{E} \left[s \zeta_{s|z}^c(s, z) \middle| z \right] \right\} dH_z(z) = \int_z \left\{ \mathbb{E} \left[(1 - \hat{g}(s, z)) s \middle| z \right] - \mathbb{E} \left[\frac{T'_z(z) + s'_{inc}(s, z) \tau_s}{1 - T'_z(z)} z \zeta_z^c(s, z) s'_{inc}(s, z) \middle| z \right] \right\} dH_z(z) \quad (39)$$

$$SN : \frac{T'_s(s^0)}{1 + T'_s(s^0)} \int_z \left\{ s^0 \zeta_{s|z}^c(s^0, z) \right\} h(s^0, z) dz = \int_z \left\{ \mathbb{E} \left[1 - \hat{g}(s, z) \middle| z, s \geq s^0 \right] \right\} dH_z(z) - \int_z \left\{ \frac{T'_z(z) + s'_{inc}(s^0, z) T'_s(s^0)}{1 - T'_z(z)} z \zeta_z^c(s^0, z) s'_{inc}(s^0, z) \right\} h(s^0, z) dz \quad (40)$$

$$LED : \mathbb{E} \left[\frac{T'_z(z) + \tau'_s(z) s + s'_{inc}(s, z) \tau_s(z)}{1 - T'_z(z) - \tau'_s(z) s} z \zeta_z^c(s, z) s \middle| z^0 \right] h_z(z^0) = \int_{z \geq z^0} \mathbb{E} \left[(1 - \hat{g}(s, z)) s \middle| z \right] dH_z(z) - \int_{z \geq z^0} \left\{ \mathbb{E} \left[\frac{\tau_s(z)}{1 + \tau_s(z)} s \zeta_{s|z}^c(s, z) \middle| z \right] + \mathbb{E} \left[\frac{T'_z(z) + \tau'_s(z) s + s'_{inc}(s, z) \tau_s(z)}{1 - T'_z(z) - \tau'_s(z) s} z \zeta_z^c(s, z) s'_{inc}(s, z) \middle| z \right] \right\} dH_z(z) \quad (41)$$

As in the unidimensional case, the s'_{inc} statistic, together with standard elasticity concepts, allow us to characterize optimal taxes on s in terms of observables. The main difference between multidimensional and unidimensional heterogeneity is that with unidimensional heterogeneity, the expectation operators $\mathbb{E}[\cdot|z]$ are not needed. In expression (39), $\hat{g}(s, z)$, $\zeta_z^c(s, z)$, and $s'_{inc}(s, z)$ are functions only of z in the unidimensional case. In expression (40), the term $\mathbb{E} \left[1 - \hat{g}(s, z) \middle| z, s \geq s^0 \right]$ reduces to $\mathbb{E} \left[1 - \hat{g}(z) \middle| z \geq z^0 \right]$ in the unidimensional case, and the functions s'_{inc} and h can be written as functions of z only. Analogous simplifications apply to expression (41).²¹

If all terms inside the expectation operators $\mathbb{E}[\cdot|z]$ in Proposition 3 are independent of each other, then the expectation can be applied to each statistic separately, and thus the unidimensional formulas are similar to the multidimensional formulas provided that all statistics are interpreted as averages conditional on z . For example, the first term in the integral in expression (39) can be written as $\left(1 - \overline{\hat{g}(z)} \right) \overline{s(z)}$, where the “bar” notation denotes income-conditional averages. The second term in the integral can be written as

$$\frac{T'_z(z) \overline{s'_{inc}(z)} + \tau_s \overline{s'_{inc}(s, z)^2}}{1 - T'_z(z)} \overline{z \zeta_z^c(z)}. \quad (42)$$

The main new effect is the square of s'_{inc} inside the integral. Because $\int (s'_{inc})^2 dH > \left(\int s'_{inc} dH \right)^2$ and because the square enters into the optimal tax expression negatively, this implies that ignoring multidimensional heterogeneity can lead to over-estimates of optimal marginal tax rates on s . The formulas in (40) and (41) also involve squares of s'_{inc} , and thus also imply that multidimensional

²¹For reference, we provide a characterization of optimal taxes on s assuming unidimensional heterogeneity and a given (potentially suboptimal) earnings tax $T_z(z)$, in Appendix A.3.

heterogeneity can lower the optimal tax rate on s through $(s'_{inc})^2$. We quantify the importance of this insight in our empirical application in Section 6. More generally, positive covariances between pairs of statistics inside the expectation operator will tend to lower the optimal tax rate on s , while negative covariances will tend to increase it.

5 Extensions

In this section we provide three key extensions. First, we generalize our results to more than two dimensions of consumptions. This allows us to cover settings where, for example, individuals have access to multiple saving vehicles that are taxed differentially. Second, we allow for the possibility that the government wants people to save more than their perceived private optima, either because of a misalignment between private and social inter-generational preferences or because of individuals' behavioral biases. Third, we consider the case where taxes can be collected both in units of c and in units of s , as is often the case for savings taxes. These extensions highlight that s'_{pref} remains a key sufficient statistic for optimal taxes, and that our previous formulas readily extend to these settings.

5.1 Multiple Goods

We now extend our analysis to a setting where agents consume $n+1$ goods, c and $\mathbf{s} = (s_1, s_2, \dots, s_n)$. For example, \mathbf{s} might correspond to different categories of saving, which the government might choose to tax in different ways. We consider a tax system $\mathcal{T}(\mathbf{s}, z) = \mathcal{T}(s_1, s_2, \dots, s_n, z)$, where we retain the normalization that the numeraire c is untaxed. We normalize $\mathbf{s} = (s_1, s_2, \dots, s_n)$ to measure consumption in units of the numeraire. An individual of type θ then maximizes $U(c, \mathbf{s}, z; \theta)$ subject to the budget constraint $c + \sum_{i=1}^n s_i \leq z - \mathcal{T}(\mathbf{s}, z)$.

We denote own-price elasticities of goods by $\zeta_{s_i|z}^c(z)$, and we define cross-substitution elasticities by $\xi_{s_j,i|z}^c(z) := -\frac{\mathcal{T}'_{s_i}(\mathbf{s}(z;\theta), z)}{s_j(z;\theta)} \frac{\partial s_j(z;\theta)}{\partial \mathcal{T}'_{s_i}(\mathbf{s}(z;\theta), z)} \Big|_{\theta=\vartheta(z)}$, where $s_j(z;\theta)$ denotes type θ consumption of good j when earning labor income z . We denote causal income effects on good s_j by $s'_{j,inc}(z) := \frac{\partial s_j(z;\theta)}{\partial z} \Big|_{\theta=\vartheta(z)}$. We continue using $\hat{g}(z)$ to denote the social marginal welfare effect of increasing a z -earner's consumption of c by one unit.²²

For the result below, as well as for Propositions 5, 6, and the supplementary results in Appendices A.3 and A.4, we assume that

Assumption 4. *The tax systems under consideration are such that at the optimum: (i) these tax systems are smooth, (ii) agents' optima are unique and their first-order and second-order conditions strictly hold, (iii) there is no bunching, (iv) c and s are smooth functions of z , and (v) when the SN system is studied, s is strictly monotonic (increasing or decreasing) in z .*

²²The formula for $\hat{g}(z)$ in this more general setting is in Appendix B.9.

