

NBER WORKING PAPER SERIES

SUFFICIENT STATISTICS FOR NONLINEAR TAX SYSTEMS
WITH GENERAL ACROSS-INCOME 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, Revised April 2022

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 or the National Bureau of Economic Research. We thank Pierre Boyer, Philippe Choné, Ashley Craig, Nathan Hendren, Laurence Jacquet, Claus Kreiner, 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, University of Michigan, and LMU Munich for their valuable comments. Supplementary material available at https://benlockwood.com/papers/FLT_Supplement.pdf.

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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.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 General Across-Income Heterogeneity
Antoine Ferey, Benjamin Lockwood, and Dmitry Taubinsky
NBER Working Paper No. 29582
December 2021, Revised April 2022
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 across-income heterogeneity in preferences, inheritances, income-shifting capabilities, and other sources. 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 general across-income 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. 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
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 covaries with preferences for commodities or with other attributes 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 modern optimal income tax literature (e.g., Saez, 2001; Rothschild and Scheuer, 2013; Sachs, Tsyvinski, and Werquin, 2020; Hendren, 2020; Bierbrauer, Boyer, and Hansen, 2020), have remained elusive. Existing results have instead studied settings with restrictions to a small number of discrete “types” or restrictions on the form of utility functions or tax functions, or they have focused on qualitative insights.¹

In this paper, we derive sufficient statistics formulas for optimal linear and nonlinear commodity taxes in a general setting where attributes other than ability—such as preferences, inheritances, or rates of return—vary across the income distribution. 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

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)

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? Unlike an optimal mechanism, a smooth tax system cannot disallow *double deviations*, where individuals can jointly alter their earnings and savings to reach bundles not chosen by any other type. The broader forms of across-income heterogeneity we consider thus introduce a complication not present in the standard income taxation model of Mirrlees (1971), nor in Atkinson and Stiglitz (1976), as double deviations are generally not attractive to individuals in these settings. Nevertheless, we show that under an extended Spence-Mirrlees condition and under modest regularity conditions, it is possible to construct a smooth tax system 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 across-income heterogeneity justifying taxes on s : 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, $s'_{het}(z) := s'(z) - s'_{inc}(z)$, is a sufficient statistic for (local) across-income 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 all else constant and (ii) the degree to which higher-ability types prefer, or are able to obtain, more s , holding earnings constant. The second component is captured by s'_{het} , 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 earnings ability and unobserved factors, like preferences.

The formula for optimal savings tax rates is a product of s'_{het} 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 $s'_{het}(z) = 0$ 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. These Pareto efficiency conditions inform whether there is a set of joint perturbations to the income and savings tax that decrease distortions from taxation without decreasing revenue, such that all individuals

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.

are at least weakly better off.

We show that these formulas apply in a variety of 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, deviations from weak separability, and the ability to engage in income shifting. 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.

Our second contribution is a characterization of what we call “simple tax systems.” Across a large number of countries, the tax system consists of a nonlinear tax on income, 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. We also provide sufficient conditions for the SN and LED systems to implement the optimal mechanism. Importantly, when focusing on simple tax systems, we extend our results to multidimensional heterogeneity and to a potentially suboptimal income tax. In these cases, 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 generalize our results to several other key applications. 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.

Second, 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).

Third, 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 our 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 $s'(z) - s'_{inc}(z) = s'_{het}(z)$, 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 $s'_{het}(z)$ 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 $s'_{het}(z) = s'(z)$. They consider multi-dimensional 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 optimal taxation in settings with other sources of across-income heterogeneity. 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, as well as 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

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 Henriët (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 Henriët (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.

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.

in relaxing the assumption of homogeneous and weakly separable preferences, but using a static (two-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 5 studies simple tax systems. Section 6 presents extensions to our results. Section 7 applies our formulas to quantify optimal savings tax rates in the United States. Section 8 concludes. All proofs are gathered in the Appendix.

2 Baseline Model and Assumptions

For concreteness, we begin by discussing preference heterogeneity, and then generalize the analysis to broader forms of across-income heterogeneity in Section 4.

Setting. There is a population of heterogeneous individuals indexed by their type $\theta \in \Theta$, where Θ is compact. They differ in their disutility from generating earnings z and in their preferences for a consumption bundle (c, s) embodied in their utility function $U(c, s, z; \theta)$. We begin with the common assumption that heterogeneity is unidimensional, $\Theta \subset \mathbb{R}$, in which case we interpret type θ as reflecting earnings ability; Section 5.2 considers multi-dimensional heterogeneity. We assume that θ has a continuously differentiable cumulative distribution function $F(\theta)$.

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 and s is another dimension of commodity consumption that could be taxed nonlinearly, such as energy or housing (Gaubert et al., 2021).

Throughout the paper, we assume the following regularity conditions for the utility function:

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.

We also 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.

Preference heterogeneity. To introduce preference heterogeneity, consider the following example. In applications where s represents savings, a frequently used functional form (e.g., Saez, 2002; Golosov et al., 2013) involves additively separability and heterogeneity in individuals' 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 for s 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 for consumption bundles when the marginal rate of substitution between c and s varies with earnings ability. We denote this marginal rate of substitution by $\mathcal{S}(c, s, z; \theta) := \frac{U'_s(c, s, z; \theta)}{U'_c(c, s, z; \theta)}$ and similarly let $\mathcal{Z}(c, s, z; \theta) := \frac{U'_z(c, s, z; \theta)}{U'_c(c, s, z; \theta)}$ be the marginal rate of substitution between consumption c and earnings z . Using the shorthand $\mathcal{S}'_\theta(c, s, z; \theta_0) := \frac{\partial}{\partial \theta} \mathcal{S}(c, s, z; \theta)|_{\theta=\theta_0}$, we formally define preference heterogeneity as follows:

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

$$\exists \theta_0, \forall (c, s, z), \mathcal{S}'_\theta(c, s, z; \theta_0) \neq 0. \quad (2)$$

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

Government. The government seeks to maximize a weighted sum of utilities:

$$\max \int_{\theta} \alpha(\theta) U(c(\theta), s(\theta), z(\theta); \theta) dF(\theta), \quad (3)$$

where $\alpha(\theta)$ represents some set of Pareto weights across types. Selecting a particular set of weights requires normative assumptions, which we discuss when introducing social marginal welfare weights in Section 3.2.2.

Type θ is private information and cannot be observed by the government; only the distribution of types, $F(\theta)$, is known. Therefore the government designs a tax and transfer function \mathcal{T} that depends on the observable variables c , s , and z , which can be written without loss of generality as a tax on s and z only.⁷ The government anticipates that individuals choose these variables to maximize their utility subject to their individual budget constraints, $c + ps \leq z - \mathcal{T}(s, z)$. Thus, an optimal tax system maximizes (3) subject to individual optimization, and subject to a resource constraint, $\int_{\theta} \mathcal{T}(s(\theta), z(\theta)) dF(\theta) \geq E$, where E is an exogenous revenue requirement.

If the tax system $\mathcal{T}(s, z)$ is unrestricted, this problem is equivalent to the problem of selecting an optimal allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_{\theta}$ to maximize the objective in (3) subject to the resource constraint

$$\int_{\theta} [z(\theta) - ps(\theta) - c(\theta)] dF(\theta) \geq E, \quad (4)$$

and subject to incentive compatibility constraints

$$\forall (\theta, \theta') \in \Theta^2, U(c(\theta), s(\theta), z(\theta); \theta) \geq U(c(\theta'), s(\theta'), z(\theta'); \theta). \quad (5)$$

We refer to an allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_{\theta}$ that maximizes (3) subject to (4) and (5) as an *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 $\mathcal{T}(s, z)$. First, we show that an optimal incentive-compatible allocation can be implemented by a *smooth* tax system. Second, we derive a sufficient statistics characterization of optimal smooth tax systems.

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: $\tilde{\mathcal{T}}(s, z) = \mathcal{T}(\mathcal{C}(s, z), s, z)$.

3.1 Implementability with Smooth Tax Systems

Our implementation result relies on regularity conditions and on an extended Spence-Mirrlees condition about the optimal incentive-compatible allocation.

Assumption 2. *In the optimal incentive-compatible allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$, we assume that: (i) c , s , and z are smooth functions of θ , (ii) 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 $\theta' \neq \theta$, and (iii) the set of types for which $z'(\theta) = 0$ is of measure zero, and, when $z'(\theta) \neq 0$, the following extended Spence-Mirrlees condition holds for any type θ' :*

$$\mathcal{S}'_\theta(c(\theta), s(\theta), z(\theta); \theta') \frac{s'(\theta)}{z'(\theta)} + \mathcal{Z}'_\theta(c(\theta), s(\theta), z(\theta); \theta') > 0. \quad (6)$$

Assumptions (i) and (ii) are standard assumptions required to apply optimal control methods to characterize the optimal allocation. The main component of Assumption 2 is an extension of the Spence-Mirrlees condition that $\mathcal{Z}'_\theta(c, s, z; \theta) > 0$. Intuitively, the standard Spence-Mirrlees condition states that higher types are more willing to trade labor for consumption. Our condition extends this to our more general setting, stating higher types are more willing to trade labor for the joint change in c and s that can be obtained along the allocation path. Paralleling Mirrlees (1971), our conditions ensure that $z(\theta)$ is strictly increasing in θ (Appendix Lemma B.1), allowing 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.

