NBER WORKING PAPER SERIES

OPTIMAL INCOME TAXATION

Louis Kaplow

Working Paper 30199
http://www.nber.org/papers/w30199

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2022

I am grateful to the editor, referees, Alan Auerbach, Ashley Craig, James Hines, Benjamin Lockwood, Casey Rothschild, Florian Scheuer, and Matthew Weinzierl for helpful comments and discussions; John Graham, Julia Huang, Xinchen Li, Leslie Liu, Nikolas Paladino, John Sullivan, and Michael Wei for research assistance; and Harvard University's John M. Olin Center for Law, Economics, and Business for financial support. Disclaimer: I consult on antitrust matters, and my spouse is a lawyer who mostly represents financial services firms. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Optimal Income Taxation
Louis Kaplow
NBER Working Paper No. 30199
July 2022
JEL No. A13,D61,D62,D63,D82,D83,H20,H21,H23,H24,H41,H43,H53,H55,J22,L40

ABSTRACT

This article explores subjects in optimal income taxation characterized by recent research interest, practical importance in light of concerns about inequality, potential for misunderstanding, and prospects for advancement. Throughout, the analysis highlights paths for further investigation. Areas of focus include multidimensional abilities and endogenous wages; asymmetric information and the income of founders; production and consumption externalities from labor effort; market power and rents; behavioral phenomena relating to perceptions of the income tax schedule, myopic labor supply, and the interactions of savings, savings policies, and labor supply; optimal income transfers; the relationship between optimal income taxation and the use of other instruments; and issues relating to the social welfare function and utility functions, including nonwelfarist objectives, welfare weights, heterogeneous preferences, and taxation of the family.

Louis Kaplow
Harvard University
Hauser 322
Cambridge, MA 02138
and NBER
meskridge@law.harvard.edu

1. *Introduction*

Mirrlees (1971) launched the field of optimal income taxation. Recent decades have seen a resurgence of interest in extending his original framework and adapting it to perform quantitative policy assessments. This renaissance has paralleled broader attention by economists and society to concerns about inequality and redistribution.

The subject of optimal income taxation is vast. It has given rise to textbooks and surveys that vary in breadth and focus. This article consciously highlights domains characterized by a combination of recent research interest, prospects for advancement, practical importance, potential for misunderstanding, and the author's comparative advantage. Among the significant omissions are empirical work, numerical methods, macroeconomics, political economy, federalism, international considerations, developing economies, most dynamic considerations, and the all-important subjects of administration, compliance, and enforcement. Nevertheless, this article covers a number of diverse and substantial subjects, striving throughout for depth over breadth and seeking to illuminate fruitful paths for further investigation.

Section 2 presents the core framework that constitutes the foundation for all that follows. Emphasis is placed on assumptions, qualifications, and intuition rather than on technical matters. Interpretation of the basic Mirrlees (1971) model is emphasized. In spite of its simple, static formulation, it can properly be understood only as a reduced form for a collapsed dynamic model wherein each individual's utility, income, consumption, and labor effort refer to experiences over a lifetime, including as a child, possibly as a parent and part of a married couple, and in retirement. Moreover, the nonlinear tax and transfer schedule being optimized actually represents much of the fiscal system, including not only all manner of taxes (including value-added taxes (VATs) and payroll taxes) but also cash and in-kind transfers and perhaps certain publicly provided private goods like health care and primary education. These often submerged features suggest important areas for theoretical exploration and call for substantial revision and extension of efforts to simulate policies.

Section 3 explores determinants of labor income that underlie the Mirrlees model. In the standard formulation, an individual's type, ability, marginal product of labor, and market wage are all taken to be equal to each other and exogenous, with effective labor supply in the economy constituting a fungible aggregate. The evolution of income inequality, however, motivates investigation of the determinants of individuals' skills and the market wages that result. Earlier explorations of the general equilibrium effects of income taxation on wages have been extended to explore a variety of impacts that may arise when abilities are multidimensional and occupational choice is endogenous. Attention is also directed at individuals who found, operate, and own significant portions of firms. They are increasingly responsible for a remarkable portion of income at the top of the distribution. Founders' labor effort is often entangled with their capital stakes, including from sweat equity, which creates complications attributable to asymmetric information between founders and external suppliers of capital—an important deviation from the canonical setting with perfect markets and a critical aspect of capital income taxation that has received little attention.

Section 4 elaborates extensions of the Mirrlees framework that address externalities and rents caused by or associated with labor effort. Optimal income taxation is directly implicated when the labor wedge itself has additional effects on social welfare or when an additional labor wedge may be present. Individuals' labor effort in certain occupations may generate positive or negative production externalities that affect other types' marginal products and hence their wages. In addition, greater labor effort that increases the actor's own consumption and utility may also raise or lower the utility of other individuals because the utility functions of the latter might depend on the circumstances of the former. Finally, market power and its associated rents interact with the optimal income taxation problem because market power has distributive effects, through both markups imposed on consumers and profits received by firms' owners, and it also influences the net return to labor effort by reducing wages' purchasing power and the derived demand for labor.

Section 5 explores the emerging subfield of behavioral optimal income taxation. One line of work considers systematic misperceptions of the income tax schedule, which have implications for optimal income taxation as well as for the interpretation of empirical work premised on individuals' accurate understandings of tax reforms. Another branch studies myopic labor supply. Finally, the substantial behavioral economics research on savings and associated corrective policies has potentially large implications for labor effort, in part arising from the very behavioral premises driving savings decisions. These labor supply effects in turn may have impacts on welfare that exceed those of improved savings and relatedly may substantially influence optimal income taxation. Many of these subjects have received only limited attention.

Section 6 turns to optimal income transfers, which in the Mirrlees framework can be understood as addressing the optimal intercept and the marginal income tax rates at the bottom of the income distribution. Transfers to the poor are highly consequential on account of their high marginal utilities of consumption and because some social welfare functions (SWFs) place greater weight on those with lower levels of utility. An important analytical tool is the introduction of separate income tax schedules—more broadly, tax schedules that depend on signals in addition to income—particularly because of the categorical nature of many transfer programs that aim at households with children, the disabled, or the elderly. Moreover, each schedule (just as when there is a single schedule) is taken to incorporate all tax and transfer programs. Simple but important lessons include that it is not meaningful to think in terms of an "optimal EITC" or of "phaseouts" (e.g., whether a universal basic income should be means-tested). Whatever is omitted under one program may be part of another; moreover, with separate schedules, only a single, cross-schedule revenue constraint applies. Work-inducing policies and the participation margin are also analyzed, generating insights that are applicable throughout the income distribution.

Section 7 elaborates a modular approach to the integration of optimal income taxation and other instruments that is useful for both theoretical and applied research. The framework has broad applicability, for example, to commodity taxation, corrective taxation, public goods provision, regulation, and estate and gift taxation. The method neither relies on functional form assumptions like the weak separability of labor in the utility function nor limits itself to explorations in the neighborhood the optimum for either the income tax or the other instruments. It addresses distributive and labor supply effects of the income tax and of the other instruments as well as distinctive policy targets such as externality correction. Of particular interest for present purposes, one of the two modules consists purely of Mirrleesian optimal income taxation analysis whereas the other module contains no such considerations but includes everything else. Accordingly, the former module can be analyzed entirely with the tools of optimal income taxation that are the subject of this article, and the latter module can be analyzed separately, without having to address the many challenges of optimal income taxation. Interestingly, this is true even though the latter module, through the use of a distributively offsetting income tax adjustment, includes a supplemental income tax schedule; that income tax adjustment is separated from the Mirrlees problem via a two-part decomposition. A major benefit of this approach is that it enables Pareto assessments of the other instruments because distribution is held constant within the second module.

Section 8 concludes the investigation by examining features of individuals' utility functions and the SWF. The focus of the former is on multidimensional heterogeneity that may influence labor effort and can have subtle implications for optimal redistribution. Possibly differing preferences regarding consumption-leisure choices can overlap with and in some settings be essentially indistinguishable from differences in ability that are at the center of the Mirrlees framework. Other forms of heterogeneity are considered as well, including the important case in which the "individuals" in the Mirrlees model are members of families or are taken as stand-ins for multimember households. The section also examines the frequent but not universal practice of employing an individualistic SWF—that is, one that depends directly and solely on individuals' utilities, a choice that is necessary to avoid conflicts with the Pareto principle—in contrast to nonwelfarist approaches that have received some attention. The relationship of both utility functions and SWFs to marginal social welfare weights is explored, with attention to how the endogeneity of those weights affects interpretations of the first-order condition for optimal marginal tax rates and of simulations that employ fixed weights when comparing policies or assessing the impact of parameter changes on optimal policy. Finally, section 9 offers brief closing remarks.

2. *Framework*

Mirrlees (1971) provides the modern framework for the study of optimal nonlinear income taxation. His article emphasizes to a remarkable degree the role of key assumptions and qualifications, the relaxation of which has provided much of the research agenda over the past half century. It is important to elaborate this framework's central elements as a foundation for the remainder of this article and a guidepost for additional areas of exploration.

Section 2.1 outlines the setup. Section 2.2 explores optimal linear income taxation because its simplicity clarifies central intuitions and surfaces important subtleties.[1] Section 2.3 presents the optimal nonlinear income tax.[2] Section 2.4 concludes with interpretative remarks that emphasize how the Mirrlees framework is a collapsed dynamic model that can be properly understood and applied only by adopting a lifetime perspective.

2.1. *Setup*

An individual's utility is $u(c, l)$, where c denotes consumption, l is labor effort, $u_c > 0$, and $u_l < 0$ (subscripts denote derivatives). An individual's consumption is given by

$$(1) \quad c = wl - T(wl),$$

where w is the individual's wage rate and T is the tax-transfer function.

The motivation for redistributive taxation is that individuals differ in their wages, also referred to as their earning abilities or types. The distribution of abilities is $F(w)$, with density $f(w)$. An individual's pre-tax income is $y = wl$. The variable l is taken to represent hours,

[1] For another two-parameter functional form, but with constant curvature, see Feldstein (1969), Bénabou (2002), and Heathcote, Storesletten, and Violante (2017).