Proposition 4. *With n taxed goods s_1, \dots, s_n , for each good i and at each bundle (c, \mathbf{s}, z) chosen by a type θ , an optimal smooth tax system satisfies*

$$\frac{\mathcal{T}'_{s_i}(\mathbf{s}, z)}{1 + \mathcal{T}'_{s_i}(\mathbf{s}, z)} = s'_{i,pref}(z) \frac{1}{s_i \zeta_{s_i|z}^c(z)} \frac{1}{h_z(z)} \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) + \underbrace{\sum_{j \neq i} \frac{\mathcal{T}'_{s_j}(\mathbf{s}, z)}{\mathcal{T}'_{s_i}(\mathbf{s}, z)} \frac{s_j \xi_{s_j,i|z}^c(z)}{s_i \zeta_{s_i|z}^c(z)}}_{\text{Tax diversion ratio}}. \quad (43)$$

Any Pareto-efficient smooth tax system satisfies

$$\frac{\mathcal{T}'_{s_i}(\mathbf{s}, z)}{1 + \mathcal{T}'_{s_i}(\mathbf{s}, z)} = s'_{i,pref}(z) \frac{z \zeta_z^c(z)}{s_i \zeta_{s_i|z}^c(z)} \frac{\mathcal{T}'_z(\mathbf{s}, z) + \sum_{j=1}^n s'_{j,inc}(z) \mathcal{T}'_{s_j}(\mathbf{s}, z)}{1 - \mathcal{T}'_z(\mathbf{s}, z)} + \underbrace{\sum_{j \neq i} \frac{\mathcal{T}'_{s_j}(\mathbf{s}, z)}{\mathcal{T}'_{s_i}(\mathbf{s}, z)} \frac{s_j \xi_{s_j,i|z}^c(z)}{s_i \zeta_{s_i|z}^c(z)}}_{\text{Tax diversion ratio}}. \quad (44)$$

Proposition 4 features all of the familiar terms of Theorem 2, and includes a novel term that captures the tax implications of substitution effects between the different goods. Intuitively, substituting from s_i to higher-taxed goods generates positive fiscal externalities that motivate higher marginal tax rates on s_i , while substitution to lower-taxed goods generates negative fiscal externalities that motivate lower marginal tax rates on s_i . These effects are summarized by what we call the tax diversion ratio—the impact on taxes collected on goods $j \neq i$ relative to the impact on taxes collected on good i , when the price of good i is increased. The higher is the diversion ratio, the more favorable are the fiscal externalities associated with substitution away from good i , and thus the higher is the optimal tax rate on good i .

5.2 Optimal Taxation when the Government Wants Agents to Save More

Our framework can be interpreted as a bequest model in which parents work and consume in the first period, and leave a bequest to their heirs in the second period. Under this interpretation, our baseline model makes the implicit assumption that the government values bequests in the same way as parents. Farhi and Werning (2010) consider a model where the weight that parents attach to the wellbeing of future generations is too low from a normative perspective. This misalignment introduces a motive to encourage bequests, which we consider in this extension.

Following Farhi and Werning (2010), we assume additively separable preferences given by

$$U(c, s, z; \theta) = u(c; \theta) - k(z; \theta) + \beta v(s; \theta), \quad (45)$$

where $u(c; \theta)$ is the utility parents derive from consumption c , $k(z; \theta)$ is the disutility parents incur to obtain earnings z , $v(s; \theta)$ is the utility heirs derive from a bequest s , and β is the weight parents attach to the wellbeing of their heirs. As in Farhi and Werning (2010), the government maximizes

$$\int_{\theta} [U(c(\theta), s(\theta), z(\theta); \theta) + \nu v(s(\theta); \theta)] dF(\theta), \quad (46)$$

where ν parametrizes the degree of misalignment. Farhi and Werning (2010) microfound ν as the Lagrange multiplier associated with a constraint that the future generation attains a required level of well-being $\int_{\theta} v(s(\theta); \theta) dF(\theta) \geq \underline{V}$.

The formal model above can be interpreted more generally beyond the bequest application, and can be used to analyze behavioral biases as a motivation for encouraging savings. In particular, suppose that $v(s; \theta) = \delta(\theta)u(s; \theta)$, where δ is the “exponential discount factor” and β is “present focus,” as in Laibson (1997). If the government utilizes the “long-run criterion” for welfare, then the degree of misalignment is given by $\nu = (1 - \beta)$.²³ More generally, β may be heterogeneous, so that misalignment is type-dependent and given by $\nu(\theta) = (1 - \beta(\theta))$. For example, Lockwood (2020) summarizes evidence suggesting that individuals with higher earnings ability have lower degrees of present focus.

Below, we characterize optimal taxation with heterogeneous misalignment, where $\beta(z)$ and $\nu(z)$ denote the parameters corresponding to a z -earner. This generalizes the result in Farhi and Werning (2010) by (i) allowing heterogeneity in preferences for s , and by (ii) allowing heterogeneity in the misalignment parameter ν .

Proposition 5. *At each bundle (c, s, z) chosen by a type θ , an optimal smooth tax system satisfies the following marginal tax rate condition*

$$\frac{\mathcal{T}'_s(s, z)}{1 + \mathcal{T}'_s(s, z)} = s'_{pref}(z) \frac{1}{s\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) - \frac{\nu(z)}{\beta(z)} g(z). \quad (47)$$

Any Pareto-efficient smooth tax system satisfies

$$\frac{\mathcal{T}'_s(s, z)}{1 + \mathcal{T}'_s(s, z)} = s'_{pref}(z) \frac{z\zeta_z^c(z)}{s\zeta_{s|z}^c(z)} \left[\frac{\mathcal{T}'_z(s, z) + s'_{inc}(z)\mathcal{T}'_s(s, z)}{1 - \mathcal{T}'_z(s, z)} + s'_{inc}(z) \frac{\nu(z)}{\beta(z)} g(z) \right] - \frac{\nu(z)}{\beta(z)} g(z). \quad (48)$$

This is an intuitive generalization of Theorem 2, where the key new term is a form of Pigovian correction, given by $\frac{\nu(z)}{\beta(z)} g(z)$. As equation (47) shows, the presence of misalignment motivates the government to lower the tax rate on s . The degree by which the government lowers the tax rate depends on the degree of misalignment (relative to the discount factor β), and on the social marginal welfare weight. Because social marginal welfare weights decline with z , equation (47) gives the “progressive estate taxation” result of Farhi and Werning (2010)—i.e., savings subsidies that decline with income—under the special assumptions that (i) $s'_{pref} \equiv 0$ and (ii) $\beta(z) \equiv \beta \in \mathbb{R}$, $\nu(z) \equiv \nu \in \mathbb{R}$. This core result of Farhi and Werning (2010) extends the standard Pigovian taxation logic to optimal screening of distortions with a nonlinear tax.

More generally, Proposition 5 provides a simple formula for balancing the “corrective motives” studied by Farhi and Werning (2010) with the additional motives to tax s in the presence of preference heterogeneity studied in this paper. This extends the Allcott et al. (2019) results for linear commodity taxes with biased consumers to study optimal screening of biases with a nonlinear

²³See Bernheim and Taubinsky (2018) for a detailed discussion of such a criterion, as well as alternative normative approaches to studying the implications of present focus.

tax. If $s'_{pref}(z) > 0$ and $\nu(z)/\beta(z)$ and $g(z)$ are decreasing with z , Proposition 5 suggests a progressive tax on s that can feature subsidies at low incomes and taxes at high incomes.

5.3 Tax Arbitrage with Heterogeneous Prices

Thus far we have considered tax functions where the tax is always paid in units of the numeraire commodity c . In some applications it is also natural to consider tax systems with *multidimensional range*, which include taxes collected in units of c and also in units of s . This is natural, for example, if c and s correspond to period 1 and period 2 consumption, respectively, and taxes must be paid in both periods. The additional richness in the range does not alter the optimal tax implications when the rates of transformation p are homogeneous; in equilibrium, the government's rate of transformation is the same as the homogenous rate for individuals, and it does not matter what portion of the total tax bill is collected in units of s —individuals will simply purchase sufficient s to cover that portion. However, we shall show that when prices are heterogeneous, there is an additional efficiency rationale for differentially taxing s . Intuitively, heterogeneity in prices motivates a form of “tax arbitrage,” in which the government collects relatively more taxes in units of s from individuals who can obtain s at a low price—or in the setting of savings, it imposes relatively higher savings taxes (and lower earnings taxes) on individuals with high rates of returns. This extension provides a bridge between our baseline results and the independent work of Gerritsen et al. (2020), which also studies the role of such efficiency effects.