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

1. \mathcal{T} satisfies individual 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 $U(c, s, z; \theta)$ subject to the constraint $c + ps + \mathcal{T}(s, z) \leq z$, for all $\theta \in \Theta$.

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

Theorem 1. *Under Assumptions 1 and 2, an optimal incentive-compatible allocation is implementable by a smooth tax system. In this smooth tax system, individuals' 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 individuals to locally adjust s and z 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.

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:

$$\mathcal{T}'_s(s(\theta), z(\theta)) = \mathcal{S}(c(\theta), s(\theta), z(\theta); \theta) - p \quad (7)$$

$$\mathcal{T}'_z(s(\theta), z(\theta)) = \mathcal{Z}(c(\theta), s(\theta), z(\theta); \theta) + 1. \quad (8)$$

In the presence of preference heterogeneity, individuals' temptation to deviate from their assigned allocation $(c(\theta), s(\theta), z(\theta))$ are stronger under a smooth tax system than under a direct 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 a direct 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 preference heterogeneity that we consider here.⁸ 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 C.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 $(c(\theta), s(\theta), z(\theta))$. 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 Appendix Lemmas C.1 and C.2, which char-

⁸As pointed out by Kocherlakota (2005), Werning (2010), and others, smooth tax systems can also generate double deviations in dynamic settings where there is a discrete set of types. Werning (2010) provides a general implementation proof for a dynamic setting where productivity is smoothly distributed. The setting studied by Werning (2010), and the proof technique, is distinct from ours because time preferences, and thus preferences for period-2 consumption, are assumed homogeneous.

acterize individuals' budget constraints and second derivatives of indirect utility functions for any tax system \mathcal{T} that preserves individuals' first-order conditions.

In the third step, we show that there exists a sufficiently large k such that for *all* types θ , $(c(\theta), s(\theta), z(\theta))$ is a *global* optimum 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 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 individuals' 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

Assumptions. To define our sufficient statistics, it is helpful to write individuals' optimization problem under a tax system $\mathcal{T}(s, z)$ as

$$\max_z \left\{ \max_{c,s} U(c, s, z; \theta) \text{ s.t. } c \leq z - ps - \mathcal{T}(s, z) \right\}, \quad (9)$$

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 .

When Assumptions 1 and 2 hold, our implementation result for smooth tax systems $\mathcal{T}(s, z)$ holds (Theorem 1) and we do not need to impose any other requirements. In cases where these assumptions fail, or when we study simpler tax systems for which this implementation result may not hold, our optimal tax formulas remain valid under the following assumption:

Assumption 3. *The tax systems under consideration are such that at the optimum: (i) these tax systems are smooth, (ii) $z(\theta)$ is smooth and strictly increasing, $c(z; \theta)$ and $s(z; \theta)$ are smooth functions of z and θ , and (iii) individuals' optima are unique and their first-order and second-order conditions strictly hold.*

Elasticities for z -choices. 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

$$\zeta_z^c(z(\theta)) := -\frac{1 - \mathcal{T}'_z(s(\theta), z(\theta))}{z(\theta)} \frac{\partial z(\theta)}{\partial \mathcal{T}'_z(s(\theta), z(\theta))} \quad (10)$$

$$\eta_z(z(\theta)) := -(1 - \mathcal{T}'_z(s(\theta), z(\theta))) \frac{\partial z(\theta)}{\partial \mathcal{T}(s(\theta), z(\theta))} \quad (11)$$

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 earnings choices take 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.⁹

Elasticities for s -choices. 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:

$$\zeta_{s|z}^c(z(\theta)) := -\frac{1 + \mathcal{T}'_s(s(z; \theta), z)}{s(z; \theta)} \frac{\partial s(z; \theta)}{\partial \mathcal{T}'_s(s(z; \theta), z)} \Big|_{z=z(\theta)} \quad (12)$$

$$\eta_{s|z}(z(\theta)) := -(1 + \mathcal{T}'_s(s(z; \theta), z)) \frac{\partial s(z; \theta)}{\partial \mathcal{T}(s(z; \theta), z)} \Big|_{z=z(\theta)} \quad (13)$$

$$s'_{inc}(z(\theta)) := \frac{\partial s(z; \theta)}{\partial z} \Big|_{z=z(\theta)} \quad (14)$$

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 accounting for nonlinearities in the tax system, holding labor income fixed at $z(\theta)$.

Note that we define the elasticity of s with respect to $1 + \mathcal{T}'_s$, rather than with respect to $p + \mathcal{T}'_s$. This choice is irrelevant when $p \equiv 1$ —a normalization that can be adopted without

⁹This corresponds to the type of circular adjustments described in, e.g., Jacquet and Lehmann (2021b).

loss of generality, as shown in our discussion of generalized budget constraints in section 4.1 below. However, defining the elasticity with respect to $p + \mathcal{T}'_s$ may be more natural in applications where s is a commodity sold at an after-tax price of $q = p + \mathcal{T}'_s$.¹⁰ For all elasticity concepts, we use the “bar” notation, as in $\overline{\zeta_{s|z}^c}$, to denote a population elasticity.

Preference heterogeneity. To quantify preference heterogeneity, we decompose the cross-sectional profile of $s(z) := s(z; \vartheta(z))$. 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 vary across earnings z . We thus define our measure of local across-income preference heterogeneity, $s'_{het}(z)$, as the difference between the cross-sectional change in s along the earnings distribution and the causal income effect $s'_{inc}(z)$:

$$s'_{het}(z(\theta)) := s'(z(\theta)) - s'_{inc}(z(\theta)) \quad (15)$$

Formally, $s'(z)$ is a total derivative equal to the sum of two partial derivatives:

$$\underbrace{\frac{ds(z, \vartheta(z))}{dz}}_{s'(z)} = \underbrace{\frac{\partial s(z'; \vartheta(z))}{\partial z'} \Big|_{z'=z}}_{s'_{inc}(z)} + \underbrace{\frac{\partial s(z; \vartheta(z'))}{\partial z'} \Big|_{z'=z}}_{s'_{het}(z)} \quad (16)$$

The first term on the right-hand side measures how a change in z affects s consumption, holding the type θ constant. The second term, $s'_{het}(z)$, measures how a change in type affects s consumption, holding earnings z constant. For instance, in example (1) above, $s'_{het}(z)$ would be directly proportional to the local change in the discount factor $\delta(\vartheta(z))$.

Violations of weak separability. The sufficient statistic s'_{het} also captures failures of weak separability between labor and consumption as in, e.g., Corlett and Hague (1953). For example, suppose that higher types get a higher hourly wage rate, and that consumption of s and leisure are complements. Then, because higher types θ obtain more leisure at a fixed level of earnings z , higher types will have a stronger preference for s holding z constant. In our setup, this amounts to $s'_{het} > 0$.

¹⁰It is straightforward to convert our results between these elasticity definitions: in this case, the appropriate formulas can be obtained from Theorem 2 and Proposition 3 by multiplying $\zeta_{s|z}^c$ by $(p + \mathcal{T}'_s)/(1 + \mathcal{T}'_s)$. The only change in Theorem 2 is that the left-hand-side in equation (20) becomes $\frac{\mathcal{T}'_s(s(z), z)}{p + \mathcal{T}'_s(s(z), z)}$, and analogously for Proposition 3.

3.2.2 Social Marginal Welfare Weights

To encode the policymaker’s redistributive tastes, we follow the literature in defining social marginal welfare weights as the marginal social welfare derived from an increase in consumption for an individual of type θ at points $s(\theta)$ and $z(\theta)$, normalized by the marginal value of public funds λ :

$$g(s(\theta), z(\theta)) := \frac{\alpha(\theta)}{\lambda} U'_c(z(\theta) - \mathcal{T}(s(\theta), z(\theta)) - ps(\theta), s(\theta), z(\theta); \theta). \quad (17)$$

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

$$\hat{g}(s, z) := g(s, z) + \mathcal{T}'_z(s, z) \frac{\eta_z(z)}{1 - \mathcal{T}'_z(s, z)} + \mathcal{T}'_s(s, z) \left(\frac{\eta_{s|z}(z)}{1 + \mathcal{T}'_s(s, z)} + s'_{inc}(z) \frac{\eta_z(z)}{1 - \mathcal{T}'_z(s, z)} \right), \quad (18)$$

where the last term comes from the fact that income effects on earnings, proportional to $\eta_z(z)$, induce changes in s consumption proportional to $s'_{inc}(z)$, affecting 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 Appendix C of Saez and Stantcheva, 2016) 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. For example, a savings tax might be viewed as an unfair tax on those with relatively high discount factors. Lockwood and Weinzierl (2015) show that this difficulty arises even in the standard Mirrlees (1971) model; because there is no formal distinction between heterogeneous earnings ability and heterogeneous preferences for exerting labor effort—both manifest as a revealed preference for a lower level of earnings.