[2] For texts and surveys, see Atkinson and Stiglitz (1980), Stiglitz (1987), Salanié (2011), Piketty and Saez (2013), and Tuomala (2016). See also the exploration of Pareto efficient tax schedules by Werning (2007).

intensity, and investments in human capital. The government perfectly and costlessly observes individuals' incomes, y, but is unable to observe ability, w, or labor effort, l (from which, if it could, would enable it to infer w).[3] As emphasized by Mirrlees (1971), this informational constraint, which renders individualized (type-specific) lump-sum taxation infeasible, lies at the heart of the second-best problem of optimal redistributive income taxation.

The tax and transfer schedule, $T(wl)$, at any income level may be positive or negative. Although much analysis focuses on $T'(wl)$, that is, marginal tax rates, the intercept, $T(0)$, is also of great interest. The value $-T(0)$ is the (uniform) lump-sum transfer received by all individuals who do not work and hence earn no income. Importantly, the function T represents the entire tax and transfer system. Regarding taxes, this includes not only income taxes but also payroll taxes and VATs (and, in a complete analysis, excise taxes, corporate income taxes, and more). Transfers include not only cash payments under the income tax and welfare programs but also social insurance payments, in-kind assistance (such as food stamps, housing assistance, and medical care), and perhaps more—such as government-provided child care and education, which raises subtleties beyond the scope of this article (Kaplow 2006b, 2008a). This breadth raises conceptual and practical challenges for empirical analysis and calibrated simulations, including many related to the fact that the model, although formally static, is taken to represent a collapsed lifetime perspective, as discussed in section 2.4.[4]

Individuals choose l to maximize $u(c, l)$ subject to their budget constraint (1). An individual's first-order condition is

$$(2) \quad w\big(1 - T'(wl)\big)u_c + u_l = 0.$$

Individuals' incentive constraints are often taken to be represented by this condition, which raises two problems. First, under various income tax schedules, including optimal ones, many individuals (those with the lowest w's) do not participate in the labor market; they (except for the marginal type in that group) are at a corner. Second, budget sets are not convex when there are fixed costs of labor force participation (discussed further in section 6.2 on optimal transfer programs) and when income tax schedules exhibit falling marginal tax rates. Both optimal schedules and many actual systems (due in large part to phaseouts of transfers) may well have falling rates toward the bottom, and some may have falling rates at the top. Nonconvexities lead to "jumpers," individuals who work discretely more or less in response to marginal changes in the tax schedule at higher or lower levels of income than that which they currently earn. (The first-order condition for an individual at a jumping margin holds at two different levels of l.) Ignoring this possibility in theoretical analysis or in simulations can produce erroneous results.[5]

[3] The perfect, costless observability of income is a large simplification that has significant consequences explored in a number of literatures (Andreoni, Erard, and Feinstein 1998; Cowell 1990; Roth, Scholz, and Witte 1989; Slemrod and Kopczuk 2002; Slemrod and Yitzhaki 2002; Slemrod and Gillitzer 2014; Slemrod 2019).

[4] As a simple example, a correct analysis includes both the payroll taxes that fund public retirement benefits and the benefits themselves. Ignoring both when the social insurance system as a whole is redistributive—or, worse, including the taxes but not the benefits—can create serious mismatches between what the model in principle represents and its application.

[5] For example, Slemrod et al. (1994) analyze the optimal two-bracket income tax, exposing an erroneous theoretical result in Sheshinski (1989) that was due to a failure to analyze jumpers; in their simulations, the optimal schedule had a lower rate in the higher bracket—that is, falling marginal tax rates—and hence a region "jumped over," with incomes that no individuals choose to earn. Regarding nonconvexities in nonlinear income tax schedules more broadly, see Mirrlees (1971) and Stiglitz (1987). As an interpretative aside, note that it would be incorrect to ignore

The government's problem is to choose a tax-transfer schedule $T(wl)$ to maximize social welfare, which can be stated as

$$(3) \quad \int W\Big(u\big(c(w), l(w)\big)\Big) f(w) dw,$$

where c and l are each expressed as functions of w to refer to the level of consumption achieved and labor effort chosen by an individual of type (ability) w.[6] This maximization is subject to a revenue constraint and to constraints regarding individuals' behavior. The former is

$$(4) \quad \int T\big(wl(w)\big) f(w) dw = R,$$

where R is an exogenously given revenue requirement. Here, revenue is interpreted as expenditures on public goods that should be understood as implicit in individuals' utility functions; because these expenditures are taken to be fixed, they need not be modeled explicitly.[7] A challenge in interpretation and application involves the previous observation that many government expenditures involve transfers—and hence are properly included as part of T—or are spent, for example, on publically provided private goods like child care and public education. The latter, depending on subtle assumptions, might likewise be appropriately included in T, particularly as part of the understanding of $T(0)$, which affects individuals' realized level of c, which in turn affects individuals' labor effort and the marginal social value of an additional dollar received by individuals of different types.

Regarding the incentive constraints, individuals are assumed to respond to the given tax schedule optimally as described by their first-order conditions. When these conditions hold, they can be differentiated with respect to a marginal adjustment of the income tax schedule to determine how labor effort will respond. Because individuals are at an optimum before this adjustment, their labor effort response has no first-order effect on their utility (the envelope theorem). Hence, the welfare implications of a tax adjustment will depend on its direct effect on utility—for example, paying a higher tax will reduce utility to an extent indicated by an individual's marginal utility of consumption—and on its revenue effects, which consist of two components. The direct ("mechanical") effect is the flipside of the effect on utility; a higher tax rate applied to existing income yields more revenue. The indirect ("behavioral") effect is due to the impact of individuals' adjustments of labor effort on revenue. Indeed, because of the envelope condition, individuals' behavioral responses are relevant only because of this revenue effect, which is often referred to as a "fiscal externality" because it is a social consequence ignored by individual actors. Note that, even when individuals discontinuously adjust their labor supply—at the participation margin or between two positive levels of labor effort—the envelope condition likewise applies, so again the only welfare-relevant impact of their behavioral response to a marginal tax change is the fiscal externality.

---

jumpers—whether at the participation margin or otherwise—on the ground that they have minimal mass because their behavioral changes are discrete, unlike those responding at the intensive margin.

[6] Section 8.1 discusses the use of welfare weights as well as nonwelfarist SWFs.

[7] Section 7.3 examines the optimal provision of public goods in the presence of a nonlinear income tax.

2.2. *Optimal Linear Income Tax*

A linear income tax is defined by

$$(5) \quad T(wl) = twl - g,$$

where t is the (constant, income-independent) marginal tax rate and g is the uniform per-capita grant, which is equal to $-T(0)$. Allowing $g \neq 0$—and, in particular, $g > 0$—is critical to what optimal tax theorists mean by a linear income tax. Note that if there were no incentive effects (and ignoring any exogenous revenue requirement), a linear income tax spans the full range of redistributive possibilities: with $t = 0$ and $g = 0$, there is no redistribution, and with $t = 1$ and $g = \bar{y}$ (mean income), there is full equalization.

A further implication is that, even allowing for incentive effects, the difference between linear and nonlinear income taxation is not that the latter entails greater redistribution but instead is more subtle. For example, if the optimal nonlinear income tax is U-shaped, as in Diamond (1998) and some other simulations, the optimal linear income tax would probably undertax both the rich and the poor and overtax the middle class. Furthermore, to the extent that much tax revenue is spent on transfers or publicly provided private goods, the overall distributive effects of a fiscal system may have more to do with the level of taxation than with whether the tax and transfer schedule is nonlinear or with the precise shape of that schedule. For example, the United States has a steeper income tax than that of some other rich countries that rely substantially on VATs and higher payroll taxes for much of their revenue, but the latter forms of taxation nevertheless finance a more generous and overall more redistributive welfare state.

Relatedly, the most straightforward role of differential commodity taxation in supplementing an income tax constrained to be linear (when the optimal nonlinear schedule is U-shaped) would be to tax both luxuries and necessities relatively highly and to tax goods consumed disproportionately by the middle class at relatively lower rates—prescriptions essentially unrelated to conventional Ramsey (1927) tax prescriptions. Nevertheless, a number of literatures—including an important strand of modern literature on optimal capital taxation—allow a linear income tax but implicitly assume (and sometimes explicitly state) that ruling out a nonlinear income tax means that $g = 0$. By contrast, work in the Mirrlees tradition, beginning with Atkinson and Stiglitz (1976), emphasizes the important role of allowing $g \neq 0$ and explains how most Ramsey-based results vanish or change qualitatively once the $g = 0$ requirement is relaxed (Stiglitz 1987, Mirrlees 1994, Kaplow 2008a).

Consider how to optimally set a linear income tax. The first-order condition for t (which implies the level of g) can usefully be expressed as

$$(6) \quad \frac{t}{1-t} = -\frac{cov\big(\alpha(w), y(w)\big)}{\int y(w)\varepsilon(w)f(w)dw},$$

where $y(w) = wl(w)$, income earned by individuals of ability w; $\varepsilon(w)$ is the compensated elasticity of labor effort of individuals of ability w; and $\alpha(w)$ is the net social marginal valuation of consumption, evaluated in dollars, of individuals of ability w.[8] The latter is given by

[8] There are many derivations of this condition, and it is expressed in a variety of ways. The presentation here is close to that in Stiglitz (1987, p. 1016, expression (29)), and his derivation appears in his note 31. See also Atkinson and Stiglitz (1980, pp. 407–408). For a more extensive analysis, see Stiglitz (1976). These derivations, it should be

$$(7)\quad \alpha(w) = \frac{W'(u(w))u_c(w)}{\lambda} + tw\left(\frac{\partial l(w)}{\partial g}\right).$$

The numerator of the first term on the right side of (7) indicates how much an additional dollar of (lump-sum) income to an individual of ability w contributes to social welfare: u_c is how much utility rises per dollar of consumption and W' is the extent to which social welfare increases per unit of utility, and this product is converted to a dollar value by dividing by λ, the shadow value of government revenue (which here corresponds to the value of raising g and thus is the average marginal social welfare weight over the population).[9] The second term, which is negative, reflects the income effect, namely that giving additional lump-sum income to an individual of ability w reduces labor effort, which in turn reduces government tax collections by tw per unit reduction in $l(w)$.