Formally, suppose that the government uses a two-part tax structure, where individuals pay a tax $T_1(z)$ in units of c and a tax $T_2(s, z)$ in units of s . For instance, in a two-period model where s is savings, T_1 represents the earnings tax levied in period 1 and paid in period-1 dollars, while T_2 represents the savings tax levied in period 2 and paid in period-2 dollars, and $p = 1/(1+r)$ is a function of the rate of return r . For concreteness, we refer to T_1 as period-1 taxes and to T_2 as period-2 taxes, though we emphasize that the presence of efficiency effects is not about dynamics per se, but rather that T_2 is collected in units of s .

Following Gahvari and Micheletto (2016), we consider heterogeneous prices $p(z, \theta)$ which are a function of gross earnings and type. For example, wealthier individuals may have access to better rates of return on savings or prices of commodities. Alternatively, higher earnings ability may be associated with a better ability to obtain high rates of return or to find better prices.

Individuals maximize their utility $U(c, s, z; \theta)$ subject to the budget constraint $c + p(z, \theta)s \leq z - T_1(z) - p(z, \theta)T_2(s, z)$. Denoting by $\vartheta(z)$ the type θ of individuals who choose earnings z , we slightly abuse notation to define $p(z) := p(z, \vartheta(z))$. The government, as before, maximizes a weighted average of utilities,

$$\int_z \left\{ \alpha(z) U\left(z - T_1(z) - p(z)(s(z) + T_2(s(z), z)), s(z), z; \vartheta(z)\right) \right\} dH_z(z), \quad (49)$$

subject to the constraints

$$\int_z T_1(z) dH_z(z) \geq E_1 \quad (50)$$

$$\int_z T_2(s(z), z) dH_z(z) \geq E_2, \quad (51)$$

which generate marginal values of public funds λ_1 and λ_2 . We continue using $\hat{g}(z)$ to denote the social marginal welfare effect of increasing a z -earner's consumption of c by one unit, and it is here normalized by the marginal value of public funds λ_1 .²⁴

Heterogeneity in p generates efficiency effects through two channels. First, for individuals with relatively low $p(z)$, it is efficient for the government to decrease T_1 and increase T_2 . This efficiency effect is present irrespective of the mechanism for the cross-sectional variation of p with z , and leads to a deviation from the Atkinson-Stiglitz Theorem.

Second, lump-sum changes in T_2 trigger novel substitution effects. This is because a lump-sum increase dT in T_2 has the same effect on an agent's utility as a $p(z)dT$ increase in T_1 , and thus changes behavior as much as a marginal tax rate change of $\frac{\partial p}{\partial z}dT$ in T_1 . We denote by $\varphi(z) := -\left(T_1'(z) + \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial z} + s'_{inc}(z) \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s}\right) \frac{z\zeta_s^c(z)}{1-T_1'(z)} \frac{\partial p}{\partial z}$ the fiscal impacts of this substitution effect at earnings z . The impact of a uniform lump-sum change in T_2 is then $\overline{\hat{g}p - \varphi}$, where the "bar" notation is used to denote a population average across all earnings levels. Thus, $\lambda_2/\lambda_1 = \overline{\hat{g}p - \varphi}$, as we formally show in Appendix B.11.3.

Taking these new considerations into account, we characterize optimal taxes on s for SN tax systems where $T_2(s, z) = T_2(s)$, and for LED tax systems where $T_2(s, z) = \tau_s(z) s$. We supplement these results with a characterization of the optimal earnings tax schedule $T_1(z)$ in Appendix B.11.4.

Proposition 6. *With heterogeneous prices, at each bundle (c, s, z) chosen by a type θ , an optimal SN two-part tax system $\{T_1(z), T_2(s)\}$ satisfies*

$$\frac{\lambda_2/\lambda_1 T_2'(s)}{1 + p(z)T_2'(s)} = \frac{1}{s\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} \left\{ s'_{pref}(z) \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) + \frac{s'(z)}{p(z)} \left(\Psi(z) + \int_{x \geq z} [\varphi(x) - \overline{\varphi}] dH_z(x) \right) \right\} \quad (52)$$

where

$$\Psi(z) := \left(1 - H_z(z)\right) \int_{x \leq z} \hat{g}(x) (p(x) - p(z)) dH_z(x) + H_z(z) \int_{x \geq z} \hat{g}(x) (p(z) - p(x)) dH_z(x), \quad (53)$$

An optimal LED two-part tax system $\{T_1(z), \tau_s(z)s\}$ satisfies

$$\begin{aligned} \frac{\lambda_2/\lambda_1 \tau_s(z)}{1 + p(z)\tau_s(z)} &= \frac{1}{s(z)\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} \left\{ s'_{pref}(z) \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) + \frac{p'(z)}{p(z)} s(z) \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) \right\} \\ &+ \frac{1}{\zeta_{s|z}^c(z)} \frac{1}{p(z)} \left\{ \overline{\hat{g}p} - \hat{g}p(z) + \varphi(z) - \overline{\varphi} \right\}. \end{aligned} \quad (54)$$

²⁴The formula for $\hat{g}(z)$ in this more general setting is in Appendix B.11.3.

Proposition 6 shows that the sufficient statistic $s'_{pref}(z)$ remains critical for optimal marginal tax rates on s . On the left-hand side of (52) and (54), the presence of $p(z)$ in the denominator is because an agent's marginal tax rate on s , translated to units of c , is $p(z)\frac{\partial T_2}{\partial s}$. The presence of λ_2/λ_1 in the numerator of the left-hand side is because fiscal externalities generated by substitution away from s must be weighted by the “period 2” marginal value of public funds.

Proposition 6 also introduces novel efficiency terms that lead to taxes on s , even when $s'_{pref} \equiv 0$. In the SN formula, there are two additional efficiency effects. These terms are both positive and thus push toward taxing s when higher earners (i) face lower prices p (e.g., higher rates of returns on savings) and choose higher levels of s , and (ii) exhibit larger substitution effects φ . The first term, proportional to $\Psi(z)$, captures the efficiency effects of increasing period-2 taxes. This term is unambiguously positive when p decreases cross-sectionally with z , and captures the intuition that with a SN system, increasing marginal tax rates on s at any point $z > z_{\min}$ increases period-2 taxes on individuals with below-average p . The second term, proportional to $\int_{x \geq z} [\varphi(x) - \bar{\varphi}] dH_z(x)$, captures the fact that increasing marginal tax rates on s motivates individuals to increase labor supply in order to get lower prices p when $\frac{\partial p}{\partial z} < 0$. The SN formula generalizes the result in Gerritsen et al. (2020) to incorporate other forms of across-income heterogeneity and makes transparent the sign of these terms in a formula employing measurable sufficient statistics.

The implications for LED tax systems are somewhat different. Assume again that p declines cross-sectionally with z (i.e., $p'(z) < 0$). The first novel term in equation (54), proportional to $p'(z)/p(z)$, reflects the fact that higher earners are less responsive to marginal changes in T_2 when $p(z)$ declines with income, since period-2 consumption is “cheaper” for them than period-1 consumption. The second term, proportional to $\bar{g}p - \bar{g}p(z) + \varphi(z) - \bar{\varphi}$, is also negative for sufficiently low values of z , as in this case both $\bar{g}p - \bar{g}p(z)$ and $\varphi(z) - \bar{\varphi}$ are negative. However, this term is positive for sufficiently high values of z . Thus, when $s'_{pref}(z) \equiv 0$, the optimal LED system features subsidies on s for lower-income individuals and taxes on s for higher-income individuals.