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.¹¹ For results in the case of unidimensional heterogeneity, welfare weights are written as a function only of earnings, $g(z(\theta))$, 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

¹¹Our empirical application in Section 7 employs a version of the inverse optimum approach, estimating optimal savings taxes consistent with the current U.S. taxes on labor income.

function of both s and z , $g(s(\theta), z(\theta))$.

3.3 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 . This result is a generalization of Lemma 1 in Saez (2002) to 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 changes will be the same for 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 or under Assumption 3, 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:*

$$\frac{\mathcal{T}'_z(s, z)}{1 - \mathcal{T}'_z(s, z)} = \frac{1}{z \zeta_z^c(z)} \frac{1}{h_z(z)} \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) - s'_{inc}(z) \frac{\mathcal{T}'_s(s, z)}{1 - \mathcal{T}'_z(s, z)} \quad (19)$$

$$\frac{\mathcal{T}'_s(s, z)}{1 + \mathcal{T}'_s(s, z)} = s'_{het}(z) \frac{1}{s \zeta_{s|z}^c(z)} \frac{1}{h_z(z)} \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x). \quad (20)$$

Any Pareto-efficient smooth tax system satisfies

$$\frac{\mathcal{T}'_s(s, z)}{1 + \mathcal{T}'_s(s, z)} = s'_{het}(z) \frac{z \zeta_z^c(z)}{s \zeta_{s|z}^c(z)} \frac{\mathcal{T}'_z(s, z) + s'_{inc}(z) \mathcal{T}'_s(s, z)}{1 - \mathcal{T}'_z(s, z)}. \quad (21)$$

Formula (19) 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 tax revenue. To see this, consider a small increase in the earnings marginal tax rate, $d\tau_z$, in a small bandwidth of earnings, dz , around z . This reform triggers a lump-sum tax increase, $d\tau_z dz$, for all individuals with higher earnings that changes tax revenue and welfare by $d\tau_z dz \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x)$. Moreover, it induces the mass of individuals in the bandwidth, $h_z(z) dz$, to reduce their earnings and thus tax revenue by $\mathcal{T}'_z(s, z) \frac{z}{1 - \mathcal{T}'_z(s, z)} \zeta_z^c(z) d\tau_z$. This earnings reduction also leads to a reduction in the consumption of s by an amount proportional to s'_{inc} , which reduces tax revenue by $\mathcal{T}'_s(s, z) s'_{inc}(z) \frac{z}{1 - \mathcal{T}'_z(s, z)} \zeta_z^c(z) d\tau_z$. Summing these effects and characterizing the optimum as a situation where they cancel each other yields (19).

Formula (20) is one of our key results: Optimal marginal tax rates on s satisfy a condition that is remarkably similar to the standard “ABC” formula, and that provides a transparent generalization of the Atkinson-Stiglitz Theorem. When the sufficient statistic s'_{het} is equal to zero, the condition implies that the optimal tax on s must equal zero as well. When $s'_{het} > 0$, implying that higher earners have a stronger relative preference for s , the condition implies that the optimal tax rate on s must be positive and that its magnitude 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 $s(z)$.

To obtain intuition for the result, combine the earnings tax *increase* discussed above with a small *decrease* in the marginal tax rate on s , $d\tau_s$, in a small bandwidth ds around $s(z)$. Set $ds = s'(z) dz$ such that the two bandwidths coincide, $h_s(s(z)) ds = h_z(z) dz$.¹² By Lemma (1), the $d\tau_s$ decrease in the marginal tax rate on s induces individuals in this bandwidth to increase their earnings by $\frac{z}{1 - \mathcal{T}'_z(s, z)} \zeta_z^c(z) s'_{inc} d\tau_s$. Hence, setting $d\tau_z = s'_{inc}(z) d\tau_s$ implies that earnings changes induced by these two reforms cancel out, leaving the earnings of individuals in the bandwidth unchanged. As a result, the joint reform (i) induces individuals in the bandwidth to increase their consumption of s and thus tax revenue by $\mathcal{T}'_s(s, z) \frac{s(z)}{1 - \mathcal{T}'_s(s, z)} \zeta_{s|z}^c(z) d\tau_s$ and (ii) generates a lump-sum tax decrease, $dT = d\tau_s ds - d\tau_z dz = (s'(z) - s'_{inc}(z)) d\tau dz$, for all individuals with earnings higher than z , which induces a $dT \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x)$ change in tax revenue. Setting the sum of these effects to zero yields (20).

We can combine conditions (19) and (20) to derive the Pareto-efficiency condition in (21).¹³ Because the condition in (21) does not feature social marginal welfare weights,

¹²Here we use the fact that $h_s(s(z)) = \frac{h_z(z)}{s'(z)}$, which highlights that this heuristic proof strategy relies on the mapping $s(z)$ being strictly monotonic. Our formal appendix proof follows a more sophisticated proof strategy that does not require strict monotonicity of $s(z)$.

¹³This corresponds to combining tax reforms with $d\tau_z = s'(z) d\tau_s$ to cancel out lump-sum tax changes

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'_{het} of how relative tastes for s covary with earnings ability. The stronger the association between relative preferences for s and earnings ability, the more efficient it is to tax s instead of z . An important implication of this Pareto-efficiency condition is that in the absence of preference heterogeneity, positive tax rates on s are Pareto dominated. On the other hand, any Pareto-efficient tax system must feature non-zero tax rates on s in the presence of preference heterogeneity.

4 Across-Income Heterogeneity and its Determinants

4.1 Budget Heterogeneity and Auxiliary Choices

So far, we have considered economies with type-specific preferences $U(c, s, z; \theta)$ but type-independent budget constraints $c \leq B(s, z) - \mathcal{T}(s, z)$, with budget domain $B(s, z) := z - ps$. In this environment, across-income heterogeneity captured by the sufficient statistic $s'_{het}(z)$ originates from preference heterogeneity only.

However, this approach readily extends to other sources of across-income heterogeneity. For example, across-income heterogeneity in prices of s (Gahvari and Micheletto, 2016; Gerritsen et al., 2020), in income shifting (Slemrod, 1995; Christiansen and Tuomala, 2008), and 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 $s'_{het}(z)$ represents these alternative sources of heterogeneity—or a combination of them—and the characterization of optimal tax schedules remains intact.

To formalize this idea, consider the following modifications to our baseline model. First, assume that individuals might face type-dependent budget constraints given by $c \leq B(s, z, \chi; \theta) - \mathcal{T}(s, z)$, where χ represents a vector of auxiliary choices made by individuals that may affect their budget domain. Second, assume that individuals may manipulate their taxable earnings z and taxable savings s , or alter their ability to produce taxable earnings z by making productivity-enhancing investments, resulting in general type-dependent preferences $U(c, \phi_s(z, s, \chi; \theta), \phi_z(z, s, \chi; \theta), \chi; \theta)$ affected by auxiliary choices χ .

We establish the following equivalence result between economies:

above earnings z . This result builds on Konishi (1995), Laroque (2005), and Kaplow (2006), who derive Pareto-optimality conditions under the more restrictive assumptions of the Atkinson-Stiglitz theorem.

Lemma 2. *For any given tax system $\mathcal{T}(s, z)$, individuals make identical choices in (1) an economy in which they have type-dependent preferences $U(c, \phi_s(z, s, \chi; \theta), \phi_z(z, s, \chi; \theta), \chi; \theta)$ and type-dependent budget domains $B(s, z, \chi; \theta)$ that are potentially affected by auxiliary choices χ , and in (2) an economy in which individuals have type-independent budget domains $B(s, z)$, no auxiliary choices, and type-dependent preferences*

$$\tilde{U}(c, s, z; \theta) := \max_{\chi} U(c + B(s, z, \chi; \theta) - B(s, z), \phi_s(z, s, \chi; \theta), \phi_z(z, s, \chi; \theta), \chi; \theta). \quad (22)$$

Lemma 2 shows that an economy that features preference heterogeneity and budget heterogeneity as well as auxiliary choices is equivalent to an economy with preference heterogeneity only, provided that preferences are suitably defined. The intuition is that under a given tax system $\mathcal{T}(s, z)$, individuals' utility maximization problems are equivalent in both economies. This means that all individuals make identical choices and attain the same level of utility, and the government collects the same tax revenue.¹⁴ Since this equivalence holds for any tax system $\mathcal{T}(s, z)$, it immediately follows that these equivalent economies share the same optimal tax system, leading to the following proposition:

Proposition 1. *In an economy with preference heterogeneity, budget heterogeneity, and auxiliary choices, an optimal smooth tax system remains characterized by Theorem 2, as long as the equivalent economy with preference heterogeneity only satisfies the assumptions of the Theorem.*

This result highlights the generality of our sufficient statistics characterization of optimal taxes, where the key sufficient statistic s'_{het} captures all relevant dimensions of across-income heterogeneity that justify taxing s . While these different sources of across-income heterogeneity 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. To emphasize this point, we provide in Appendix B.3 a structural characterization of s'_{inc} and s'_{het} in economies with preference heterogeneity, budget heterogeneity, and auxiliary choices.