Expression (6) indicates how various factors affect the optimal level of a linear income tax. Beginning with the numerator, a greater (in magnitude) covariance between $\alpha(w)$ and $y(w)$ favors a higher tax rate. The net marginal social valuation of income, $\alpha(w)$, will be falling with income under assumptions ordinarily postulated (although the income effect can qualify this). In the present setting, a larger covariance does not refer to a closer (negative) correlation, which is always taken to be present, but rather to higher dispersions (standard deviations) of $\alpha(w)$ and $y(w)$. The dispersion of $\alpha(w)$ will tend to be greater the more concave (egalitarian) is the social welfare function W and the more concave is utility as a function of consumption (that is, the greater the rate at which marginal utility falls with consumption). The dispersion of $y(w)$ will be greater when (again, under standard assumptions) the distribution of underlying abilities is more unequal. In sum, more egalitarian social preferences, more concave utility as a function of consumption, and higher underlying inequality all favor a higher t.

The denominator on the right side of (6) indicates that a higher compensated labor supply elasticity favors a lower tax rate. The entire denominator is a weighted average; the elasticity matters more for high-income individuals (because more revenue is lost for a given percentage reduction in labor effort) and at ability levels where there are more individuals (typically the middle of the income distribution). If this compensated elasticity were constant, the denominator would equal that elasticity times average income, $\bar{y}$.

In focusing on expression (6)—and likewise for the first-order condition for the optimal nonlinear income tax, below—some major caveats are in order. First, income effects are relevant, here because they influence the value of $\alpha(w)$ through the second term in (7) (and also through the shadow price λ). Second, most of the values on the right side of (6)—including those entering via (7)—are endogenous. Thus, if one undertakes a comparison that postulates, say, a different labor supply elasticity—implicitly, a different utility function—essentially everything except $f(w)$ changes, including the shadow price of the government revenue

---

noted, typically do not take into account that some individuals (those of low ability) will choose not to work in many applications of the analysis.

[9] One can solve for λ using the first-order condition for g and substitute into expression (7) to yield the following more explicit statement that shows more fully the influence of income effects:

$$\alpha(w) = \frac{W'(u(w))u_c(w)}{\int W'\big(u(\omega)\big)u_c(\omega)f(\omega)d\omega}\left(1 - \int t\omega\left(\frac{\partial l(\omega)}{\partial g}\right)f(\omega)d\omega\right) + tw\left(\frac{\partial l(w)}{\partial g}\right).$$

constraint.[10] Accordingly, it is treacherous to make confident statements regarding the effect of changing one or another parameter without more elaborate analysis. This point will be relevant throughout this article; see, for example, the applications in section 5 on behavioral optimal income taxation. And nearly all of the effects and caveats developed here for the simple linear income tax model arise as well under nonlinear income taxation.

2.3. *Optimal Nonlinear Income Tax*

Returning to the more general formulation of the optimal income taxation problem described in section 2.1, the first-order condition is presented in a variety of ways in the literature.[11] Under a commonly used simplification with no income effects, it can be expressed as

$$(8) \qquad \frac{T'(wl(w))}{1-T'(wl(w))} = \frac{1-F(w)}{\xi(w)wf(w)} \frac{\int_w^\infty \left(1 - \frac{W'(u(\omega))u_c(\omega)}{\lambda}\right) f(\omega)d\omega}{1-F(w)},$$

where $\xi(w)$ is related to the elasticity of labor supply.[12] Note that because $l(w)$, the level of labor effort optimally chosen by an individual of type w, is endogenous, the optimal marginal tax rate T' at a given level of income will refer to the income of a type that depends on how the schedule is set. Moreover, the level of the grant, $-T(0)$, is implicit in the schedule of optimal marginal tax rates when the government's revenue constraint is met.[13]

This first-order condition is most easily understood by contemplating a local perturbation that raises the marginal income tax rate in a small interval in the neighborhood of some income level $y(w)$ (that corresponds to the earnings $wl(w)$ of some type w). This marginal rate increase will mechanically (inframarginally) raise a unit of revenue from all individuals who earn more than $y(w)$, which will be $1-F(w)$ of the population, the numerator in the first term on the right side of expression (8). (With income effects, there will be a further revenue increase from these inframarginal individuals.[14]) There also will be a behavioral (marginal) effect on individuals of the type w who earn $y(w)$, given by the denominator of the first term: $\xi(w)$ is the elasticity factor indicating how much $l(w)$ falls, w is the earnings reduction per unit decrease in $l(w)$, and $f(w)$ is the density of individuals thus affected.

Hence, the first term on the right side of expression (8) is a sort of benefit-cost ratio

[10] For example, if changing a parameter or relaxing an assumption reduces revenue, the result is analogous to a higher revenue requirement R (in either case, a given t now funds a lower g), which (in standard cases) tends to increase the variation in the marginal social valuation (7) and thus favor a higher t.

[11] See Saez (2001) (expressions (15)-(17) on page 215, evaluating expression (17) for the case in which the uncompensated and compensated elasticities are the same). See also Atkinson and Stiglitz (1980) (expression (13-54) on page 417, with different notation and arrangement of terms, including that their term corresponding to ξ appears in the numerator rather than in the denominator because it is defined as the reciprocal); Stiglitz (1987) (expression (25) on page 1007 and the expression in note 17 on page 1008); Diamond (1998) (expression (10) on page 86); Dahan and Strawczynski (2000) (expression (2) on page 682); and Auerbach and Hines (2002) (expressions (4.12) and (4.15) on pages 1381–1382).

[12] When utility takes a simple quasi-linear form, $\xi(w) = \varepsilon(w)/(1+\varepsilon(w))$.

[13] See also Scheuer and Werning (2018) on how one can establish a theoretical linkage between Mirrlees (1971) and Diamond and Mirrlees (1971) despite the fact that the former features a nonzero intercept whereas the latter, on the surface, does not allow one.

[14] Under a nonlinear income tax, the induced additional earnings themselves change individuals' marginal tax rates, which complicates the analysis (Saez 2001).

regarding the mechanical and behavioral effects on tax revenue of the marginal tax rate increase.[15] A larger mechanical effect favors higher marginal tax rates; $1 - F(w)$ is greater at lower incomes, helping to explain why many simulations have high, and falling, optimal marginal tax rates at the bottom. A larger behavioral effect favors lower marginal tax rates; w is greater at higher incomes, $f(w)$ is greatest in the middle of the income distribution, and $\xi(w)$ may vary with income in different ways depending on the utility function. (Regarding the latter, as a practical matter the elasticity of taxable income is relevant, and this elasticity is often thought to rise with income due to evasion and avoidance opportunities, although this is endogenous to tax design and administration and also depends on occupation, particularly regarding the importance of the cash economy at different income levels.[16])

The second term on the right side of expression (8) is a distributive weight. The integral from w to ∞ that is in turn divided by $1 - F(w)$ gives an average weight for individuals of types above w, reflecting that the redistribution is from them to the population as a whole. The expression in large parentheses in the integrand is the difference between the marginal dollar that is raised from each individual above type w and the dollar equivalent of the loss in welfare that occurs on account of that individual paying more tax. The numerator of $W'\big(u(\omega)\big)u_c(\omega)/\lambda$ indicates the marginal utility of a dollar of consumption to the type ω being integrated over, multiplied by the marginal contribution of that increment to social welfare. This product is divided by the shadow value of government funds, which can be thought of as the marginal social value of a dollar averaged over the population. The higher the type ω, the lower will be this factor as a whole and thus the greater will be the value one minus this expression. Accordingly, the higher is $y(w)$ and hence the type w whose marginal income tax rate we contemplate increasing, the greater will be the average of this distributive term, reflecting a greater social welfare gain associated with redistributing from this inframarginal group to the population as a whole. Hence, this second, distributive term is a force for rising marginal income tax rates.

As emphasized previously, one must be careful in interpreting such first-order conditions due to the endogeneity of many variables on the right side. Here, pertinent endogeneity includes the three elements of $W'\big(u(\omega)\big)u_c(\omega)/\lambda$. The marginal utility of consumption of any type ω depends on $-T(0)$ (which itself depends on the amount of revenue raised from the entire population) and on the schedule of marginal tax rates up to that type. The marginal contribution to social welfare depends on the realized utility of type ω, unless the SWF is utilitarian, in which case W' is constant. Perhaps less obviously, the shadow value of government revenue is itself a weighted average of endogenous values over the population (interpreting this shadow price, as with the optimal linear income tax, as the marginal social value of raising the grant). The significance of this endogeneity will be noted at many points in this article; corresponding warnings regarding the proper use of marginal social welfare weights are elaborated in section 8.1.

Starting with Mirrlees (1971), simulations have been used to explore the shape of the optimal nonlinear income tax. Although not the focus of this theoretical investigation, a few

---

[15] Recall that the behavioral effect on social welfare consists solely of the revenue effect (fiscal externality) because of individuals' envelope condition. Note further that the ratio $(1 - F(w))/f(w)$, which features in the discussion to follow, is the inverse of the hazard rate of the distribution of w---and that $(1 - F(w))/wf(w)$ is the inverse of the local Pareto parameter of this distribution, a property noted below in discussing the optimal asymptotic top marginal income tax rate.

[16] See the analyses and surveys in Feldstein (1999), Gruber and Saez (2002), Slemrod and Kopczuk (2002), Chetty (2009), Saez, Slemrod, and Giertz (2012), and Mertens and Montiel Olea (2018).

results will be noted. Departing from earlier work that used a lognormal distribution of abilities, Diamond (1998) examines a Pareto distribution, under which the $1 - F(w)/f(w)$ component of expression (8) rises more steeply at the upper end of the income distribution. He finds that optimal marginal tax rates are rising at the top. Dahan and Strawczynski's (2000) simulations indicate that Diamond's result was driven in part by his additional assumption that preferences were quasi-linear, which eliminates income effects. Nevertheless, their diagrams suggest, consistent with Diamond's claim, that moving from a lognormal to a Pareto distribution favors higher rates—still falling, but notably less rapidly—at the top of the income distribution. Saez (2001), using annual income distribution data in the United States from 1992 and 1993, finds in his simulation with a utilitarian welfare function, a compensated elasticity of labor supply of 0.5, and a functional form for utility that has income effects, that the optimal schedule has a marginal rate near 80% at the bottom of the income distribution that falls to approximately 40% at \$80,000, and then rises to nearly 70% at the upper end, where it roughly levels off.[17]

There has also been theoretical exploration of the optimal marginal income tax rates at the bottom and top of the income distribution. If the lowest type supplies positive labor, the optimal bottom marginal tax rate is zero: there is no redistributive gain because the second, distributive term in expression (8) is zero reflecting that the inframarginal population is the entire population (Brito and Oakland 1977, Seade 1977, Ebert 1992). And if there is a highest type, it can be shown that the optimal top marginal income tax rate is zero because there is no inframarginal revenue gain, leaving only the marginal distortion (Phelps 1973, Sadka 1976, Seade 1977). But neither of these results is regarded to have much practical relevance. Suppose instead that the ability distribution is unbounded at the top; then a simple approximation can be obtained for the optimal top marginal rate under a number of assumptions that have greater appeal (Diamond and Saez 2011). First, assume that the marginal social utility of a dollar falls to zero in the limit as income rises. In that event, the second term in expression (8) equals 1, and the overall expression greatly simplifies. If one additionally posits a constant (limiting) uncompensated elasticity e and that the distribution is approximately Pareto with parameter a at the top end, it can be shown that the limiting top marginal income tax rate approaches $1/(1 + ae)$.