The contrast in implications for SN versus LED tax systems—everywhere-positive tax rates in the former, subsidies followed by taxes in the latter—highlights that the new efficiency considerations from heterogeneous rates of return depend on the types of restrictions imposed on the tax system. The reason for this dependence is because positive tax rates on s are a consequence of a missing instrument problem. In a fully flexible tax system, the efficiency gains of taxing a person in period 2 instead of period 1 could be obtained by shifting each individual's total tax burden onto their lowest-cost tax base up to the point that heterogeneous prices are arbitrated away, without the distortion of increasing marginal tax rates on s . But less flexible tax systems can only generate this shifting of the tax burden by altering marginal tax rates on s , and the optimal means of doing this depend on the nature of the restricted tax system.

6 Empirical Application

We now apply our formulas to the question of savings taxes in the United States. We first calibrate the relevant sufficient statistics from micro data and empirical studies, devoting particular attention to the calibration of the sufficient statistic $s'_{pref}(z)$. We then use the Pareto-efficiency conditions derived in Proposition 2 to compute the SL, SN and LED savings tax schedules that would be consistent with the status quo income tax schedule. This allows us to study the welfare-improving reforms that could be made to the existing tax system, taking as given the distributional preferences already embedded in the existing income tax. As is typical for calculations based on sufficient statistics formulas, these results are approximations, as they do not account for changes in the underlying distributions and sufficient statistics that might arise if the savings tax were reformed. These results suggest that across-income heterogeneity leads to a (mostly) positive and progressive schedule of savings tax rates, which range from approximately 0% at the bottom of the income distribution up to between 15% and 20% at higher incomes in our baseline calibration.

6.1 Calibration

We calibrate a model of the U.S. economy that can be interpreted through the lens of our model with a joint savings and income tax function $\mathcal{T}(s, z)$, expressed in terms of the three simple tax systems described in Table I. Appendix C discusses details of this calibration; here, we summarize the key steps. We calibrate a two-period model economy with a fine grid of incomes, where the first period corresponds to work-life and the second to retirement. We assume that these periods are separated by 20 years, and we assume a constant (and, in our baseline, homogeneous) annual rate of return of 3.8% before taxes, drawing from Fagereng et al. (2020). We calibrate a version of the economy with unidimensional heterogeneity (i.e., a single level of savings at each income) and a version with multidimensional heterogeneity, reporting results for each below.

Note that because our model builds on standard models of commodity taxation, it implicitly assumes that z and $\mathcal{T}(s, z)$ are measured in the same units as consumption, which in a dynamic setting corresponds to “period-1” dollars. In practice, savings taxes are typically levied after returns, and they are thus measured in “period-2” dollars. We accordingly translate all tax rates into units of period-2 dollars when reporting results, so that a marginal savings tax rate of 10% indicates that if an individual’s total wealth at retirement increases by \$1, then they must pay an additional \$0.10 in taxes when they retire. Appendix C.1 discusses details of our calibration and this conversion.

To calibrate the earnings and savings distributions—and thus the across-income savings profile $s(z)$ —we use the Distributional National Accounts micro-files of Piketty et al. (2018). We use 2019 measures of pretax labor income (*plinc*) and net personal wealth (*hweal*) at the individual level, as well as the age category (20 to 44 years old, 45 to 64, and above 65). Discretizing the income distribution into percentiles by age group, our measure of annualized earnings during the working life z at the n th percentile is constructed by averaging earnings at the n th percentile across those

aged 20 to 44 and those aged 45 to 64. Our measure of s is the average value of net personal wealth, $hweal$, projected forward to age 65 based on the value observed at each income percentile in the 45-64 age bucket. This measure of wealth includes housing assets, business assets, and financial assets, net of liabilities, as well as defined-contribution pension and life insurance assets.²⁵ We normalize both labor earnings and retirement savings by the number of years worked.

Figure I plots our estimate of gross (i.e., after-returns and before-tax) savings per year worked, across the income distribution. This does not include Social Security, which we model as lump-sum forced savings that are received during retirement. The figure shows that savings upon retirement are approximately zero at low incomes, but increase substantially with income. We convert this to net-of-tax savings using a calibration of savings tax rates across the earnings distribution in the U.S., derived by computing the weighted average of savings tax rates using the asset composition of savings portfolios reported in Bricker et al. (2019); see Appendix C.1.2 for details. The convex shape of the savings profile, which persists after accounting for taxes, indicates that the cross-sectional slope $s'(z)$ rises with income, as shown by the solid blue line in Figure II.

To calibrate the causal income effect on savings, $s'_{inc}(z)$, and thus our measure of local preference heterogeneity $s'_{pref}(z) = s'(z) - s'_{inc}(z)$, we draw from two sources. The first is Fagereng et al. (2021), who estimate the marginal propensity to consume (MPC) out of windfall income across the earnings distribution using information on lottery prizes linked with administrative data in Norway.²⁶ Lottery consumption is widespread in Norway—over 70% of adults from all income groups participated in 2012—and administrative records of asset and wealth holdings allow for direct measures of savings and consumption responses to lottery winnings. They find that individuals' consumption peaks during the winning year and gradually reverts to their previous value afterwards. Over a 5-year horizon, they estimate that winners consume close to 90% of the prize (see their Figure 2, “aggregate consumption response”) which translates into a long-run MPC of 0.9, and a marginal propensity to save of 0.1. They do not find significant heterogeneity across incomes in this MPC. We convert this MPC into a response of net retirement savings to changes in pre-tax labor income using our calibrated schedules of income and savings tax rates.

Our second source of data on $s'_{inc}(z)$ is a new probability-based survey representing the U.S. adult population, conducted on the AmeriSpeak panel in the spring of 2021. In the survey we asked respondents to report how much more they would save each year if they received a hypothetical raise that increased their household's income by \$1000 over the coming years. The relative advantages of this survey are that it is based on the U.S. population and that it asks directly about a modest, persistent change in pre-tax income, rather than a large one-time windfall. The survey results

²⁵The ongoing methodological discussion regarding the different ways to measure wealth (See e.g. Saez and Zucman, 2020; Smith et al., 2021) has important implications for estimates of wealth in the top 1% , but has little impact on the wealth distribution of the rest of the population that we are using here.

²⁶Two other recent studies point to the promise of estimating such causal marginal propensities in a variety of settings. Golosov et al. (2021) study the response to lottery prize winnings in the U.S., although the absence of third-party administrative reporting of wealth in the U.S. complicates the measurement of marginal propensities to save. Straub (2018) estimates the propensity to save out of permanent income, although the absence of quasi-experimental variation in earnings makes it difficult to separate causal income effects from across-income heterogeneity.

suggest a short-run MPC close to that reported in Fagereng et al. (2021), with little variation across incomes. We translate this into a long-run MPC using the response profile of Fagereng et al. (2021). Figure II reports these two schedules of $s'_{inc}(z)$. There is a substantial difference between $s'(z)$ and $s'_{inc}(z)$, which is positive across most of the income distribution and rises with income.²⁷ This is the key force that pushes toward a progressive and mostly positive schedule of optimal savings taxes.

We assume a constant compensated earnings elasticity of $\bar{\zeta}_z^c = 0.33$, drawn from the meta-analysis of Chetty (2012). The value of the savings elasticity $\zeta_{s|z}^c$ is related to the elasticity of taxable wealth (e.g., Jakobsen et al. (2020)) and to the elasticity of capital gains realizations with respect to the capital gains tax (e.g., Agersnap and Zidar (2021)). However, studies that use tax reforms as quasi-experimental variation for identification estimate elasticities that are likely inflated by cross-base responses, as taxpayers re-optimize their savings portfolio towards savings vehicles that are relatively less taxed after the reform.²⁸ We report results for a broad range of values spanning $\zeta_{s|z}^c = 0.7$ to $\zeta_{s|z}^c = 3$, with a baseline of $\zeta_{s|z}^c = 1$, which approximately aligns with the baseline calibration considered in Golosov et al. (2013), in which the intertemporal elasticity of substitution is set to one. Appendix C.1.4 discusses this conversion.