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

¹⁴This holds because the tax $\mathcal{T}(s, z)$ is measured in dollars and paid out of earnings. If instead the tax had a two-part structure where individuals pay T_1 in units of c (e.g., dollars) and T_2 in units of s (e.g., liters of soda), then budget heterogeneity would affect tax revenue. Such a system is relatively common for 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 individuals with higher returns more heavily in units of s . We explore such arbitrage motives in Section 6.3.

of whether it is due to real labor supply responses or costly avoidance behavior.¹⁵ We now discuss some examples of budget heterogeneity and auxiliary choices, focusing on those that are relevant in the context of savings taxation.

Heterogeneous Prices. One widespread form of budget heterogeneity is price heterogeneity. Suppose that individuals face prices $p(s, z, \chi; \theta)$ for s that depend on their level of s , earnings z , effort χ to seek out lower prices, or types θ , such that their budget domain is $B(s, z, \chi; \theta) = z - p(s, z, \chi; \theta)s$ and their preferences are $U(c, s, z, \chi; \theta)$, allowing for monetary or psychic costs of effort χ . For instance, in the context of savings taxation, higher savings might allow individuals to lock in better interest rates, higher income might generate beneficial network effects that expose individuals to better opportunities (both are examples of “scale dependence”), and higher ability types may be better at finding lower prices or higher returns on investments (an example of “type dependence”).

Lemma 2 shows that this economy is equivalent to an economy with budget domain $B(s, z) := z - s$, with the price normalized to $p \equiv 1$, where individuals’ utility function is $\tilde{U}(c, s, z; \theta) := \max_{\chi} U(c + (1 - p(s, z, \chi; \theta))s, s, z, \chi; \theta)$. This also demonstrates that in our baseline model, the price p can be normalized to 1 without loss of generality—a feature we employ in some appendix proofs. Intuitively, with a price of $p = p'$ instead of $p = 1$, individuals enjoy $(1 - p')s$ more consumption of c at a given choice s .

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 $s'_{inc}(z)$ and $s'_{het}(z)$ correctly distinguishes between them. In particular, *type*-dependent heterogeneity in prices will generally lead to $s'_{het}(z) \neq 0$ and thus to deviations from the Atkinson-Stiglitz result, whereas *scale*-dependent heterogeneity in prices would contribute to $s'_{inc}(z)$ and would thus preserve the Atkinson-Stiglitz result.

Heterogeneous Endowments. Another widespread form of budget heterogeneity is heterogeneity in initial endowments. Suppose that individuals have endowments $y_0(\theta)$, from inheritance or other sources, such that their budget domain is $B(s, z, \chi; \theta) = z + y_0(\theta) - s$. In the presence of income effects, endowments would influence individuals’ decisions, and *type*-dependent endowments would therefore enter s'_{het} . For instance, this could represent a situation in which higher productivity individuals have higher savings in part because they have larger inheritances, resulting in $s'_{het} > 0$ and justifying taxes on s .

¹⁵Chetty (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 6.2 and 6.3, respectively.

Income Shifting. Suppose that some auxiliary actions allow individuals to shift some of their labor income to capital income. 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 taxable savings (including capital income), and let $m(s, z, \chi; \theta)$ denote any financial costs involved in income shifting. Individuals' taxable labor income is thus $z = \tilde{z} - \chi$ and their taxable savings are $s = \tilde{s} + \chi$. This describes an economy with budget domain $B(s, z, \chi; \theta) = z - s + 2\chi - m(s, z, \chi; \theta)$ and preferences $U(c, s - \chi, z + \chi, \chi; \theta)$, where χ might directly influence utility because of effort or psychic costs. In this example, $\phi_s(s, \chi) = s - \chi$ and $\phi_z(z, \chi) = z + \chi$. If individuals with higher earnings ability are better able to engage in income shifting, this type dependence will effectively translate in $s'_{het} > 0$. In contrast, any scale dependence in income shifting would affect s'_{inc} , but not s'_{het} .

Human Capital Investments. Our framework can also be related to models studying human capital investments. Suppose that s represents human capital investments such as education, which affects productivity during working life (Bovenberg and Jacobs, 2005; Stantcheva, 2017). Following the static framework in Bovenberg and Jacobs (2005), the productivity-enhancing effect of human capital may be directly captured in preferences through, e.g., $\phi_z(z, s; \theta)$. Using Lemma 2, this economy is equivalent to one with preferences $\tilde{U}(c, s, z; \theta) := U(c, s, \phi_z(z, s; \theta); \theta)$. The functional form assumptions in Bovenberg and Jacobs (2005) imply that $s'_{het}(z) < 0$ 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.

4.2 Measuring s'_{het}

Because the key sufficient statistic $s'_{het}(z)$ is equal to the difference $s'(z) - s'_{inc}(z)$, it can be measured as a residual using familiar estimation strategies. The term $s'(z)$ represents the cross-sectional variation of s across the income distribution, which 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 types of quasi-experimental variation.

Proposition 2. *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:*

M1. If preferences are weakly separable and the tax system is separable, $s'_{inc}(z) = \frac{1-T'_z(z)}{1+T'_s(s(z))} \eta_{s|z}(z)$.

M2. If wage rates w and hours h are observable, $s'_{inc}(z) = s(z) \frac{\xi_w^s(z)}{w(z)+h(z)\xi_h^s(z)}$.

M3. 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 .¹⁶ 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.¹⁷

5 Simple Tax Systems and Multidimensional Heterogeneity

In practice, tax systems must be defined by policymakers and implemented by institutions, which may impose constraints on the degree of complexity in the tax function. In this section, we apply our sufficient statistics methods to characterize the optimal policy for a few

¹⁶A 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 above, 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.

¹⁷Fagereng 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 7.

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 in 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.

classes of tax systems with restrictions such as separability or linearity. Focusing on more restricted systems also allows us to characterize optimal policy 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. A common approach is thus to characterize the optimal policy within a restricted class of tax systems using conventional perturbation 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.

5.1 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 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 A1, provided as Supplementary Material, 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, we provide supplementary information.¹⁹

In the United States, for example, 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

¹⁸See Piketty and Saez (2013), Diamond and Spinnewijn (2011), and Gauthier and Henriët (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.

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. Taxes on long-term capital gains and qualified dividends take the form of linear earnings-dependent taxes.²⁰ 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.²¹

5.2 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 B.5. The preference heterogeneity statistic s'_{het} remains the key sufficient statistic for the marginal tax rate on s . For SL systems, it is convenient to introduce the term $s_{het}(z) := \int_{x=z_{min}}^z s'_{het}(x)dx$, which integrates local preference heterogeneity across incomes to obtain a measure of total preference heterogeneity up to earnings level z .

Proposition 3. *Suppose that the optimal SL, SN, and LED systems satisfy Assumption 3, and that 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:*

²⁰This is a slight approximation since the linear capital gains tax rate in the U.S. is a function of total income (including capital gains) rather than solely labor income.

²¹One 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, 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 B.7 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.

$$SL : \frac{\tau_s}{1 + \tau_s} = \frac{1}{\bar{s}\zeta_{s|z}^c} \int_z s'_{het}(z) \left[\int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) \right] dz \quad (23)$$

$$= -\frac{1}{\bar{s}\zeta_{s|z}^c} Cov[s_{het}(z), \hat{g}(z)] \quad (24)$$

$$SN : \frac{T'_s(s)}{1 + T'_s(s)} = \frac{1}{s\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} s'_{het}(z) \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) \quad (25)$$

$$LED : \frac{\tau_s(z)}{1 + \tau_s(z)} = \frac{1}{s\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} s'_{het}(z) \int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) \quad (26)$$

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:

$$SL : \frac{\tau_s}{1 + \tau_s} = \frac{1}{\bar{s}\zeta_{s|z}^c} \int_z s'_{het}(z) z\zeta_z^c(z) \frac{T'_z(z) + s'_{inc}(z)\tau_s}{1 - T'_z(z)} dH_z(z) \quad (27)$$

$$SN : \frac{T'_s(s)}{1 + T'_s(s)} = s'_{het}(z) \frac{z\zeta_z^c(z)}{s\zeta_{s|z}^c(z)} \frac{T'_z(z) + s'_{inc}(z)T'_s(s)}{1 - T'_z(z)} \quad (28)$$

$$LED : \frac{\tau_s(z)}{1 + \tau_s(z)} = s'_{het}(z) \frac{z\zeta_z^c(z)}{s\zeta_{s|z}^c(z)} \frac{T'_z(z) + \tau'_s(z)s + s'_{inc}(z)\tau_s(z)}{1 - T'_z(z) - \tau'_s(z)s} \quad (29)$$

The optimal tax formulas and the Pareto-efficiency conditions for SN and LED systems are analogous to the conditions for T'_s in the fully flexible smooth tax systems derived in Theorem 2. Appendix B.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 (23), 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 (25) and (26). 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 (27). Expression (24) 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 $s_{het}(z) = s(z)$; i.e., when there are no income effects and thus all consumption differences are driven by preference heterogeneity.²² In contrast, when all consumption differences are driven by heterogeneity in income such that $s_{het}(z) \equiv 0$, it 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 $s'_{het}(z) \neq 0$.