2.4. *Interpretation*

In order to understand what can be learned, develop appropriate extensions, and perform policy simulations, models of optimal income taxation must be interpreted appropriately. This section offers some observations and caveats. Many concern the fact that the standard Mirrlees framework not only can be, but really must be interpreted as a collapsed dynamic model of individuals over their lifetimes. This problem is complex, among other reasons because individuals start their lives as children, typically living in households with other adults, and then progress with various orderings, durations, and probabilities through periods of formal education (involving the acquisition of human capital), life as a single adult, time as part of a married couple (which may include periods that are childless, with young children, and with older

[17] His functional form for utility has income effects that rise with income to an extent that the uncompensated elasticity approaches zero as w increases, which favors higher marginal rates at the top than otherwise. For further exploration of the optimal asymptotic marginal income tax rate, see Dahan and Strawczynski (2012).

children), and retirement.[18] Any individual's utility function thus is a stand-in for a sum or integral of these experiences over a lifetime.[19]

The first component of $u(c,l)$ is best understood as a reduced form for lifetime consumption, something not well captured by a snapshot of annual income at some point in an individual's adult life (Aaberge and Mogstad 2015, Scheuer and Slemrod 2020).[20] Analysis confined to working adults excludes children and retirees from assessments of behavior and of social welfare. Likewise, the second component of $u(c,l)$ includes not only hours worked and the intensity of work but also the development of human capital, whether as a child, a young adult pursuing higher education, or a worker who is learning by doing. Disentangling ability and effort in $y(w) = wl(w)$ is not straightforward, which also makes it more difficult to extract $f(w)$ from an observed distribution of earnings, even apart from the aforementioned complications regarding childhood and retirement.

Another challenge involves determining the functional form and parameters of $u(c,l)$, which itself no doubt depends on one's stage in the lifecycle. These choices regarding u, often made for reasons of tractability, are important not only to properly identify the elasticity of labor effort (or of taxable income), which has received a great deal of attention, but also because the functional form and parameters of u are directly relevant to assessments of social welfare, as elaborated in section 8. The lifetime perspective is highly relevant to both, a point already suggested by changes in an individual's family and work status over time.

Regarding labor effort, a central challenge—limiting attention now to (potentially) working adults—is in measuring long-run elasticities, which are the relevant parameters for determining the ultimate effects of reforms.[21] For reasons of data availability and identification, timeframes are often fairly short, generating the serious possibility that measured elasticities significantly understate long-run elasticities. Many individuals may not yet even be aware of how recent, subtle reforms affect their budget sets. More broadly, many margins of adjustment—ranging from investment in human capital to occupational and lifestyle choices—can take many years or even a generation to emerge.[22]

There is also a significant conundrum pointing in the reverse direction. Looking at developed countries over the past century or two, real wages are now an order of magnitude (or more) higher, the disutility of labor effort has fallen dramatically, and labor-saving substitutes for home production (from consumer durables to home heating to sliced bread) have expanded to a remarkable degree.[23] So why have we seen significant decreases in hours worked over this time frame rather than large increases? One answer is that income effects may be high

[18] A number of literatures have extended the optimal income tax framework to incorporate savings, uncertainty over future earnings, borrowing constraints, and other factors (Golosov, Tsyvinski, and Werning 2007; Stantcheva 2020). Many of the points emphasized here, which are qualitatively different, have received less attention.

[19] For a preliminary theoretical exploration, see Scheuer and Werning (2018), and for suggestive empirical evidence, see Altig et al. (2020). See also Kremer (2002), Weinzierl (2011), Bastani, Blomquist, and Micheletto (2013), Michau (2014), and Heathcote, Storesletten, and Violante (2020) on age-dependent income taxation.

[20] Interestingly, Vickrey (1939) proposed lifetime income averaging in light of graduated marginal income tax rates.

[21] For empirical explorations, see Chetty, Friedman, Olsen, and Pistaferri (2011), Keane (2011), Chetty (2012), Keane and Rogerson (2012), and Gelber, Jones, and Sacks (2020).

[22] A lifetime perspective is important for myriad issues. For example, some of the behavioral phenomena examined in section 5 imply that individuals may overspend or underspend their budgets, which seems less likely in the long run, suggesting either that the phenomena may not significantly alter optimal policy or that there exist other effects that need to be taken into account.

[23] One might add that adults in developed countries have many fewer children compared to centuries past, which greatly reduces the need to spend time at home rather than in the labor market. This, of course, helps to explain the increase in female labor force participation over the past half century.

(Restuccia and Vandenbroucke 2013; Bick, Fuchs-Schündeln, Lagakos, and Tsujiyama 2019; Boppart and Krusell 2020). Yet in research on optimal income taxation, these often are thought to be small and many analyses take them to be zero.[24] Another is the huge increase in leisure activities. Whatever is the explanation, one needs to employ and calibrate a reduced-form utility function that is consistent with such phenomena.

Because this is a theoretical exploration whereas many of the challenges suggested here are empirical, they will not be pursued further. The foregoing considerations outline a substantial research agenda, even before exploring extensions of the basic model that are the focus of this article. This perspective also calls for significant caution in drawing policy implications from optimal income tax simulations that, for practical reasons, are calibrated without regard to most of these considerations.

3. *Labor Income*

The canonical formulation of the optimal nonlinear income taxation problem addresses the taxation of labor income. A central set of assumptions to relax and questions to explore involves the concept of labor income itself. This section examines developments in this realm and identifies topics for further research.[25]

As background, section 3.1 elaborates individuals' types. In the basic Mirrlees' model, we have $y = wl$, with $f(w)$ as the density function. The variable w equivalently represents type, ability, the marginal product of labor, and the market wage, all taken to be equal to each other and exogenous, with effective labor supply in the economy being a fungible aggregate. But these need not be the same. Moreover, many proffered sources of increases in inequality entail changes in different types' wages over time, so it is important to consider how this evolution can be analyzed in the standard model or extensions thereof.

Section 3.2 addresses multidimensional abilities. Although multidimensional screening problems can be much more challenging to analyze, important work has considered tractable variations that enable the study of important forces. This work takes wages to be endogenous, examining how the distribution of different skills and endogenous labor effort determine the distribution of individuals' realized wages. Adjustments to the income tax schedule change labor supply differentially for different skill types and thus have welfare-relevant feedbacks on the wage distribution.

Section 3.3 relaxes the implicit assumption that other market participants can observe individuals' w's and l's (even though the government cannot). Analysis focuses on an application to the founders of firms. A practical motivation is that a significant portion of income at the top of the income distribution—and of recent increases therein—is attributable to those who founded, operate, and own significant fractions of their firms. These individuals' ownership involves capital (including sweat equity) and hence what might be viewed as capital income, but this is entangled with their labor income. These ownership stakes—which on pure

[24] In both modeling and applications, analysts often use a quasi-linear utility function of the form $u(c - v(l))$, which is more tractable because it has no income effects. It implies, for example, that if c has risen by an order of magnitude or more—and if, moreover, u is even moderately concave (say, taking the log form)—then the marginal disutility of labor effort, $u'v'$, has fallen massively, even setting aside any reduction in v', which greatly magnifies the puzzle while ruling out by assumption perhaps the major explanation. See Chetty (2006) for bounds on this effect. Another implication of this functional form is that higher labor effort, by raising $v(l)$, raises the marginal utility of consumption, despite leaving less time available to engage in consumption activities.

[25] Additional dimensions are explored elsewhere, notably an important literature on income taxation and human capital, surveyed in Stantcheva (2020).

diversification grounds should not exist (or should be negative)—are intimately related to founders' labor supply on account of information asymmetries between founders and external suppliers of capital.

3.1. *Ability, Marginal Product, and Wages*

Individuals' abilities determine their marginal products, which in turn determine their wages. In the standard model, each of these is identical, all labor effort is fungible (subject to a linear scaling by ability), and wages are fixed, taken to be equal to similarly fixed marginal products, with perfectly competitive labor markets.

Begin with individuals' abilities, which are taken to be unidimensional and homogeneous in the sense that production depends only on the total quantity of effective labor effort, which for each individual equals the product of labor supply, l, and a scaling factor, w. It is apparent, however, that individuals' abilities are multidimensional. Each type can thus better be characterized by a vector that represents a skillset. At a broad level, we might distinguish between cognitive, physical, and emotional abilities, but each of these (and more) has important subcomponents. Subtle differences—such as in coordination for professional athletes or interpersonal skills for managers—can translate to large differences in marginal products.

Furthermore, the marginal product associated with an individual's skill vector depends on many factors external to the individual. One is the matching of ability vectors (workers) to occupations: most leading scientists would be poor athletes, and vice versa. A given ability vector will yield different marginal products in different occupations, depending on how close is the match, the supply and demand for that skillset, and the technological frontier. Centuries ago and in many parts of the world today, brawn was more valuable than brains. Relative scarcity is quite important, which in turns depends on technology and preferences as well as policies, notably, concerning trade and infrastructure. There is no intrinsic mapping of ability vectors to marginal products. Not only the magnitudes but also the orderings are endogenous to much else in the economy.

Another key determinant of marginal products and hence wages is the income tax schedule itself. Because the income tax influences labor supply and different adjustments will change the relative labor effort of different types, the tax system influences the marginal products and hence the wages associated with various ability vectors, a subject explored in section 3.2.