By way of comparison, Golosov et al. (2013) estimate preference heterogeneity by estimating differences in discount factors across ability levels. They infer discount factors from a simple parametric model of savings choice applied to survey data on individuals' household income paths and net worth, and they use survey respondents' results to the Armed Forces Qualification Test (AFQT) as a proxy for ability. In contrast to our findings, their estimation strategy finds very little measured preference heterogeneity, amounting to less than 1% of the cross-sectional variation in savings (see Appendix C.1.3). This discrepancy could be driven by attenuation bias due to measurement error in their proxy for ability—an issue we avoid by computing preference heterogeneity directly as a difference of two statistics rather than from regression analysis. It could also be driven by their use of a narrower measure of across-income heterogeneity based only on time preferences, as opposed to all of the possible forms of across-income heterogeneity that our statistic comprises.

6.2 Results

Figure III reports the schedule of marginal tax rates for SL, SN and LED tax systems that satisfy the Pareto efficiency formulas in Proposition 2, taking the existing U.S. income tax schedule and income distribution as given. In each case, we translate the tax into a marginal tax rate on gross savings at retirement, measured in period-2 dollars. Each panel reports results for a different value of the savings elasticity. For SL tax systems, the linear savings tax rate τ_s is by definition

²⁷Our measure of $s'_{pref}(z)$ appears to be slightly negative at low incomes, which in our simulations gives rise to slight savings subsidies at low incomes. However we note that this could be driven by limitations in our ability to measure $s'_{inc}(z)$ at low incomes. This emphasizes the value of additional empirical research on this statistic.

²⁸Our extension to many goods (Section 5.1) shows how the inclusion of cross-base responses affect optimal savings tax formulas. It could be used to compute the optimal savings tax on different savings vehicles, if there was a larger body of empirical evidence on savings elasticities and cross-base responses.

constant across earnings levels. For LED tax systems, the linear savings tax rate $\tau_s(z)$ is earnings-dependent and we thus report the linear savings tax rate at each earnings level. For SN tax systems, the nonlinear savings tax schedule $T_s(s)$ depends on the value of savings s , and not on earnings z . But to make the SN system visually comparable to the other systems, we plot the marginal savings tax rate faced at the margin by each earner, given their level of saving (represented on Figure I).

In each panel, marginal savings tax rates are mostly positive, and the nonlinear tax schedules are progressive, with marginal rates increasing with income. The magnitudes depend on the value of the savings elasticity parameter. In the baseline case of $\zeta_{s|z}^c = 1$, savings tax rates in SN and LED tax systems average approximately 0% for annual incomes below \$50,000, then steadily increase up to nearly 20% for annual incomes around \$200,000, remaining stable thereafter. The savings tax rate in a SL tax system is approximately 6%. Changing the savings elasticity parameter scales the efficient savings tax rates without affecting the overall pattern: across-income heterogeneity calls for (mostly) positive and progressive savings tax rates. At the lower elasticity values, our estimates of optimal tax rates are substantially higher than the prevailing savings tax rates in the U.S., which are also shown in Figure III.

Figure IV considers two key extensions to these results: multidimensional heterogeneity, and heterogeneous rates of return with “tax arbitrage” efficiency effects, as discussed in Section 5.3. For comparability with our baseline results, all other parameters, including elasticity parameters and income-dependent welfare weights, are held fixed at the values from our baseline calibration. These results are computed using our baseline savings elasticity of $\zeta_{s|z}^c = 1$. We plot both types of nonlinear tax schedules, LED and SN, omitting the separable linear plots for legibility.

In the case of multidimensional heterogeneity, we use the same measure of gross savings, but rather than compute average savings at each income, we partition the population into four levels of savings at each level of income, representing quartiles of the income-conditional savings distribution.

In the case of heterogeneous rates of return, we follow Gerritsen et al. (2020) who, relying on empirical work by Fagereng et al. (2020), assume that rates of return rise by 1.4 percentage points from the bottom to the top of the income distribution. We linearly interpolate this difference across income percentiles, centered on our 3.8% baseline rate of return.

Consistent with the intuition described in Section 4.3, the top two panels of Figure IV show that incorporating multidimensional heterogeneity reduces the magnitude of optimal tax rates LED systems (top left panel) and in SN systems (top right panel). The effect is particularly pronounced for SN systems, where savings tax rates are plotted as a function of total savings at the time of retirement, since agents with the same income save different amounts and thus face different savings tax rates. In this extension, marginal savings tax rates are still progressive, and are above status quo savings tax rates across high incomes in our baseline specification.

The bottom two panels show that the presence of heterogeneous rates of returns tends to significantly raise optimal savings tax rates, reflecting the efficiency effects of tax arbitrage highlighted in Proposition 6.²⁹ The bottom right panel shows that tax rates in the SN system are higher at

²⁹Consistent with the tax arbitrage interpretation, these efficiency effects are (almost) unaffected by whether return

all levels of income, consistent with our discussion of the formula for SN systems in Proposition 6. On the other hand, recall that the formula for LED systems implied lower savings tax rates at low incomes and higher tax rates at higher incomes. Consistent with this, the bottom left panel shows that relative to the baseline, the optimal savings tax rates with heterogeneous rates of return are even more progressive. For example, substantial savings subsidies are optimal for incomes below about \$40,000, whereas savings taxes are substantially higher at higher incomes.

Taken together, our empirical results show a robust role for progressive savings taxes, stemming from across-income heterogeneity captured in the s'_{pref} statistic. This highlights the importance of this sufficient statistic and motivates additional empirical work estimating the long-run marginal propensity to save out of earned income, as well as across-income consumption profiles and causal income effects in other applications. Moreover, our empirical results show that policy implications depend in important ways on the type of simple tax system in question, and they demonstrate the quantitative role of multidimensional heterogeneity and additional tax arbitrage efficiency effects.

7 Conclusion

This paper characterizes optimal smooth tax systems on earnings and savings (or other dimensions of consumption) in the presence of across-income heterogeneity. We first prove that with unidimensional heterogeneity, the optimal allocation can be implemented by a smooth tax on earnings and savings. We then derive formulas which characterize the optimal smooth tax system, expressed using familiar empirical statistics, as well as a key sufficient statistic for preference heterogeneity, $s'_{pref}(z)$. This statistic can be estimated from empirical data, and can also accommodate other dimensions of heterogeneity such as heterogeneous rates of return, endowments, or income-shifting abilities. We then consider a set of “simple” separable tax systems that are widely used in practice. We derive intuitive sufficient statistics formulas for these separable tax systems, under both unidimensional and multidimensional heterogeneity. We also provide tractable extensions to multiple goods, corrective motives, and heterogeneous prices with “tax arbitrage” efficiency effects. Finally, we apply our theoretical formulas to the setting of savings taxes in the U.S.. Results suggest that the savings tax rates that would be consistent with the existing income tax are progressive and (mostly) positive. Together, these results provide a practical and general method for quantifying optimal tax systems for savings, inheritances, and other commodities in the presence of across-income heterogeneity.

heterogeneity is driven by income (scale-dependence) or by ability (type-dependence).