5.3 Multidimensional Heterogeneity, Suboptimal Earnings Taxes

Our next result generalizes Proposition 3 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) can be two-dimensional. Second, it characterizes optimal taxes on s even when the earnings tax $T_z(z)$ is not necessarily optimal.²³ Proposition B.6 in the Appendix characterizes optimal earnings tax schedules. The combination of Proposition B.6 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 4. *Consider smooth simple tax systems with (potentially suboptimal) earnings tax schedules $T_z(z)$ and optimally set marginal tax rates on s . Assume that individuals’ 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^0 , the marginal tax rates on s in SL/SN/LED systems must satisfy the following optimality conditions:*

²²In this special case of no income effects, the SN optimal tax formula (25) also nests a previous result: The optimal nonlinear tax formula of Saez and Stantcheva (2018), in which wealth choices originate entirely from preferences for wealth. Use $h_s(s(z)) = \frac{h_z(z)}{s'(z)}$ to recover the equivalence with formula (11) in Saez and Stantcheva (2018).

²³For 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 B.4.

$$\begin{aligned}
SL : \frac{\tau_s}{1 + \tau_s} \int_z \left\{ \mathbb{E} \left[s \zeta_{s|z}^c(s, z) \middle| z \right] \right\} dH_z(z) &= \int_z \left\{ \mathbb{E} \left[(1 - \hat{g}(s, z)) s \middle| z \right] \right. \\
&\quad \left. - \mathbb{E} \left[\frac{T'_z(z) + s'_{inc}(s, z) \tau_s}{1 - T'_z(z)} z \zeta_z^c(s, z) s'_{inc}(s, z) \middle| z \right] \right\} dH_z(z)
\end{aligned} \tag{30}$$

$$\begin{aligned}
SN : \frac{T'_s(s^0)}{1 + T'_s(s^0)} \int_z \left\{ s^0 \zeta_{s|z}^c(s^0, z) \right\} h(s^0, z) dz &= \int_z \left\{ \mathbb{E} \left[1 - \hat{g}(s, z) \middle| z, s \geq s^0 \right] \right\} dH_z(z) \\
&\quad - \int_z \left\{ \frac{T'_z(z) + s'_{inc}(s^0, z) T'_s(s^0)}{1 - T'_z(z)} z \zeta_z^c(s^0, z) s'_{inc}(s^0, z) \right\} h(s^0, z) dz
\end{aligned} \tag{31}$$

$$\begin{aligned}
LED : \mathbb{E} \left[\frac{T'_z(z) + \tau'_s(z) s + s'_{inc}(s, z) \tau_s(z)}{1 - T'_z(z) - \tau'_s(z) s} z \zeta_z^c(s, z) s \middle| z^0 \right] h_z(z^0) &= \int_{z \geq z^0} \mathbb{E} \left[(1 - \hat{g}(s, z)) s \middle| z \right] dH_z(z) \\
&\quad - \int_{z \geq z^0} \left\{ \mathbb{E} \left[\frac{\tau_s(z)}{1 + \tau_s(z)} s \zeta_{s|z}^c(s, z) \middle| z \right] + \mathbb{E} \left[\frac{T'_z(z) + \tau'_s(z) s + s'_{inc}(s, z) \tau_s(z)}{1 - T'_z(z) - \tau'_s(z) s} z \zeta_z^c(s, z) s'_{inc}(s, z) \middle| z \right] \right\} dH_z(z)
\end{aligned} \tag{32}$$

Because Lemma 1 still holds, the s'_{inc} statistic, together with standard elasticity concepts, allow us to characterize optimal taxes on s in terms of observables. If all terms inside the expectation operators $\mathbb{E}[\cdot|z]$ in Proposition 4 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 (30) 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

$$\frac{T'_z(z) \overline{s'_{inc}(z)} + \tau_s \overline{s'_{inc}(s, z)^2}}{1 - T'_z(z)} z \overline{\zeta_z^c(z)}. \tag{33}$$

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 . Formulas (31) and (32) also involve squares of s'_{inc} , implying that multidimensional heterogeneity can similarly lower the optimal tax rate on s through $(s'_{inc})^2$. We quantify the importance of this insight in our empirical application in Section 7. 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.

6 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'_{het} remains a key sufficient statistic for optimal taxes, and that our previous formulas readily extend to these settings.

6.1 Multiple Goods

Suppose individuals 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.²⁴

Proposition 5. *Under Assumption 3, for each taxed good $i = 1, \dots, n$ and at each bundle (c, \mathbf{s}, z) chosen by a type θ , an optimal smooth tax system satisfies*

$$\frac{\mathcal{T}'_{s_i}(\mathbf{s}, z)}{1 + \mathcal{T}'_{s_i}(\mathbf{s}, z)} = s'_{i,hets}(z) \frac{1}{s_i \zeta_{s_i|z}^c(z)} \frac{1}{h_z(z)} \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) + \underbrace{\sum_{j \neq i} \frac{\mathcal{T}'_{s_j}(\mathbf{s}, z)}{\mathcal{T}'_{s_i}(\mathbf{s}, z)} \frac{s_j \xi_{s_j,i|z}^c(z)}{s_i \zeta_{s_i|z}^c(z)}}_{\text{Tax diversion ratio}}. \quad (34)$$

²⁴The formula for $\hat{g}(z)$ in this more general setting is in Appendix C.10.

Any Pareto-efficient smooth tax system satisfies

$$\frac{\mathcal{T}'_{s_i}(\mathbf{s}, z)}{1 + \mathcal{T}'_{s_i}(\mathbf{s}, z)} = s'_{i,het}(z) \frac{z \zeta_z^c(z)}{s_i \zeta_{s_i|z}^c(z)} \frac{\mathcal{T}'_z(\mathbf{s}, z) + \sum_{j=1}^n s'_{j,inc}(z) \mathcal{T}'_{s_j}(\mathbf{s}, z)}{1 - \mathcal{T}'_z(\mathbf{s}, z)} + \underbrace{\sum_{j \neq i} \frac{\mathcal{T}'_{s_j}(\mathbf{s}, z)}{\mathcal{T}'_{s_i}(\mathbf{s}, z)} \frac{s_j \zeta_{s_j|i}^c(z)}{s_i \zeta_{s_i|z}^c(z)}}_{\text{Tax diversion ratio}}. \quad (35)$$

Proposition 5 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 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 .

6.2 Optimal Taxation with Corrective Motives

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 well-being 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

$$U(c, s, z; \theta) = u(c; \theta) - k(z; \theta) + \beta v(s; \theta), \quad (36)$$

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 well-being of their heirs. As in Farhi and Werning (2010), the government maximizes

$$\int_{\theta} [U(c(\theta), s(\theta), z(\theta); \theta) + \nu v(s(\theta); \theta)] dF(\theta), \quad (37)$$

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 $\nu = (1 - \beta)$.²⁵ 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 6. *Under Assumption 3, at each bundle (c, s, z) chosen by a type θ , an optimal smooth tax system satisfies the following marginal tax rate condition*

$$\frac{\mathcal{T}'_s(s, z)}{1 + \mathcal{T}'_s(s, z)} = s'_{het}(z) \frac{1}{s\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) - \frac{\nu(z)}{\beta(z)} g(z). \quad (38)$$

Any Pareto-efficient smooth tax system satisfies

$$\frac{\mathcal{T}'_s(s, z)}{1 + \mathcal{T}'_s(s, z)} = s'_{het}(z) \frac{z\zeta_z^c(z)}{s\zeta_{s|z}^c(z)} \left[\frac{\mathcal{T}'_z(s, z) + s'_{inc}(z)\mathcal{T}'_s(s, z)}{1 - \mathcal{T}'_z(s, z)} + s'_{inc}(z) \frac{\nu(z)}{\beta(z)} g(z) \right] - \frac{\nu(z)}{\beta(z)} g(z). \quad (39)$$

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 (38) 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 (38) 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) $s'_{het} \equiv 0$ and (ii) $\beta(z) \equiv \beta \in \mathbb{R}$, $\nu(z) \equiv \nu \in \mathbb{R}$. This core result of

²⁵See 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.

Farhi and Werning (2010) extends the standard Pigovian taxation logic to optimal screening of distortions with a nonlinear tax.

More generally, Proposition 6 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 tax. If $s'_{het}(z) > 0$ and $\nu(z)/\beta(z)$ and $g(z)$ are decreasing with z , Proposition 6 suggests a progressive tax on s that can feature subsidies at low incomes and taxes at high incomes.

6.3 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 . However, we show that when prices are heterogeneous, there is an additional efficiency rationale for differentially taxing s . Heterogeneity in prices motivates a form of “tax arbitrage,” where the government collects relatively more taxes in units of s from individuals who can obtain s at a low price or, in the case of savings, it imposes relatively higher savings taxes (and lower earnings taxes) on individuals with high rates of returns. This extension provides a generalization both of 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 $p(z, \theta)$

that 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 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,

$$\int_z \left\{ \alpha(z) U\left(z - T_1(z) - p(z)(s(z) + T_2(s(z), z)), s(z), z; \vartheta(z)\right) \right\} dH_z(z), \quad (40)$$

subject to the constraints

$$\int_z T_1(z) dH_z(z) \geq E_1 \quad \text{and} \quad \int_z T_2(s(z), z) dH_z(z) \geq E_2, \quad (41)$$

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, normalized by the marginal value of public funds λ_1 .²⁶

Heterogeneity in p generates efficiency effects through two channels, which are represented in Proposition 7 below. 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 individual'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 $\varphi(z) := -\left(T_1'(z) + \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial z} + s'_{inc}(z) \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s}\right) \frac{z\zeta_z^c(z)}{1-T_1'(z)} \frac{\partial p}{\partial z}$ 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 C.12.3.