In addition, the standard framework assumes perfect competition, so wages equal corresponding marginal products. Relatedly, employers and financiers are abstracted from in the basic setup, so there is no occasion for possible asymmetric information in labor markets. The unobservability of individuals' w's and l's to the government, which is at the core of the optimal income taxation problem, is imagined not to infect market interaction. Market actors will often know much more than does the government, especially about those with whom they have direct dealings. Nevertheless, market participants' information about each other may be imperfect, and sometimes in ways that may mirror the government's limitations.

Each of these elements—and more—indicate important assumptions to relax and complications to explore. Only some have been examined extensively in prior literature, and only a selection is considered here. Before turning to those topics, however, it is useful to reflect on how the foregoing relates to inequality, particularly increases in the inequality of labor income in recent decades in many developed economies.

To a substantial extent, one can apply the standard model as is. When $f(w)$ is taken to be fixed, there is no possibility of increasing inequality. After all, w is innate ability, one's marginal product, and the market wage. Genetic evolution is far too slow to bear on even centuries of changes in wages and the distribution of labor income. But technological change, a focus of much work on inequality, is central. Likewise for policies, for example, relating to international trade, even setting aside the income tax. An additional feature, often accelerated by changes in communications technology, involves changing preferences, which are socially influenced. These factors and more determine not only the overall degree of inequality but also involve changes in rankings. As explained, the marginal product associated with different ability vectors is endogenous. Most highly compensated coders today would have performed very different tasks even a decade ago, many being associated with lower marginal products but also some with higher marginal products that have been eroded by technological change.

Abstracting from the endogeneity of marginal products, and hence wages, to the income tax itself, many of these phenomena can be examined in the standard model. We may now interpret individuals' w's in a given era as the prevailing marginal products and wages. Given current technology, preferences, and policies, we may think of the standard formulation as a reduced form, where a more explicit statement might posit a function $w(\phi)$, where ϕ indicates individuals' underlying multidimensional ability vectors and the resulting w for each type in a given economy and era is the type's wage, which is a sufficient statistic for optimal income tax analysis regarding both behavior and normative assessment.

Advances in technology do not inherently increase inequality, although many of them have done so in modern history. Counterexamples may include technologies that boost the productivity of individuals' with physical challenges, raising the relative wages of individuals disproportionately at the lower end of the income distribution. As mentioned, in earlier times large portions of the population lacking in physical strength were relatively disadvantaged. Or a technology may raise the productivity of an ability type that was compensated near the top, but if demand is not expanding sufficiently, this boost in effective supply might cause that type's wage to fall.

Taking the case of increasing inequality in the effective distribution of wages, which attracts current attention, it is often supposed that this type of change favors greater redistribution, taking the social welfare function as given. The analysis, however, is more subtle when one considers a number of matters explored in section 2: there are ambiguities in the meaning of a more redistributive income tax, different tax adjustments may be optimal at different parts of the income distribution, and, regarding the underlying change in $f(w)$, subtle differences in the shape matter. Moreover, changes in the mean are also relevant to social welfare in ways that influence optimal redistribution. Finally, even with no change in the tax schedule, individuals with increasing incomes will automatically pay more if marginal tax rates are positive. For many possible causes of changes in the distribution of wages, there is no qualitative change in the analysis of optimal income taxation, which takes $f(w)$, whatever it may be, as a parameter. It is also true that many of the avenues of research now receiving greater attention due to rising concerns about inequality have long been important but were underdeveloped; greater exploration is warranted regardless of whether changing circumstances require changes in methods of analysis.

3.2. *Multidimensional Abilities and Endogenous Wages*

When different types' labor effort is not fungible and, moreover, wages are endogenous, income taxation has additional distributive effects. Early work by Feldstein (1973), Allen

(1982), Stiglitz (1982), and others introduced the subject and reached conclusions that provide a benchmark for subsequent work. In Stiglitz's (1982) two-type model, starting with the familiar result that the optimal marginal rate on the high type is zero, a reduction in that rate (to a negative value) now raises social welfare because it tends to equalize the distribution of equilibrium wages. High types increase labor supply, which reduces their wage but, due to complementary in production, their greater labor supply raises the wage of low types. Similarly, a higher marginal tax rate on low types, in reducing their labor supply, now boosts their wage relative to that of high types, providing a redistributive benefit.[26]

Before considering more recent work that extends this analysis, it is worth noting why pecuniary externalities—here, the effects of an individual's labor effort on others' wages—are relevant to welfare in this setting. Socially costly redistribution through income taxation is employed because individualized lump-sum taxation is infeasible, which means that the second fundamental theorem of welfare economics is inapplicable. Greater equality in the pretax wage distribution reduces the need to rely on distortionary taxation. Furthermore, once marginal income tax rates are positive, any effects on labor supply—including how some individuals' responses affect the labor supply of others—entail fiscal externalities. In short, pecuniary externalities that are irrelevant in many policy analyses are central here, even though we continue to assume that markets are otherwise perfect.

Sachs, Tsyvinski, and Werquin (2020) extend Stiglitz's (1982) two-type model by allowing for a continuum of types.[27] In their model, ability levels correspond to distinct, fixed occupations. Raising the marginal income tax rate at any point in the income distribution directly benefits workers of the type who earn precisely that income: their tax payments do not increase directly (all the income they earned is inframarginal) and, by the envelope theorem, their reduction in labor effort is a matter of indifference; but the fall in their labor effort increases their type's wage. If one posits universal complementarity, all other types' wages fall and, through this channel, so does their utility. Much of Sachs, Tsyvinski, and Werquin's (2020) investigation analyzes particular additional assumptions that enhance the model's tractability. In the spirit of Stiglitz (1982), they find that the optimal asymptotic marginal tax rate is lower when wages are endogenous.[28]

Rothschild and Scheuer (2013) analyze a model that departs further from Stiglitz (1982) and identifies additional effects that arise when wages are endogenous. Individuals are characterized by two ability parameters corresponding to their skill in two distinct occupations that are complementary in production.[29] This depiction can be taken as a simplified, reduced

[26] An implication is that sharp marginal rate reductions on low types, such as through the EITC, reduce their wages, an effect found in Rothstein (2010).

[27] See also Ales, Kurnaz, and Sleet (2015) and Chen and Rothschild (2015).

[28] Sachs, Tsyvinski, and Werquin (2020) devote substantial attention to a special case in which the opposing effects of endogenous wage changes on individuals' incomes are equal. As a result, if the income tax were linear, the resulting revenue impact from this component would net to zero, whereas this term for the general equilibrium impact on government revenue would be positive if the targeted type had an above-average marginal tax rate, such as may be present toward the bottom and top of the income distribution under a U-shaped nonlinear income tax schedule. A full analysis requires consideration of the other effects, including the direct effects of adjusting wages on individuals' utilities as well as effects that arise in a model with exogenous wages, which is taken as the benchmark in much of Sachs, Tsyvinski, and Werquin's (2020) analysis.

[29] Rothschild and Scheuer (2014) extend this analysis to the case of many occupations. Of further note, some literature that examines two occupations refers to those in the higher-wage occupation as managers or entrepreneurs, but such designations matter only if they correspond to changes in the model. For example, Scheuer and Werning (2017) consider the possibility that some individuals, perhaps including some managers, may have marginal products and thus wages that rise in their own labor supply (in their model, by being moved to tasks with higher

form in which, as elaborated in section 3.1, individuals are each characterized by many traits, but it is sufficient to know how these traits map to marginal productivity in the two occupations. Using the Roy (1951) model, each individual chooses the occupation that yields the highest wage, and the individual's labor effort depends on the net-of-tax wage in the chosen occupation. In each occupation, there will be some resulting equilibrium distribution of abilities and thus realized wages. It is helpful to interpret the model for the case in which one of the distributions unambiguously dominates the other (even though they overlap for all types), so we can refer to a high- and a low-wage occupation. They further assume that the nonlinear income tax schedule is a function of individuals' incomes but not occupations, because occupations are intrinsically difficult to observe, classifications are manipulable, and there may be political economy problems if separate schedules were contemplated.[30]

Local perturbations of the income tax schedule cause differential effects on the two occupations. If marginal tax rates are raised near the top, for example, there will be relatively more suppression of labor effort in the high-wage occupation. A direct effect will be to raise relative wages for individuals in that occupation, wherever they are in the ability distribution for individuals pursuing that occupation. This effect dampens but does not reverse the effect of the higher marginal tax rate in reducing labor effort in that occupation. In addition, this reduction in labor effort will, due to complementarity, reduce relative wages and hence labor effort in the low-wage occupation. Note that, because these distributions overlap, these relative wage effects will be partially muted because the tax increase near the top hits some workers in the generally lower-wage occupation. This overlap also mutes the resulting reduction in redistributive effects because some in the high-wage occupation that experience higher wages had low wages to begin with and some in the low-wage occupation who experience lower wages had high wages at the outset.

An additional channel in the Rothschild and Scheuer (2013) model is that types who were at the occupational choice margin will shift from the low- to the high-wage occupation, which dampens the wage rise in that occupation as well as the wage fall in the low-wage occupation. Keep in mind that the posited increase in the marginal tax rate was on high income, not on the high-wage occupation as such; hence, it does not directly affect the occupational choice margin of any individuals, so all occupational switching is caused by the general equilibrium change in wages between the occupations. Moreover, the effect of individuals' switching occupations is limited to the impact that these switches have on relative wages: these individuals at the occupational choice margin experience no change in utility by the envelope theorem, and these shifts do not directly affect tax revenue because these marginal individuals earn the same income in either occupation.[31]

Taken together, we can see that relaxing the assumption that wages are exogenous means that a facially more redistributive income tax will redistribute less than in the standard model due

---

marginal products due to greater complementarities), which favors lower marginal tax rates. Section 3.3 examines founders, whose compensation, as a consequence of information asymmetries in their labor market, comes through ownership in their firms and hence is qualitatively different from that in the standard model.

[30] If the tax schedule could depend on the occupation—even if it was observed imperfectly—the framework introduced in section 6.1 would be applicable. Because marginal tax rates could be targeted at occupations as such, relative wages could be influenced more directly.