References

- Agersnap, O. and Zidar, O. M. (2021). The tax elasticity of capital gains and revenue-maximizing rates. *American Economic Review: Insights*, 3(4):399–416.
- Allcott, H., Lockwood, B. B., and Taubinsky, D. (2019). Regressive sin taxes, with an application to the optimal soda tax. *Quarterly Journal of Economics*, 134(3):1557–1626.
- Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: direct versus indirect taxation. *Journal of Public Economics*, 6(1-2):55–75.
- Bernheim, B. D. and Taubinsky, D. (2018). Behavioral public economics. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics*, volume 1, pages 381–516. Elsevier, New York.
- Boadway, R., Marchand, M., and Pestieau, P. (2000). Redistribution with unobservable bequests: A case for taxing capital income. *The Scandinavian Journal of Economics*, 102(2):253–267.
- Bourguignon, F. and Spadaro, A. (2012). Tax–benefit revealed social preferences. *The Journal of Economic Inequality*, 10(1):75–108.
- Bovenberg, A. L. and Jacobs, B. (2005). Redistribution and education subsidies are siamese twins. *Journal of Public Economics*, 89(11-12).
- Bovenberg, A. L. and Jacobs, B. (2011). Optimal taxation of human capital and the earnings function. *Journal of Public Economic Theory*, 13(6):957–971.
- Bricker, J., Moore, K. B., and Thompson, J. P. (2019). Trends in household portfolio composition. *Finance and Economics Discussion Series 2019-069*. Washington: Board of Governors of the Federal Reserve System.
- Chetty, R. (2009). Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance. *American Economic Journal: Economic Policy*, 1(2):31–52.
- Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018.
- Christiansen, V. and Tuomala, M. (2008). On taxing capital income with income shifting. *International Tax and Public Finance*, 15:527–545.
- Corlett, W. J. and Hague, D. C. (1953). Complementarity and the excess burden of taxation. *The Review of Economic Studies*, 21(1):21–30.
- Cremer, H., Pestieau, P., and Rochet, J.-C. (2003). Capital income taxation when inherited wealth is not observable. *Journal of Public Economics*, 87:2475–2490.

- Diamond, P. (1975). A many-person ramsey tax rule. *Journal of Public Economics*, 4(4):335–342.
- Diamond, P. and Spinnewijn, J. (2011). Capital income taxes with heterogeneous discount rates. *American Economic Journal: Economic Policy*, 3(4):52–76.
- Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2020). Heterogeneity and persistence in returns to wealth. *Econometrica*, 88(1):115–170.
- Fagereng, A., Holm, M. B., and Natvik, G. J. (2021). MPC heterogeneity and household balance sheets. *American Economic Journal: Macroeconomics*, 13(4):1–54.
- Farhi, E. and Werning, I. (2010). Progressive estate taxation. *Quarterly Journal of Economics*, 125(2):635–673.
- Farhi, E. and Werning, I. (2013). Insurance and taxation over the life cycle. *Review of Economic Studies*, 80:596–635.
- Feldstein, M. S. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics*, 81(4):674–680.
- Gahvari, F. and Micheletto, L. (2016). Capital income taxation and the Atkinson–Stiglitz theorem. *Economics Letters*, 147:86–89.
- Gaubert, C., Kline, P., and Yagan, D. (2021). Place-based redistribution. *Working Paper no. 28337, National Bureau of Economic Research*.
- Gauthier, S. and Henriët, F. (2018). Commodity taxes and taste heterogeneity. *European Economic Review*, 101:284–296.
- Gerritsen, A., Jacobs, B., Rusu, A. V., and Spiritus, K. (2020). Optimal taxation of capital income with heterogeneous rates of return. *Working Paper no. 8395, CESifo*.
- Golosov, M., Graber, M., Mogstad, M., and Novgorodsky, D. (2021). How Americans respond to idiosyncratic and exogenous changes in household wealth and unearned income. *Working Paper no. 29000, National Bureau of Economic Research*.
- Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003). Optimal indirect and capital taxation. *Review of Economic Studies*, 70(3):569–588.
- Golosov, M., Troshkin, M., and Tsyvinski, A. (2016). Redistribution and social insurance. *American Economic Review*, 106(2):359–386.
- Golosov, M., Troshkin, M., Tsyvinski, A., and Weinzierl, M. (2013). Preference heterogeneity and optimal capital income taxation. *Journal of Public Economics*, 97:160–175.
- Golosov, M. and Tsyvinski, A. (2006). New dynamic public finance: a user’s guide. *NBER Macroeconomics Annual*, 317–379.

- Golosov, M. and Tsyvinski, A. (2015). Policy implications of dynamic public finance. *Annual Reviews of Economics*, 7:147–171.
- Jacquet, L. and Lehmann, E. (2021a). How to tax different incomes? *Discussion Paper no. 14739, IZA Institute of Labor Economics*.
- Jacquet, L. and Lehmann, E. (2021b). Optimal income taxation with composition effects. *Journal of the European Economic Association*, 19(2):1299–1341.
- Jakobsen, K., Jakobsen, K., Kleven, H., and Zucman, G. (2020). Wealth taxation and wealth accumulation: Theory and evidence from Denmark. *The Quarterly Journal of Economics*, 135(1):329–388.
- Kaplow, L. (2006). On the undesirability of commodity taxation even when income taxation is not optimal. *Journal of Public Economics*, 90(6-7):1235–1250.
- Konishi, H. (1995). A Pareto-improving commodity tax reform under a smooth nonlinear income tax. *Journal of Public Economics*, 56(3):413–446.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–478.
- Laroque, G. R. (2005). Indirect taxation is superfluous under separability and taste homogeneity: A simple proof. *Economics Letters*, 87(1):141–144.
- Lockwood, B. B. (2020). Optimal income taxation with present bias. *American Economic Journal: Economic Policy*, 12(4):298–327.
- Lockwood, B. B. and Weinzierl, M. (2015). De gustibus non est taxandum: Heterogeneity in preferences and optimal redistribution. *Journal of Public Economics*, 124:74–80.
- Lockwood, B. B. and Weinzierl, M. (2016). Positive and normative judgments implicit in U.S. tax policy, and the costs of unequal growth and recessions. *Journal of Monetary Economics*, 77:30–47.
- Milgrom, P. and Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601.
- Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208.
- Mirrlees, J. A. (1976). Optimal tax theory: a synthesis. *Journal of Public Economics*, 6(4):327–358.
- Moser, C. and Olea de Souza e Silva, P. (2019). Optimal paternalistic savings policies. *Institute Working Paper 17, Federal Reserve Bank of Minneapolis*.

- Piketty, T. and Saez, E. (2013). A theory of optimal inheritance taxation. *Econometrica*, 81(5):1851–1886.
- Piketty, T., Saez, E., and Zucman, G. (2018). Distributional national accounts: methods and estimates for the United States. *The Quarterly Journal of Economics*, 133(2):553–609.
- Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The Review of Economic Studies*, 68(1):205–229.
- Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics*, 83(2):217–230.
- Saez, E. and Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, 106(1):24–45.
- Saez, E. and Stantcheva, S. (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics*, 162:120–142.
- Saez, E. and Zucman, G. (2020). Trends in US income and wealth inequality: Revising after the revisionists. *Working Paper no. 27921, National Bureau of Economic Research*.
- Scheuer, F. and Wolitzky, A. (2016). Capital taxation under political constraints. *American Economic Review*, 106:2304–2328.
- Slemrod, J. (1995). Income creation or income shifting? Behavioral responses to the Tax Reform Act of 1986. *The American Economic Review*, 85(2):175–180.
- Smith, M., Zidar, O. M., and Zwick, E. (2021). Top wealth in America: New estimates and implications for taxing the rich. *Working Paper no. 29374, National Bureau of Economic Research*.
- Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy*, 125(6):1931–1990.
- Stantcheva, S. (2020). Dynamic taxation. *Annual Review of Economics*, 12:801–831.
- Stiglitz, J. E. (2018). Pareto efficient taxation and expenditures: pre-and re-distribution. *Journal of Public Economics*, 162:101–119.
- Straub, L. (2018). Consumption, savings, and the distribution of permanent income. *Unpublished manuscript, Harvard University*.
- Thakral, N. and To, L. (2021). Anticipation and consumption. *working paper*.