Proposition 7. *Under Assumption 3 and under the assumption that for SN systems s is strictly monotonic (increasing or decreasing) in z , at each bundle (c, s, z) chosen by a type*

²⁶The formula for $\hat{g}(z)$ in this more general setting is in Appendix C.12.3.

θ , an optimal SN two-part tax system $\{T_1(z), T_2(s)\}$ satisfies

$$\frac{\lambda_2/\lambda_1 T_2'(s)}{1 + p(z)T_2'(s)} = \frac{1}{s\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} \left\{ s'_{het}(z) \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) + \frac{s'(z)}{p(z)} \left(\Psi(z) + \int_{x \geq z} [\varphi(x) - \bar{\varphi}] dH_z(x) \right) \right\} \quad (42)$$

where

$$\Psi(z) := \left(1 - H_z(z)\right) \int_{x \leq z} \hat{g}(x) (p(x) - p(z)) dH_z(x) + H_z(z) \int_{x \geq z} \hat{g}(x) (p(z) - p(x)) dH_z(x). \quad (43)$$

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

$$\begin{aligned} \frac{\lambda_2/\lambda_1 \tau_s(z)}{1 + p(z)\tau_s(z)} &= \frac{1}{s(z)\zeta_{s|z}^c(z)} \frac{1}{h_z(z)} \left\{ s'_{het}(z) \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) + \frac{p'(z)}{p(z)} s(z) \int_{x \geq z} [1 - \hat{g}(x)] dH_z(x) \right\} \\ &+ \frac{1}{\zeta_{s|z}^c(z)} \frac{1}{p(z)} \left\{ \bar{g}p - \bar{g}p(z) + \varphi(z) - \bar{\varphi} \right\}. \end{aligned} \quad (44)$$

Proposition 7 shows that the sufficient statistic $s'_{het}(z)$ continues to play a central role.²⁷ On the left-hand side of (42) and (44), the presence of $p(z)$ in the denominator is because an individual'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 7 also introduces novel efficiency terms that lead to taxes on s , even when $s'_{het} \equiv 0$. 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 de-

²⁷A characterization of the optimal earnings tax schedule $T_1(z)$ is in Appendix C.12.4.

clines cross-sectionally with z (i.e., $p'(z) < 0$). The first novel term in equation (44), 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 $\overline{\hat{g}p} - \hat{g}p(z) + \varphi(z) - \overline{\varphi}$, is also negative for sufficiently low values of z , as in this case both $\overline{\hat{g}p} - \hat{g}p(z)$ and $\varphi(z) - \overline{\varphi}$ are negative. However, this term is positive for sufficiently high values of z . Thus, when $s'_{het}(z) \equiv 0$, 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.

7 Empirical Application

We 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. In section 7.2, we devote particular attention to the calibration of the sufficient statistic $s'_{het}(z)$. We then use the Pareto-efficiency conditions derived in Proposition 3 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.

7.1 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 D 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 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.

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 D.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.²⁸ This provides us with a population of representative individuals at each percentile of the income distribution, for whom period 1 represents their working life, with a representative age of 45, and period 2 represents retirement, which occurs 20 years later at age 65. We normalize both labor earnings and retirement savings by the number of years worked.

²⁸The 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.

Figure A1 plots our estimate of gross (i.e., after-returns and before-tax) discretionary 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 the stock of 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 D.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 I.

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 measure elasticities from responses to tax reforms 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.²⁹ 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 1. Appendix D.1.4 discusses this conversion.

7.2 Estimation of s'_{het}

To estimate our measure of local preference heterogeneity $s'_{het}(z) = s'(z) - s'_{inc}(z)$, we employ two complementary strategies.

The first estimation strategy is motivated by the Proposition 2 result that when preferences are separable in s and z , the causal income effect of windfall income on s consumption identifies $s'_{inc}(z)$. To the extent that separability is plausible, we can exploit exogenous shocks to unearned income in order to estimate $s'_{inc}(z)$. To implement this strategy, we draw from 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.³⁰ Lottery consumption is widespread in Nor-

²⁹Our extension to many goods (Section 6.1) 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.

³⁰Two 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

way—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. Fagereng et al. (2021) find that individuals’ consumption peaks during the winning year and then gradually reverts to the pre-win baseline. 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 homogeneous 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—which introduces a small amount of heterogeneity across incomes—in order to arrive at $s'_{inc}(z)$. The resulting profile is plotted as the dashed red line in Figure I. The difference between this estimated schedule of $s'_{inc}(z)$ and the cross-sectional profile $s'(z)$, also plotted in Figure I, is $s'_{het}(z)$.

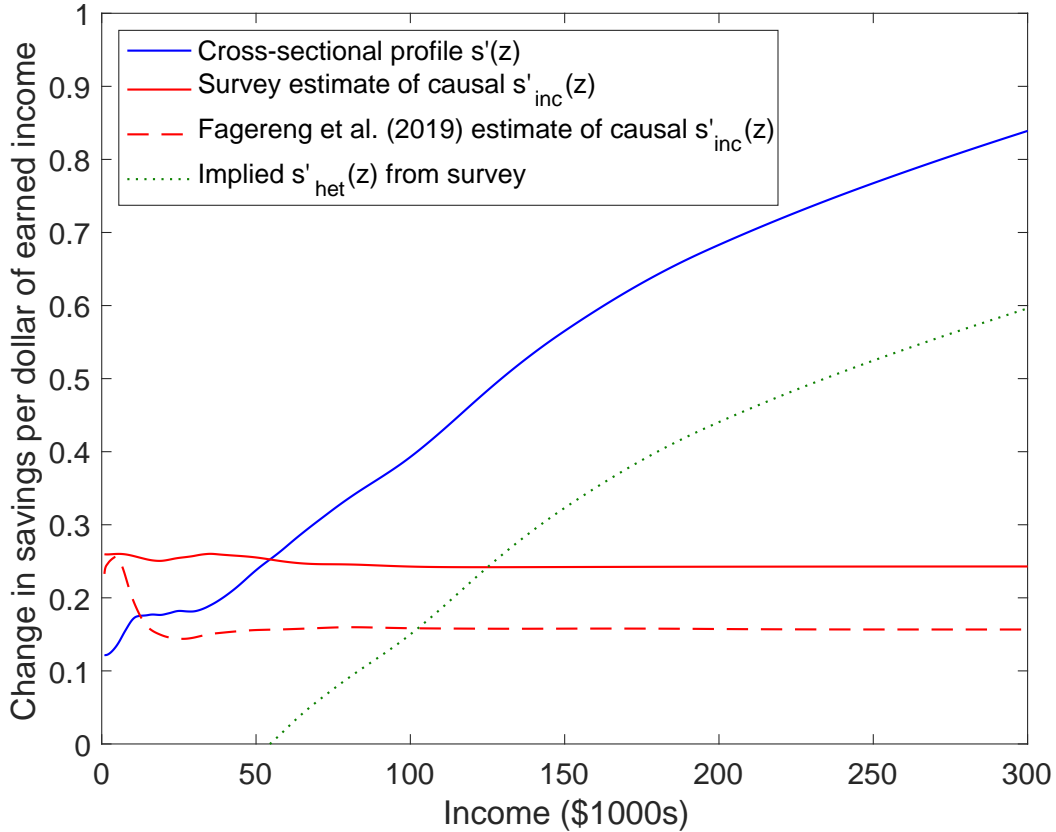
Our second strategy for estimating $s'_{het}(z)$ utilizes a new probability-based survey representing the U.S. adult population, conducted on the AmeriSpeak panel in the spring of 2021. We asked respondents to report how much more they would save each year if they received a hypothetical raise that increased their household’s annual income by \$1000 in the coming years. This strategy has the advantage of being based on the U.S. population, and of asking directly about a modest, *persistent* change in *pre-tax earned* income, rather than a large one-time windfall, so that it does not require a weak separability assumption. The survey results suggest a short-run marginal propensity to save of 0.60, 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), which we in turn convert to the across-income schedule of $s'_{inc}(z)$ displayed as the solid red line in Figure I. See Appendix D.1.3 for additional details about the survey results.

The two estimation strategies provide remarkably consistent estimates of $s'_{inc}(z)$, as shown in Figure I. There is a substantial difference between the cross-income profile $s'(z)$ and the causal income effect $s'_{inc}(z)$, suggesting that $s'_{het}(z)$ is positive across most of the income distribution and rises with income.³¹ We use the survey-based estimate as the baseline for our numerical results.

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.

³¹Our measure of $s'_{het}(z)$ 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.

Figure I: Decomposition of Cross-Sectional Savings Profile



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 7 and Appendix D.1 for details.

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 D.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.