[31] If the occupations differed in nonpecuniary ways, for example if one of them tended to generate more disutility of labor effort, compensating wage differentials would arise, in which event marginal switchers would cause fiscal externalities that would influence optimal tax rates (a feature of Lockwood, Nathanson, and Weyl 2017a, discussed in section 4.1). This complication makes the multidimensional screening problem more challenging and also may introduce welfare assessment issues related to heterogeneous preferences, which are examined in section 8.2.

to partially offsetting effects on the distribution of pretax wages. Nevertheless, the overall impact of endogenous wages in offsetting redistribution through income taxation is less in Rothschild and Scheuer's (2013) model than it is in the simpler Stiglitz (1982) model, which has only two distinct (and thus nonoverlapping) types and no occupational choice margin.

The examination of multidimensional abilities and occupational choice in a setting with endogenous wages constitutes an important advance as well as a subject warranting further study. In the past, the present, and one suspects the future, much of the evolution in inequality is through changes in the relative returns to different skills. Greater analysis of differential substitutability versus complementarity seems central. For example, at the lower end of the wage distribution, many occupations involve few skills so there may be approximate fungibility, with individuals moving across occupations in response to changes in demand but not earning differential equilibrium wages as a consequence. In more skilled jobs, there may be less substitutability, particularly in the short run, but if underlying cognitive talents can be developed and deployed in different ways through occupation-specific investments in human capital, there may be less complementarity in the long run than meets the eye. Nevertheless, as discussed in section 3.1, ability is deeply multidimensional, substitution across basic talents is often limited, and returns to different traits have varied substantially over time. Because equilibrium wages are endogenous and, moreover, influenced by the income tax schedule, richer models have the potential to advance understanding beyond what can be gleaned from merely examining how the optimal income tax changes given some exogenous change in $f(w)$.

3.3. *Asymmetric Information and the Income of Founders*

This section examines the income of individuals who supply labor effort to firms that they have founded, manage, and continue to own in whole or in part. For simplicity, these individuals will be referred to as founders.[32] A sizeable portion of total income, especially at the upper end of the income distribution, is earned by founders. The optimal taxation of this income raises distinctive practical and conceptual questions reflecting that founders' earnings are often entangled with their own supply of capital. This phenomenon, in turn, arises in large part because of asymmetric information that generates moral hazard and adverse selection problems that standard optimal income tax models assume afflict the government but not market actors. Although varying strands of literature, most not part of the formal analysis of optimal income taxation, have considered aspects of this subject, further exploration that is more explicitly embedded in the Mirrlees framework seems promising.[33]

Schumpeter (1947) recognized the distinctive feature of founders' efforts and struggled to classify them, seeing their earnings as arising from labor but distinct from ordinary wages. More recently, Kerr, Nanda, and Rhodes-Kropf (2014) emphasize the experimental and hence risky nature of many of these enterprises. Their economic importance and centrality to inequality are highlighted by Smith et al. (2019), who find that the top 1% of individuals in the U.S. income distribution earn over \$500 billion a year in pass-through income and that over \$400 billion of this income (sometimes treated as capital income in measures of wealth inequality and of the

[32] Another natural label would be "entrepreneurs," but this term is avoided because it has frequently been used in the optimal tax literature, other economics literatures, and more broadly to mean a variety of things, many of which are unrelated to the present focus. For example, much discussion and data about entrepreneurship covers many millions of individuals, most of whom moonlight, work in the gig economy, or otherwise operate fairly simple enterprises—although the challenges examined here can arise even at small scales.

[33] For example, Cullen and Gordon (2007) and Gentry (2016) consider a number of issues, many generated by complexities of actual income tax systems. Some of the analysis here builds on this and other prior work.

labor share) might best be understood as the labor income of founders.[34] This sum dwarfs earnings of top CEOs and many other groups receiving much attention, including in work on optimal income taxation (Ales and Sleet 2016; Piketty, Saez, and Stantcheva 2014). Many individuals at the top of Forbes lists in the U.S. likewise are founders (and many of the rest are heirs of founders). Also notable are Bhandari and McGrattan's (2021) measures of sweat equity, which are of comparable magnitude to annual GDP and public market capitalization.

Given the attention to the top of the income and wealth distribution, the range of proposals to raise taxes on this group, and the sheer magnitudes involved, it would be useful to extend optimal income taxation models to capture the relevant behavior. Prior extensions to the Mirrlees model that include both labor and capital income typically feature each in a pure, separate form, with the latter usually taken to be risk-free interest or the return on a (common) passive investment portfolio. With founders, by contrast, labor and capital are often entangled. In accord with seminal finance literature in the 1970s (Leland and Pyle 1977, Ross 1977), founders often hold significant stakes in their own companies—forgoing diversification—as a consequence of moral hazard and adverse selection. In both cases, the pertinent information problems are intertwined with founders' provision of labor effort.

This phenomenon raises a number of additional issues. Founders' returns to what may be deemed their savings—prior savings invested in their firms or earnings retained in these enterprises (the aforementioned sweat equity)—involve risk that is both idiosyncratic and highly correlated with founders' own human capital. These features are central drivers of Hall and Woodward's (2010) finding that, in their base case, founders of tech firms funded by top Silicon Valley VCs approximately broke even (rather than earning great riches) when computed on an ex ante, risk-adjusted basis, reflecting in significant part the large portion of these firms that were unsuccessful. Indeed, one cannot properly examine the optimal taxation of highly successful founders without considering as well how failures are to be taxed.

Building models that incorporate these features should be a high priority for researchers in optimal income taxation. Moreover, whether different tax instruments are used optimally may be highly consequential, not only regarding the core tradeoff of redistribution and labor supply distortion but also with respect to possible externalities of the sort that will be explored in section 4.1: some startups may generate substantial spillovers through various forms of innovation that creators cannot fully appropriate (Nordhaus 2004, Jones forthcoming). In addition, otherwise optimal schemes may require the observability of investments, returns, or valuations that, with privately held companies, are exceedingly difficult for tax authorities to measure, particularly given that the firms' financiers themselves often have trouble doing so. Hence, the optimal use of more restricted instrument sets needs to be considered as well.

Analysis of the optimal taxation of labor and capital income has tended to treat the two types of income as distinct, even when earned by the same individuals. For example, in the familiar application of Atkinson and Stiglitz (1976) to savings, it is assumed that individuals earn ordinary labor income that they may either consume or instead save, using the principal and interest (capital income) to fund future consumption. Subsequent work, including the new dynamic public finance literature, extends this simple story in a number of directions but largely maintains this distinction (Golosov, Tsyvinski, and Werning 2007).

Analysis of the effects of income taxation on portfolio allocations began with Domar and

[34] Their data includes sole proprietors, partners, and owners of S-corporations and hence these remarkable totals omit the income of most operating venture-capital-backed firms that are usually organized as taxable corporations, which undoubtedly contribute significantly at the top of the income distribution and should be analyzed similarly (although prevailing tax rules are different).

Musgrave (1944) and was extended from partial to general equilibrium analysis in Bulow and Summers (1984), Gordon (1985), and Kaplow (1994). In these models, individuals allocate savings (maintaining the distinction from labor income) between a riskless asset and a market portfolio. The latter investment is taken to be commonly available to all investors (and, in Kaplow 1994, to the government) and to involve systematic risk that the government cannot eliminate.

By contrast, founders' investments in their own firms typically are private—available to a handful of specialized investors—and involve the bearing of idiosyncratic risk that could be diversified by idealized markets or, in principle, by the government. Founders' utility would be higher if they diversified their holdings, indeed, if they held no equity in their firms or even negative positions that offset the idiosyncratic risk associated with their firm-specific human capital. However, as the finance literature in the 1970s emphasized, moral hazard and adverse selection limit the extent of such diversification. Indeed, many founders own most or all of the equity in their firms, perhaps along with family and friends.

One might try to embed this problem in the standard optimal income taxation framework by attempting to disentangle founders' labor and capital income and apply the models accordingly. However, such a disaggregation, in addition to being practically difficult and subject to manipulation given that these are private firms (that even specialized financiers have difficulty fully penetrating), is not the correct approach even in theory. In this setting, labor and capital income are not Platonically distinct categories (and, even if they were, they would not be the right categories). The return to founders' capital is determined by their own labor effort. Moreover, founders own equity in their firms for reasons entangled with their supplying of labor. This is obvious regarding moral hazard, where an ownership share is retained to motivate effort. For adverse selection, much of founders' asymmetric information is generated by their prior labor effort, and yet more concerns their information about their own future efforts. For example, their attempts midstream to sell equity may be taken by the market as a signal of their desire to curtail their involvement with the firm.

Begin with moral hazard. When financiers own equity in a firm, the founder's incentive to supply labor is correspondingly dulled.[35] It is natural to ask whether labor income taxation should accordingly be viewed in a new light because we now have two labor wedges on the same incentive margin. Compare what might appear to be the analogous context in which the government provides insurance, such as disaster relief, in settings that also feature private insurance—or even in settings with purely private insurance but from multiple insurers. The layering of two or more mechanisms of this type, each contributing to moral hazard, is an inefficient means of addressing risk (Pauly 1974, Kaplow 1991). The reason is that a private insurer and an insured party will devise contracts that maximize their joint surplus, equating the marginal risk mitigation benefit to the marginal incentive cost that they bear, ignoring that reduced incentives also impose a negative externality on other insurers, including the government. Indeed, if moral hazard were the only problem (there being no redistributive motive, say, because all individuals were identical), the optimal income tax would be zero at all income levels, and moral hazard would be optimally addressed by market transactions. Alternatively, if the income tax was designed to optimally trade off risk and incentives, it would be necessary to prohibit private insurance or other risk-reducing financial arrangements that would generate the aforementioned fiscal externality.

[35] This problem is, of course, ubiquitous, as it arises in public companies due to the separation of ownership and control and in basic employer-employee relationships, even if the principal is the sole owner or is taken to be a perfect representative of the owner. Much of the analysis here is applicable more broadly.

When there is a redistributive motive, however, positive marginal income taxation is optimal, raising the familiar incentive tradeoff. The underlying analysis is much the same when extended to market arrangements that themselves involve moral hazard. In the standard model, despite otherwise-perfect private markets, positive marginal tax rates distort labor supply. Individuals supplying labor—whether contracting with firms or financiers or simply optimizing in their roles as self-employed, one-person firms—ignore the negative fiscal externality imposed by their reductions in labor supply. This phenomenon is largely the same when one introduces moral hazard in the labor market.