Tables & Figures

Table I: Types of simple tax systems

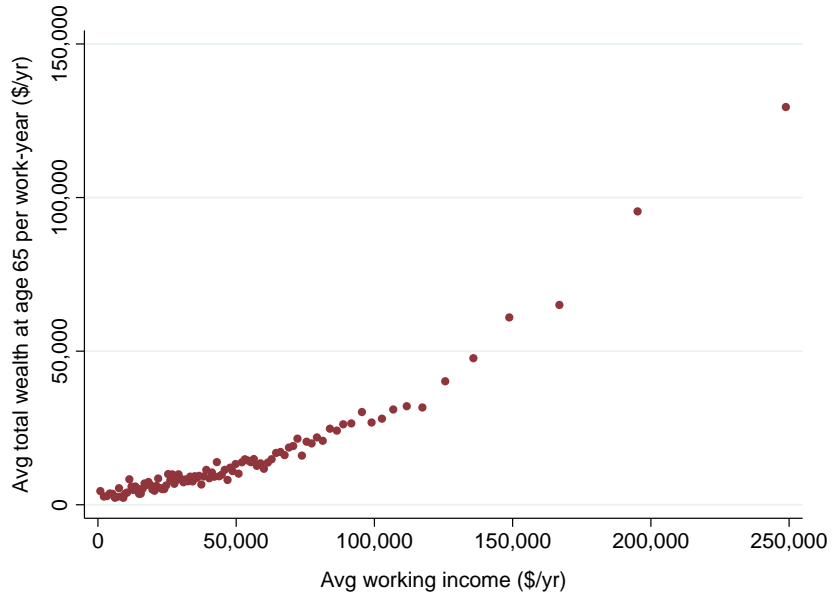
Type of simple tax system	$\mathcal{T}(s, z)$	$\mathcal{T}'_s(s, z)$	$\mathcal{T}'_z(s, z)$
SL: separable linear	$\tau_s s + T_z(z)$	τ_s	$T'_z(z)$
SN: separable nonlinear	$T_s(s) + T_z(z)$	$T'_s(s)$	$T'_z(z)$
LED: linear earnings-dependent	$\tau_s(z) s + T_z(z)$	$\tau_s(z)$	$T'_z(z) + \tau'_s(z) s$

Table II: Tax systems applied to different savings vehicles, by country.

Country	Wealth	Capital Gains	Property	Pensions	Inheritance
Australia	–	Other	SL, SN	SL	–
Austria	–	Other	SL, SN	SN	–
Canada	–	Other	SL	SN	–
Denmark	–	SN	SL, SN	SL, SN	SN
France	–	Other	Other	SL, SN	SN
Germany	–	Other	SL	SN	SN
Ireland	–	SN	SL, SN	SN	SN
Israel	–	Other	Other	SN	–
Italy	SL, SN	SL	SL	SL	SL, SN
Japan	–	SL, SN	SN	SN	SN
Netherlands	SN	SL	SL, SN	SN	SN
New Zealand	–	Other	SN	SL, LED	–
Norway	SN	SL	SL	SN	–
Portugal	–	SL	Other	SN	SL
Singapore	–	Other	SN	SN	–
South Korea	–	SN	SN	SN	SN
Spain	SN	SN	SL, SN	SN	SN
Switzerland	SN	SN	SL, SN	SN	SN
Taiwan	–	SL, SN	SL, SN	SN	SN
United Kingdom	–	Other	SN	SN	SN
United States	–	LED	SL	SN	SN

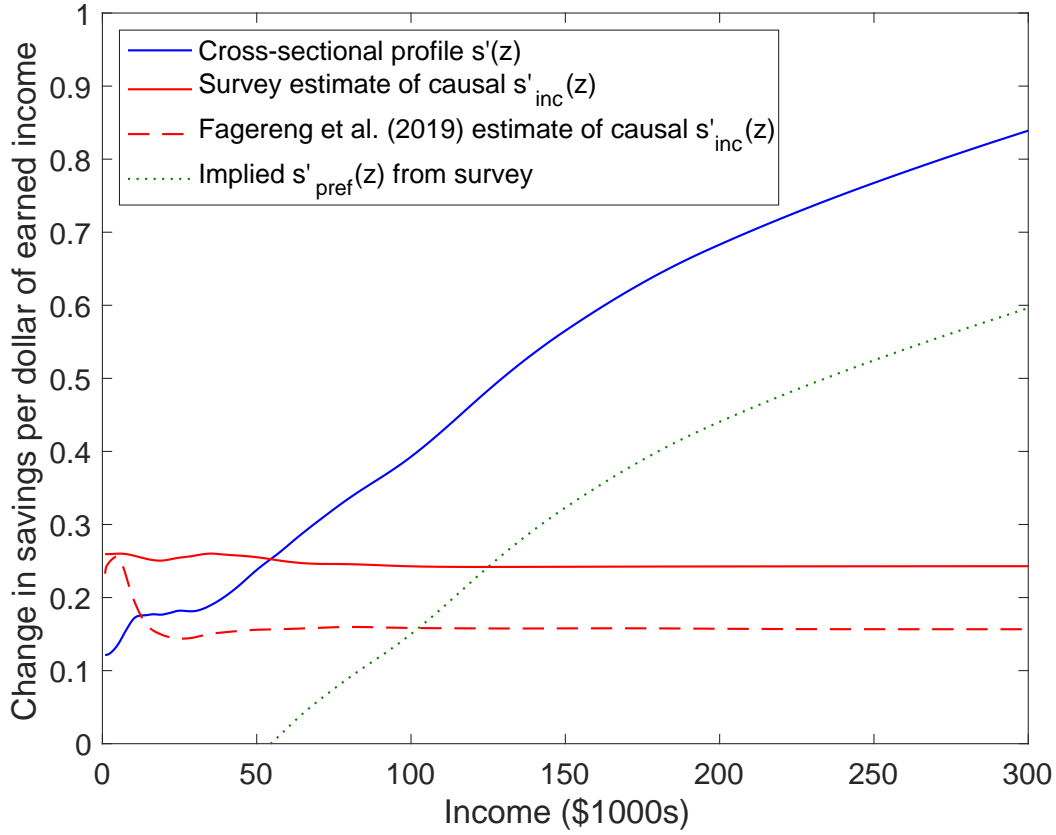
Notes: This table classifies tax systems applied to different savings vehicles across countries in 2020 according to the types in Table I. See Appendix D for further details.