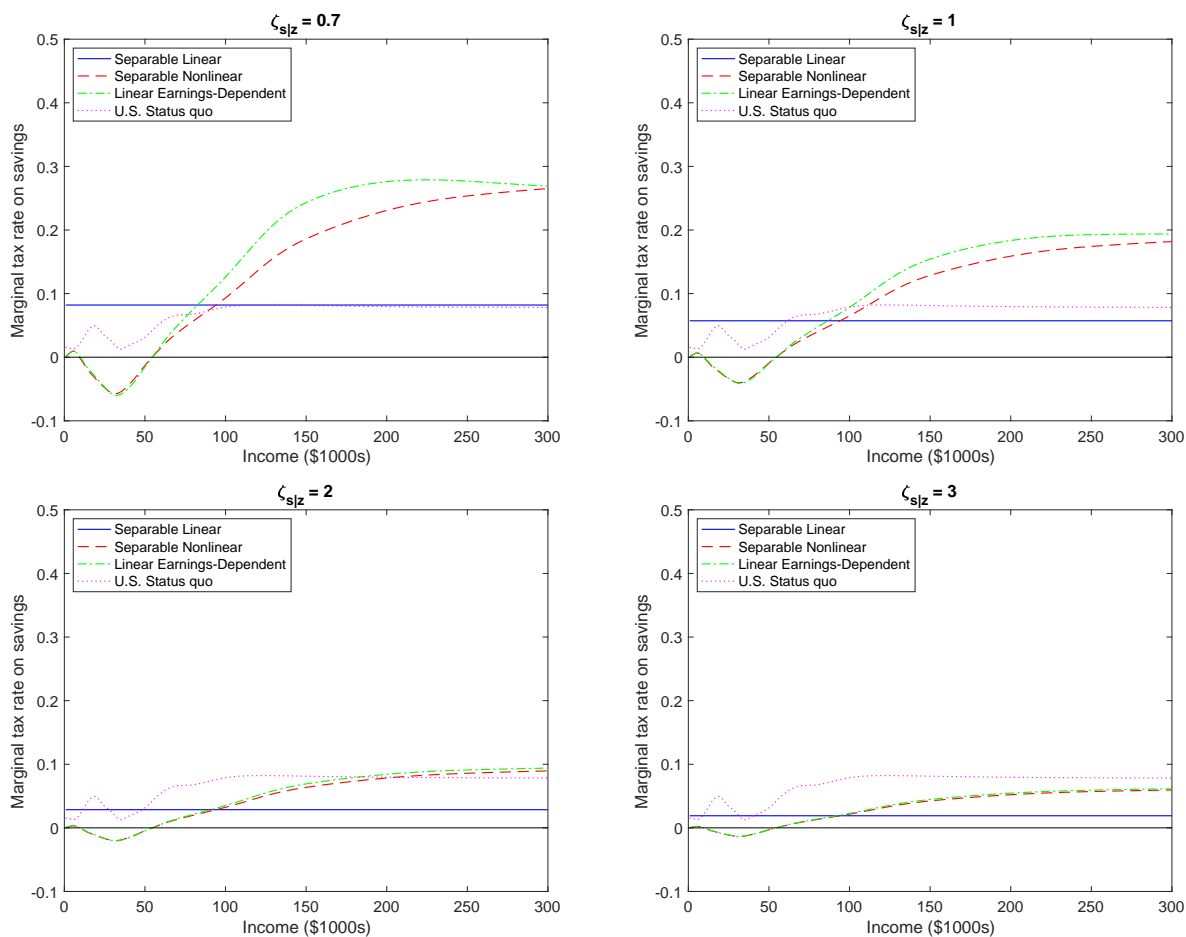
7.3 Results

Figure II reports the schedule of marginal tax rates for SL, SN and LED tax systems that satisfy the Pareto efficiency formulas in Proposition 3, 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 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 A1).

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 II.

Figure III considers two key extensions to these results: multidimensional heterogeneity, as in Section 5.3, and heterogeneous rates of return with “tax arbitrage” efficiency effects, as in Section 6.3. 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.

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

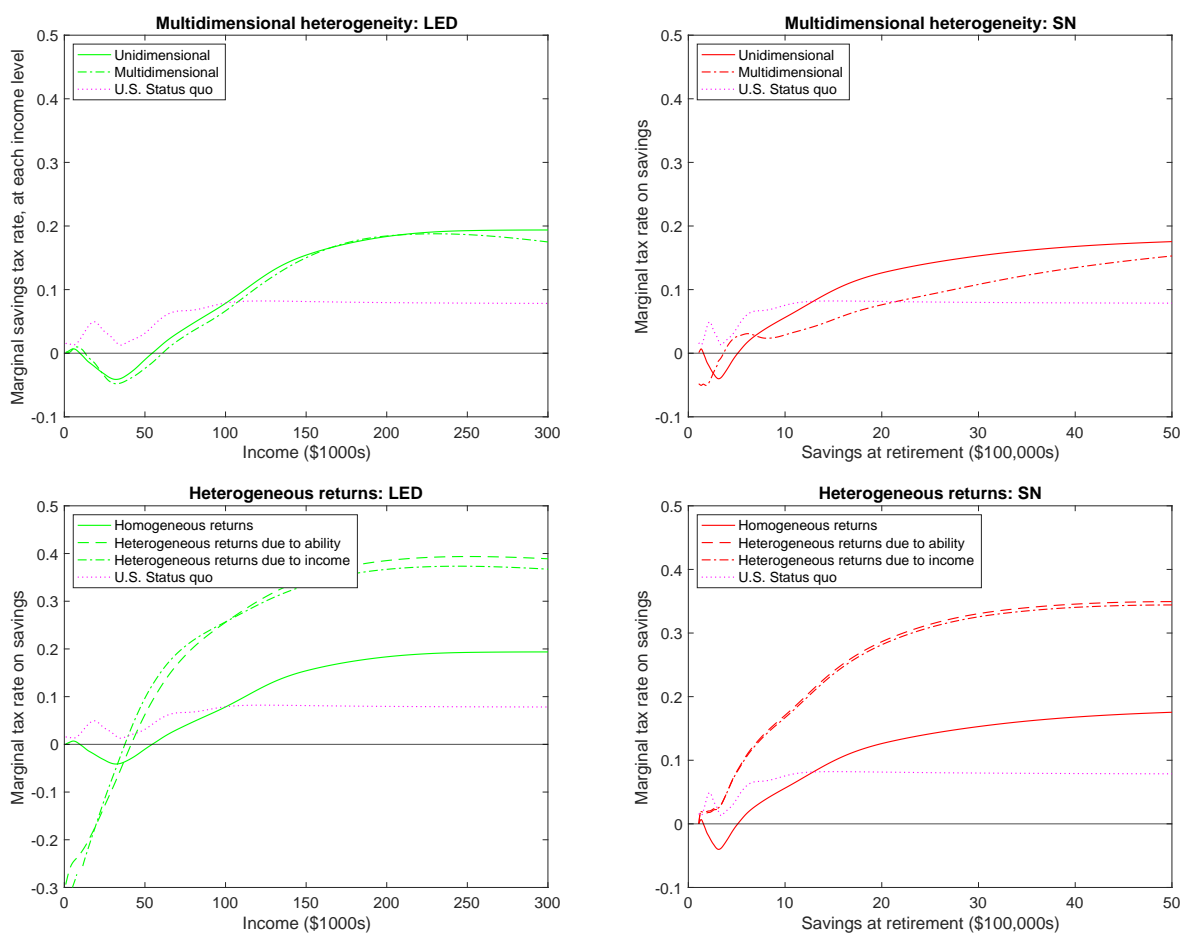


Notes: This figure presents the marginal savings tax rates values that satisfy the Pareto-efficiency formulas in Proposition 3, plotted against the earnings level to which they apply. We plot these schedules for four different values of the savings elasticity $\zeta_{s|z}$, with $\zeta_{s|z} = 1$ representing our baseline case.

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.

Figure III: 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 6: multidimensional heterogeneity (top row), and heterogeneous returns (bottom row). All plots also reproduce the Pareto-efficient savings schedules from Figure II 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 7 and Appendices D.2 and D.3 for details.

Consistent with the intuition described in Section 5.3, the top two panels of Figure III show that incorporating multidimensional heterogeneity reduces the magnitude of optimal tax rates in 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 individuals 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 7.³² The bottom right panel shows that tax rates in the SN system are higher at all levels of income, consistent with our discussion of the formula for SN systems in Proposition 7. 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 by the s'_{het} statistic. This highlights the importance of this 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.

8 Conclusion

This paper characterizes optimal tax systems on earnings and savings (or other dimensions of consumption) in the presence of general 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 that characterize the optimal smooth tax system through familiar empirical statistics, as well as a key sufficient statistic for across-income heterogeneity, $s'_{het}(z)$. This statistic is empirically estimable and captures heterogeneity in preferences, heterogeneous rates of return, endowments, or income-shifting abilities. We then use the same sufficient statistics to characterize a set of “simple” separable tax systems that are widely used in practice, considering 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.

³²Consistent with the tax arbitrage interpretation, these efficiency effects are (almost) unaffected by whether return heterogeneity is driven by income (scale dependence) or by ability (type dependence).