To explore this claim, return to the case of moral hazard between founders and financiers (or between employers and employees). Compare three scenarios. (1) The contracting parties employ an incentive scheme to limit moral hazard: founders retain more equity than is optimal purely on risk-bearing grounds, or employees receive performance pay.[36] (2) Financiers (or employers) pay wages to other workers to undertake monitoring of the founders (or employees), and founders (or employees) compensation is certain. Suppose further that this certain compensation equals the certainty equivalent under (1) for each level of effort and that the marginal resource cost of inducing incremental effort through monitoring precisely equals the additional risk premium associated with inducing the same incremental effort using the incentive scheme in (1). (Observe that the risk premium measures the true social cost of these schemes, for it is the difference between what the financier or employer pays and the certainty equivalent of what the founder or employee receives.) (3) There is no moral hazard, but the production function is such that the marginal effort induced by the same compensation as in (2), along with an additional labor input costing the same as the monitoring effort in (2), generates the same output as in (2). That is, we have the same production function and compensation to our founder or employee, but we now simply reinterpret the monitoring input as some other input to production.

It is clear that these three scenarios are economically equivalent in most respects. Financiers (or employers) make the same expenditures and receive the same returns. Founders (or employees) exert the same labor effort and receive the same compensation in utility terms. The one possible difference is that, even though the income received by taxable individuals is the same, labor income tax revenue received by the government will not be the same in scenario (1), compared to the other two scenarios, if the income tax schedule is nonlinear. This reflects the familiar point, developed further below in the discussion of adverse selection, that rising (falling) marginal tax rates treat risky payouts less (more) favorably than under a linear tax, where only total (or expected) income matters.

Moral hazard by itself thus does not substantially alter the optimal income taxation problem as applied to founders. Note further that moral hazard may not be an important consideration for many founders in any event because they fully own their own firms, including (approximately) some cases in which some finance is provided by family and friends or when there are a few owners who observe each other's efforts to a substantial extent.

Now consider adverse selection.[37] Whether at a firm's creation, in subsequent periods in

[36] Doligalski, Ndiaye, and Werquin (2020) explore a case in the latter setting in which the employee chooses effort after learning the resolution of uncertainty.

[37] This section focuses on adverse selection in the financing of founder enterprises. More broadly, the interaction of adverse selection in labor markets and optimal income taxation deserves further study. Craig (forthcoming) shows how adverse selection dulls incentives to invest in human capital (such investments create positive spillovers on other workers with whom one will be grouped by employers), a factor that reduces optimal income tax rates. The analogue here would be that prospective founders tend to underinvest in developing potential projects to the extent that they intend to raise external funds. See also Scheuer and Netzer (2007), Scheuer (2013), and Stantcheva (2014).

which founders wish to raise additional funds, or at any point when founders wish to sell some or all of their stakes, financial arrangements are often impeded by asymmetric information—here, about founders' abilities as well as their ideas and knowledge of the environment. If moral hazard were the only problem, risk-averse founders would always wish to sell at least some equity because the incentive distortion is initially zero at the margin, although once some equity is held by financiers, moral hazard may be an important reason that those holdings are not larger.[38] Substantial expenditures by financiers in the selection and oversight of their investments as well as founders' frequent retention of large (even complete) equity stakes suggest that adverse selection is a significant problem in the finance of founder enterprises.

With adverse selection, unlike moral hazard, private contracting is inefficient among actual or potentially contracting parties. Government intervention can sometimes improve social welfare, such as by providing insurance when markets would otherwise unravel. The income tax has likewise been viewed in this manner, including with respect to risky investments. In this regard, it is important to distinguish systematic risk, particularly that associated with publicly traded firms, and idiosyncratic risk, which is central for founders. Regarding the former, if the tax schedule is linear (including in the negative range, that is, allowing for full loss offsets), it is understood that income taxation has no effect regarding risky returns because individuals adjust their portfolios in a manner that maintains their net positions. For example, with a riskless asset and a market portfolio (as in the capital asset pricing model), introduction of a 50% tax on risky returns (i.e., returns net of one's investment, and abstracting from any tax on the riskless return) induces individuals to double their holdings of the market portfolio and continue to receive the same net return in every state.[39]

Founders' holdings in their own firms differ in important respects. Because these are unique assets, founders cannot simply and fluidly gross up their holdings. Instead, when the government imposes a linear income tax, the treasury becomes a financier for each founder, by fiat rather than by contract and accordingly not impeded by asymmetric information. If implemented using a cash-flow tax—or, equivalently in this simple setting, if founders are permitted to expense their investments—the founder's optimal investment would not change, viewing the investment decision in isolation.[40] (For example, with a 50% tax rate, the cost of any investment would be half as much, and half the return would be enjoyed, so the optimal investment, holding labor effort fixed, would be the same.) Such an income tax, however, applies to the total return of the firm, which incorporates the return to labor effort as well; hence, this tax would also reduce the net-of-tax return to labor and diminish labor effort. This latter effect, of course, is standard, and would arise even without an investment decision by the taxpayer. Finally, note that if the marginal products of labor and capital were complementary,

---

[38] Nevertheless, the contracting parties on this account will leave the founder holding too little equity because they do not internalize the positive fiscal externality from reductions in moral hazard. This force opposes the tendency for founders to hold too much equity on account of adverse selection. Put another way, the lack of internalization of the fiscal externality acts as a government subsidy that encourages equity sales in a market in which they are inefficiently low on account of adverse selection. However, as noted, many founders are at a corner solution wherein they own all of the equity in their firms, so these forces are inoperative at the margin.

[39] For this result to hold in general equilibrium, we can likewise suppose that the treasury, which now in essence owns the market portfolio to the extent of the tax rate, sells or shorts the market portfolio in an offsetting manner (Kaplow 1994).

[40] The exposition in the text can be understood as taking advantage of a collapsed dynamic interpretation in which risk is instantaneously resolved. By contrast to the text's assumption that investment costs are deductible (or, if spread over time, depreciable), as in many income tax systems, Scheuer (2014) takes, these expenses to be unobservable and hence not deductible, which makes it optimal to reduce entrepreneurs' income tax rates, which improperly but unavoidably apply to gross rather than net income in this setting.

this reduction in labor effort would reduce the marginal return to capital as well. That too is standard, in the sense that optimal investment decisions in an economy with a labor income tax reflect the level of labor supply that will arise in equilibrium, taking into account the effect of income taxation on labor effort.

The foregoing analysis of founders' investment decisions, however, is importantly incomplete because it has yet to factor in risk aversion. Higher income taxation reduces the dispersion of the founder's net-of-tax returns, which conveys a risk-reduction benefit. Moreover, we are supposing that this benefit was not fully available by contract due to asymmetric information. It may be helpful to contemplate cases in which the founder owns the entire firm because adverse selection is sufficiently serious.

Furthermore, this reduction in the riskiness of the founder's investment returns tends to encourage investment. Unlike the earlier example with a market portfolio, however, increasing the founder's investment in what is taken to be a unique asset will be associated with diminishing returns. Hence, we would not expect the founder to restore her original exposure to risk.[41] Note further that this tax-induced increase in investment will, with complementarity, increase labor effort, that is, relative to the lower level described above that involves the usual labor supply reduction associated with income taxation.

In summary, with founder enterprises, raising income tax rates tends to increase total investment in founder firms that are assumed to be private, subject to idiosyncratic risk, and suffering from adverse selection on account of asymmetric information. The corresponding increase in labor effort does not itself increase utility due to the founder's envelope condition, but (relative to the ordinary reduction in labor effort) causes a positive fiscal externality. Furthermore, because the government is absorbing idiosyncratic risk from many enterprises, the utility gain from reducing founders' risk exposure is a real welfare gain to the economy, in contrast to systematic risk which can be reshuffled but not extinguished.

The analysis of founders raises a number of further questions regarding optimal income taxation. First, the foregoing discussion assumes that the income tax is linear, including full loss offsets. As mentioned earlier, if marginal tax rates are rising (falling), risk-taking—here, founders raising their investment in response to higher tax rates—will be less (more) than otherwise on that account. Of particular importance, there are often important limitations on the deductibility (including refundability) of losses.[42] In the simple case of no loss offsets, risk-taking may be sharply penalized: founders share their gains with the government but bear the entirety of their losses. This asymmetry both reduces expected returns (abstracting from risk aversion)—unlike with the purely neutral scheme that reduced investment costs and returns by the same proportion—and also eliminates the most attractive portion of risk sharing, since offsets in the lowest realization states are the most valuable component of an income tax system's implicit insurance. Finally, these limitations are particularly significant in the present setting because, for each highly successful founder enterprise, there are typically large numbers of failures. Treating the latter disadvantageously diminishes ex ante incentives.[43]

---

[41] In addition, founders are often wealth constrained, which limits their ability to increase their investments, particularly at early stages. Note, however, that wealth constraints tend not to be absolute. Instead, they are a matter of degree that reflects founders' willingness to reduce their consumption. Hence, the risk-sharing created by income taxation influences this very decision rather generally.

[42] These exist in part for practical reasons: it may be difficult to police fraudulent claims of losses, and some consumption (certain hobbies) may be disguised as business activity in order to generate inappropriate deductions.

[43] This and other differences between actual and idealized income tax schedules—such as those involving the treatment of appreciation (including capital gains rules and lock-in) and complications of organizational form and

Second, it is assumed here that a founder's enterprise simply generates "income" that is subject to the income tax. However, many theoretical analyses and most actual income tax regimes in practice distinguish labor and capital income, taxing them at different rates and often using qualitatively different methods. To the extent that such contrasting treatment is not optimal, it would be best to merge the two.[44] Otherwise, substantial challenges in the taxation of founders' income must be confronted, for we are imaging a single production function that transforms financial investments and founders' labor effort into some return. Moreover, in the present setting even specialized financiers who hold significant stakes in these firms have great difficulty observing labor effort and valuing enterprises (i.e., they cannot ascertain the production function).

Third, suppose that further research, including empirical evidence, suggests that the optimal taxation of founders differs nontrivially from the optimal taxation of others who earn similar incomes, be they CEOs, engineers, or blue-collar workers. Ideally, separate tax schedules may best be employed, but as noted in section 3.2, this will often be infeasible due to difficulties of observability, manipulation of classifications, and political economy concerns. Hence, as there, one would be constrained to set optimal tax rates at different levels of income in a manner that reflects the proportions of different types earning those incomes.