Figure I: Savings Across Incomes in the United States



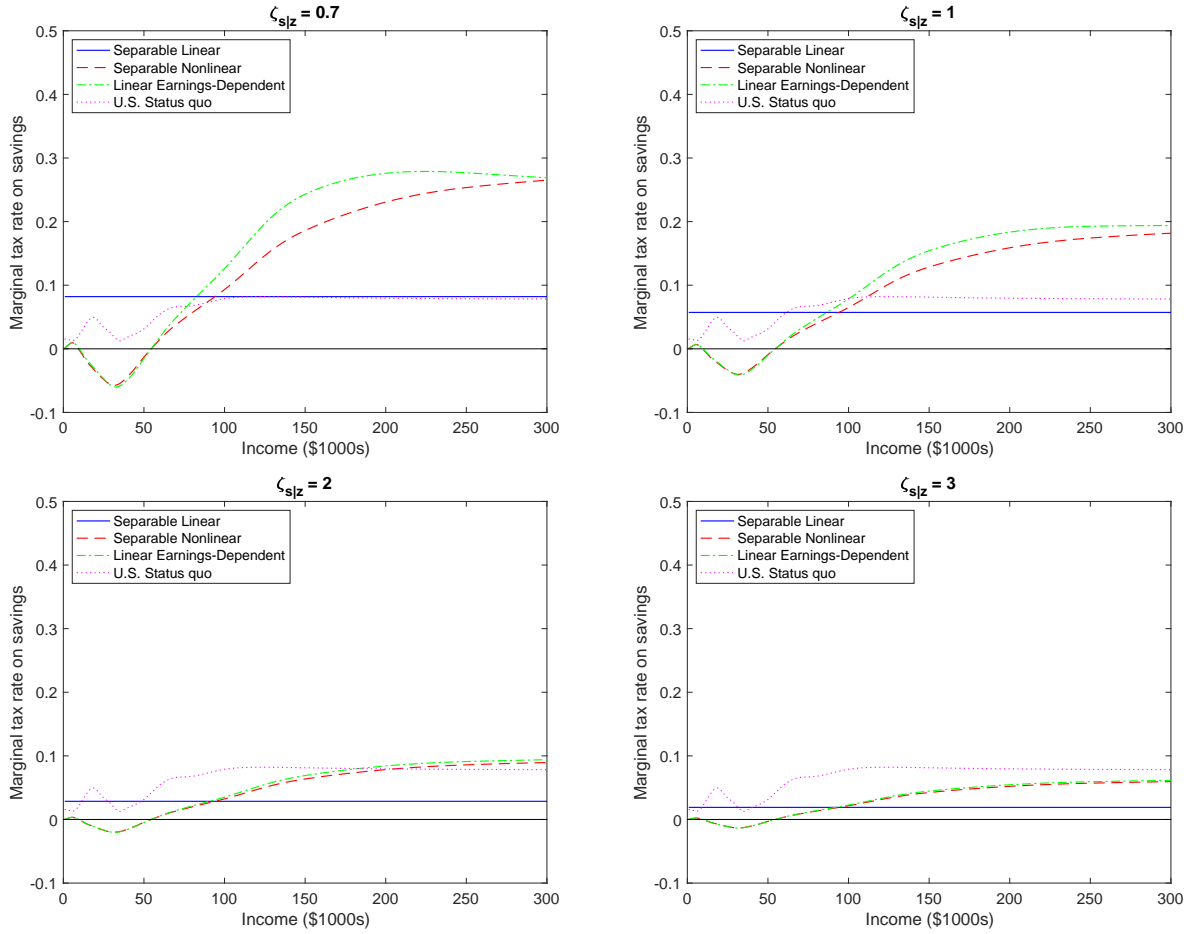
Notes: The earnings and savings distribution in the U.S. is calibrated based on the Distributional National Accounts micro-files of Piketty et al. (2018). We use 2019 measures of pretax income (*plinc*) and net personal wealth (*hweal*) at the individual level, as well as the age category (20 to 44 years old, 45 to 64, and above 65) to impute gross savings at the time of retirement, which we normalize by the number of work years. See Appendix C.1 for further details.

Figure II: Decomposition of Cross-Sectional Savings Profile into Income Effects and Preference Heterogeneity



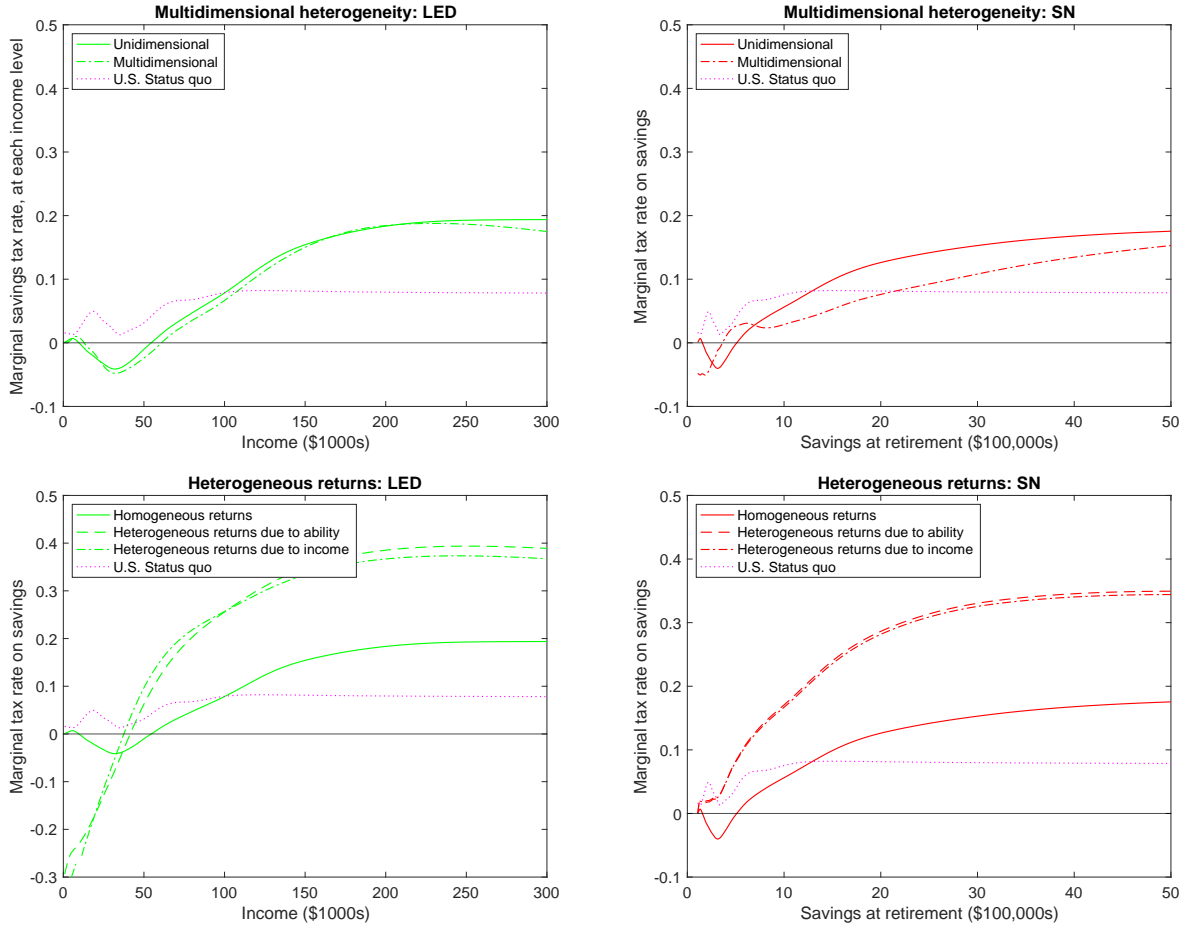
Notes: This figure reports the slope of the cross-sectional profile of savings $s'(z)$ (blue), as well as our calibrations of $s'_{inc}(z)$ based on causal income effects, derived from Fagereng et al. (2021) and from a new nationally representative survey. See Section 6 and Appendix C.1 for details.

Figure III: Savings Tax Rates Implied by Pareto-Efficiency Formulas



Notes: This figure presents the marginal savings tax rates values that satisfy the Pareto-efficiency formulas in Proposition 2, plotted against the earnings level to which they apply. We plot these schedules for four different values of the savings elasticity $\zeta_{s|z}$, with $\zeta_{s|z} = 1$ representing our baseline case.

Figure IV: Effects of Multidimensional Heterogeneity and Heterogeneous Returns



Notes: This figure plots the marginal savings tax rate schedules which are optimal, according to the first-order condition formulas presented in the text, for two extensions discussed in Section 5: multidimensional heterogeneity (top row), and heterogeneous returns (bottom row). All plots also reproduce the Pareto-efficient savings schedules from Figure III for comparison, as well as the status quo U.S. savings taxes. These plots use the same set of social welfare weights, calibrated to rationalize the status quo income tax in the unidimensional model. The Linear Earnings-Dependent (LED) schedules, in the left column, are plotted across earnings during work-life. The Separable Nonlinear (SN) schedules, in the right column, cannot be plotted this way, because individuals with a given income have different levels of savings and are thus subject to different savings taxes. We therefore plot them over total savings at the time of retirement. See Section 6 and Appendices C.2 and C.3 for details.