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 general across-income heterogeneity.

References

- AGERSNAP, O. AND O. M. ZIDAR (2021): “The Tax Elasticity of Capital Gains and Revenue-Maximizing Rates,” *American Economic Review: Insights*, 3, 399–416.
- ALLCOTT, H., B. B. LOCKWOOD, AND D. TAUBINSKY (2019): “Regressive Sin Taxes, with an Application to the Optimal Soda Tax,” *Quarterly Journal of Economics*, 134, 1557–1626.
- ATKINSON, A. B. AND J. E. STIGLITZ (1976): “The design of tax structure: direct versus indirect taxation,” *Journal of Public Economics*, 6, 55–75.
- BERNHEIM, B. D. AND D. TAUBINSKY (2018): “Behavioral Public Economics,” in *Handbook of Behavioral Economics*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, New York: Elsevier, vol. 1, 381–516.
- BIERBRAUER, F., P. BOYER, AND E. HANSEN (2020): “Pareto-improving tax reforms and the Earned Income Tax Credit,” *Working Paper no. 8358, CESifo*.
- BOADWAY, R., M. MARCHAND, AND P. PESTIEAU (2000): “Redistribution with Unobservable Bequests: A Case for Taxing Capital Income,” *The Scandinavian Journal of Economics*, 102, 253–267.
- BOURGUIGNON, F. AND A. SPADARO (2012): “Tax–benefit revealed social preferences,” *The Journal of Economic Inequality*, 10, 75–108.
- BOVENBERG, A. L. AND B. JACOBS (2005): “Redistribution and education subsidies are Siamese twins,” *Journal of Public Economics*, 89.
- (2011): “Optimal taxation of human capital and the earnings function,” *Journal of Public Economic Theory*, 13, 957–971.
- BRICKER, J., K. B. MOORE, AND J. P. THOMPSON (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, 31–52.
- (2012): “Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply,” *Econometrica*, 80, 969–1018.
- CHRISTIANSEN, V. AND M. TUOMALA (2008): “On taxing capital income with income shifting,” *International Tax and Public Finance*, 15, 527–545.
- CORLETT, W. J. AND D. C. HAGUE (1953): “Complementarity and the excess burden of taxation,” *The Review of Economic Studies*, 21, 21–30.

- CREMER, H., P. PESTIEAU, AND J.-C. ROCHET (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, 335–342.
- DIAMOND, P. AND J. SPINNEWIJN (2011): “Capital income taxes with heterogeneous discount rates,” *American Economic Journal: Economic Policy*, 3, 52–76.
- FAGERENG, A., L. GUIISO, D. MALACRINO, AND L. PISTAFERRI (2020): “Heterogeneity and persistence in returns to wealth,” *Econometrica*, 88, 115–170.
- FAGERENG, A., M. B. HOLM, AND G. J. NATVIK (2021): “MPC heterogeneity and household balance sheets,” *American Economic Journal: Macroeconomics*, 13, 1–54.
- FARHI, E. AND I. WERNING (2010): “Progressive Estate Taxation,” *Quarterly Journal of Economics*, 125, 635–673.
- (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, 674–680.
- GAHVARI, F. AND L. MICHELETTO (2016): “Capital income taxation and the Atkinson–Stiglitz theorem,” *Economics Letters*, 147, 86–89.
- GAUBERT, C., P. KLINE, AND D. YAGAN (2021): “Place-Based Redistribution,” *Working Paper no. 28337, National Bureau of Economic Research*.
- GAUTHIER, S. AND F. HENRIET (2018): “Commodity taxes and taste heterogeneity,” *European Economic Review*, 101, 284–296.
- GERRITSEN, A., B. JACOBS, A. V. RUSU, AND K. SPIRITUS (2020): “Optimal Taxation of Capital Income with Heterogeneous Rates of Return,” *Working Paper no. 8395, CESifo*.
- GOLOSOV, M., M. GRABER, M. MOGSTAD, AND D. NOVGORODSKY (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., N. KOCHERLAKOTA, AND A. TSYVINSKI (2003): “Optimal Indirect and Capital Taxation,” *Review of Economic Studies*, 70, 569–588.
- GOLOSOV, M., M. TROSHKIN, AND A. TSYVINSKI (2016): “Redistribution and Social Insurance,” *American Economic Review*, 106, 359–386.
- GOLOSOV, M., M. TROSHKIN, A. TSYVINSKI, AND M. WEINZIERL (2013): “Preference heterogeneity and optimal capital income taxation,” *Journal of Public Economics*, 97, 160–175.
- GOLOSOV, M. AND A. TSYVINSKI (2006): “New Dynamic Public Finance: a User’s Guide,” *NBER Macroeconomics Annual*, 317–379.
- (2015): “Policy Implications of Dynamic Public Finance,” *Annual Reviews of Economics*, 7, 147–171.
- HENDREN, N. (2020): “Measuring economic efficiency using inverse-optimum weights,” *Journal of Public Economics*, 187, 1–13.
- JACQUET, L. AND E. LEHMANN (2021a): “How to Tax Different Incomes?” *Discussion*

- Paper no. 14739, IZA Institute of Labor Economics.*
- (2021b): “Optimal income taxation with composition effects,” *Journal of the European Economic Association*, 19, 1299–1341.
- JAKOBSEN, K., K. JAKOBSEN, H. KLEVEN, AND G. ZUCMAN (2020): “Wealth taxation and wealth accumulation: Theory and evidence from Denmark,” *The Quarterly Journal of Economics*, 135, 329–388.
- KAPLOW, L. (2006): “On the undesirability of commodity taxation even when income taxation is not optimal,” *Journal of Public Economics*, 90, 1235–1250.
- KOCHERLAKOTA, N. R. (2005): “Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation,” *Econometrica*, 73, 1587–1621.
- KONISHI, H. (1995): “A Pareto-improving commodity tax reform under a smooth nonlinear income tax,” *Journal of Public Economics*, 56, 413–446.
- LAIBSON, D. (1997): “Golden Eggs and Hyperbolic Discounting,” *The Quarterly Journal of Economics*, 112, 443–478.
- LAROQUE, G. R. (2005): “Indirect taxation is superfluous under separability and taste homogeneity: A simple proof,” *Economics Letters*, 87, 141–144.
- LOCKWOOD, B. B. (2020): “Optimal income taxation with present bias,” *American Economic Journal: Economic Policy*, 12, 298–327.
- LOCKWOOD, B. B. AND M. WEINZIERL (2015): “De Gustibus non est Taxandum: Heterogeneity in preferences and optimal redistribution,” *Journal of Public Economics*, 124, 74–80.
- (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 I. SEGAL (2002): “Envelope Theorems for Arbitrary Choice sets,” *Econometrica*, 70, 583–601.
- MIRRLEES, J. A. (1971): “An Exploration in the Theory of Optimum Income Taxation,” *The Review of Economic Studies*, 38, 175–208.
- (1976): “Optimal tax theory: a synthesis,” *Journal of Public Economics*, 6, 327–358.
- MOSER, C. AND P. OLEA DE SOUZA E SILVA (2019): “Optimal paternalistic savings policies,” *Institute Working Paper 17, Federal Reserve Bank of Minneapolis*.
- PIKETTY, T. AND E. SAEZ (2013): “A theory of optimal inheritance taxation,” *Econometrica*, 81, 1851–1886.
- PIKETTY, T., E. SAEZ, AND G. ZUCMAN (2018): “Distributional national accounts: methods and estimates for the United States,” *The Quarterly Journal of Economics*, 133, 553–609.
- ROTHSCHILD, C. AND F. SCHEUER (2013): “Redistributive Taxation in the Roy Model,” *Quarterly Journal of Economics*, 128, 623–668.
- SACHS, D., A. TSYVINSKI, AND N. WERQUIN (2020): “Nonlinear tax incidence and optimal taxation in general equilibrium,” *Econometrica*, 88, 469–493.
- SAEZ, E. (2001): “Using elasticities to derive optimal income tax rates,” *The Review of Economic Studies*, 68, 205–229.

- (2002): “The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes,” *Journal of Public Economics*, 83, 217–230.
- SAEZ, E. AND S. STANTCHEVA (2016): “Generalized social marginal welfare weights for optimal tax theory,” *American Economic Review*, 106, 24–45.
- (2018): “A simpler theory of optimal capital taxation,” *Journal of Public Economics*, 162, 120–142.
- SAEZ, E. AND G. ZUCMAN (2020): “Trends in US Income and Wealth Inequality: Revising After the Revisionists,” *Working Paper no. 27921, National Bureau of Economic Research*.
- SCHEUER, F. AND A. WOLITZKY (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, 175–180.
- SMITH, M., O. M. ZIDAR, AND E. ZWICK (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, 1931–1990.
- (2020): “Dynamic taxation,” *Annual Review of Economics*, 12, 801–831.
- STIGLITZ, J. E. (2018): “Pareto efficient taxation and expenditures: pre-and redistribution,” *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 L. TO (2021): “Anticipation and Consumption,” *working paper*.
- WERNING, I. (2010): “Nonlinear Capital Taxation,” *working paper*.