Relatedly, one would also wish to take into account general equilibrium effects as well as endogenous occupational choice. In the analysis in section 3.2, the latter was relevant only on account of general equilibrium effects on wages: because marginal individuals' incomes are the same regardless of their occupational choice, these choices involve no fiscal externalities. By contrast, prospective founders who are at this margin would generally make different investment decisions, bear different levels of risk, and exert different amounts of labor effort depending on their choices. Hence, the decision to become a founder instead of a worker will often be associated with significant fiscal externalities. In particular, if founders—despite the insurance provided by the income tax—bear more risk, their expected earnings as founders need to be greater than what they would earn as workers (assuming that there is no uncertainty in that occupational choice, or at least less uncertainty). Therefore, those choosing to be founders pay more taxes, and because the risk in those payments is substantially idiosyncratic, the government's higher expected tax receipts raise social welfare. Finally, to the extent that founders on average generate positive spillovers, as some suggest, there is an additional reason that income taxation should favor founders.

Finally, consider briefly the application of this section's analysis of founders to some of their financiers, notably, venture capital funds as well as private equity funds and some other entities sharing some of these traits. Like founders, many individuals who work at such financial

multi-level taxation—are examined in Cullen and Gordon (2007) and Gentry (2016), largely outside the formal optimal income taxation framework.

[44] The simple statement in the text is susceptible to multiple interpretations. Under a pure, accrual, Haig-Simons income tax, all income, from both labor and capital, is taxed continuously, at a common rate. But many deviations therefrom reflect the challenges of implementing such a system. Another interpretation involves the use of a cash-flow tax, which need not distinguish labor and capital income. In idealized forms, the difference between these two regimes involves only the tax burden on the riskless, real return to capital. Because this return has been near zero for quite some time, and because of difficulties in administering an idealized Haig-Simons income tax, some analysts support a cash-flow tax (even if they in principle favor taxing capital income). Moreover, some suggest that such a cash-flow tax would capture a significantly greater portion of founders' "labor" income, by preventing it being disguised as capital income and then taxed lightly or not at all (such as through the use of misleadingly low valuations of equity compensation, deferral, and step-up basis at death). The most relevant point for present purposes is that many of the challenges of income taxation in many guises, as well as of wealth taxation, are particularly great with regard to founders' income.

enterprises supply labor—in selecting investments, serving on boards, and offering managerial services that the founder firms lack. Some of this labor effort is precisely to combat moral hazard and adverse selection regarding the founder firms. However, when we view these financial intermediaries as themselves firms with their own financiers, we can see that often the pattern is replicated. A venture fund's financiers likewise are concerned about about both moral hazard and also asymmetric information, finding it difficult to value the fund managers' skills and opportunities. For these reasons, venture fund's (and other fund types') principals often own stakes or are compensated in ways that involve their retention of equity-like positions. As a consequence, many of the issues considered in this section have similar implications for how those individuals should optimally be taxed.

4. *Externalities and Rents*

The correction of conventional externalities is usually best accomplished using targeted instruments, even when such correction influences labor effort and thus interacts with income taxation, a subject explored in section 7. This section, by contrast, addresses externalities and rents directly caused by or associated with labor effort. In this setting, optimal income taxation is more fundamentally implicated because the labor wedge itself has additional effects on social welfare or an additional labor wedge may be present. Put another way, in a world with homogeneous individuals and thus no distributive concern, the income tax may be a corrective instrument. Hence, the full mechanism design problem can be stated as setting optimal labor wedges on different types in light of the effects on distribution and externalities associated with labor effort as such.

Section 4.1 examines the possibility that labor effort, particularly that of individuals in certain occupations, may cause positive or negative production externalities. In the standard Mirrlees setting, these are usually taken to operate through effects on others' wages, continuing to employ the background assumption that wages equal marginal products. Section 4.2 considers a wide variety of cases in which additional labor effort—by increasing actors' utility or consumption—raises or lowers the utility of other individuals because the utility functions of the latter depend on the circumstances of the former. As will be explained, in some instances these two sorts of externalities are similar. Section 4.3 addresses how market power and its associated rents interact with the optimal income taxation problem. Market power has distributive effects, through markups imposed on consumers and profits received by firms' owners, and it also influences the net return to labor effort by reducing wages' purchasing power and the derived demand for labor.

4.1. *Production Externalities from Labor Effort*

Suppose that some individuals' labor effort raises or lowers the productivity of others' effort or otherwise influences the productivity of others' use of resources. To the extent that positive or negative externalities are caused by labor effort as such, it might seem natural that the income tax, which directly influences the labor wedge, would be a useful corrective instrument. Before analyzing how such externalities affect optimal labor income taxation, some framing comments are helpful.[45]

[45] It is important to distinguish this phenomenon from the general equilibrium effects examined in section 3.2. Those involve pecuniary externalities that affect social welfare on account of the labor wedge induced by income taxation, whereas production externalities are welfare relevant even in an economy with no income taxation.

Many externalities—whether from research, teaching, or finance, to mention some of the occupations commonly noted in this strand of optimal income taxation literature—typically arise from certain outputs rather than from labor input as such. Hence, conventional instruments that directly target the relevant outputs tend to be the most efficient means of correction. And indeed, research is subsidized in various ways and much education is publically provided or otherwise subsidized. Securities transaction taxes have been proposed on externality grounds and, to the extent that such taxes are poorly targeted (because many taxed activities are efficient), one suspects that income taxes weighing more heavily on part of the income distribution featuring more financial professionals who undertake particular sorts of tasks would fare much worse on that score. Most analysis to date has focused on general income taxation as a corrective tool rather than on occupation-specific income tax schedules (higher rates for financiers or certain types of financiers, lower rates for researchers), largely for reasons noted in section 3.2: difficulties of observing occupations, strategic reclassification, and political economy concerns. As we will see, this constraint greatly complicates and limits the efficacy of externality correction via income taxation.

To analyze optimal income taxation, we can supplement the basic Mirrlees model from section 2 with externalities, following some of the lines of analysis pursued in Lockwood, Nathanson, and Weyl (2017a) (although they consider only Pigouvian correction and not redistribution). Raising the marginal tax rate at a point (or in a neighborhood) in the income distribution reduces labor effort of those thereby targeted and, as a result, imposes a negative fiscal externality. This concept captures the idea that individuals who reduce their labor effort are indifferent at the margin (by their envelope condition) but do not take account of the reduction in tax revenue that results. If the labor effort of workers in some occupations causes positive and negative externalities of the conventional sort, we would add a weighted-average conventional externality component to the fiscal externality term. If, at the targeted level of income, workers' marginal labor effort across all occupations causes net negative externalities, for example, higher marginal tax rates than otherwise would be optimal. And if there were income effects, those earning higher incomes would work more, so in addition to the resulting fiscal externality offset, we would also have a conventional externality offset (or augmentation, if the sign of the net externality at higher incomes differs). For example, if marginal labor effort of those at higher incomes similarly caused net negative externalities on average, these income effects would mitigate (and could outweigh) the welfare gain from the net correction at the targeted income level.

This adjustment is straightforward conceptually but would be daunting to calibrate in practice. Perhaps the very top of the income distribution is dominated by founder-entrepreneurs whose great success generates positive spillovers, but a bit further down there are more financiers who generate negative externalities, with a mix of engineers, CEOs, lawyers, other founders, and real estate developers below that, generating all manner of positive and negative externalities. The optimal externality supplement to the conventional fiscal externality adjustment would be difficult to quantify and probably even to sign.

Next, recognizing that different externalities are naturally associated with different occupations, section 3.2's extension of the Mirrlees model to include occupational choice and general equilibrium effects on wages is appropriately introduced here as well. Focusing for simplicity on individuals directly targeted by the change in the marginal tax rate, raising the income tax reduces labor effort, which will raise (lower) relative wages in occupations that are more (less) prevalent at the targeted income level, partly mitigating (accentuating) the direct effect on labor effort. Moreover, as discussed previously, the effects on wages apply to everyone in such occupations, not just those at the targeted income level. In addition, some individuals will change occupations, moving from those with reduced wages to those with increased wages (which, recall, involves no fiscal externality because switchers' incomes are equal between the two occupations in a model in which heterogeneity is confined to abilities).

All of these effects will further complicate the determination of the net effect on conventional externalities. For example, if a negative-externality-causing occupation is particularly prevalent at the income level targeted by an increase in the marginal tax rate, the rise in the relative wage will, through both the intensive margin adjustments in labor supply at all levels of income and occupational switching, mitigate the corrective benefit. Likewise, if a positive-externality-causing occupation is most prevalent, the welfare loss from reducing generation of the positive externality will be muted.

The analysis can be further enriched in another way by introducing, as in Lockwood, Nathanson, and Weyl (2017a), a further dimension of heterogeneity. Specifically, suppose that individuals differ not only in their abilities in each occupation, a feature present in Rothschild and Scheuer's (2013) analysis that was examined in section 3.2, but also in their nonpecuniary utilities from each occupation. Occupations that are more dangerous, stressful, or boring will, in equilibrium, be associated with positive compensating wage differentials, so many individuals may be choosing between higher-paid but less pleasant occupations and lower-paid but more appealing ones. A basic feature of such a world—even without occupation-specific conventional externalities—is that an income tax favors lower-paying occupational choices: when an individual receives more compensation to take a less pleasant job, the additional pay is taxed, but the (unobserved) offsetting utility loss is not deducted.[46] An immediate implication is that individuals at the margin, who are indifferent between two occupations, impose a negative fiscal externality should they choose the one that pays less.[47]

Consider how this occupational choice margin has additional effects when one includes conventional externalities. Abstracting from income effects, consider a local increase in the marginal tax rate. In addition to all the earlier effects, we now have that every individual who earns income above the targeted level pays more tax, so those who were at the margin with another occupation that pays below the targeted income level will switch to the lower-paying occupation. Only such switches become appealing because other occupations paying more than the targeted income are equivalently subject to the tax increase and occupational pairs below the targeted income are unaffected.

These switchers, as always, are indifferent at the margin, but their changes in behavior have two first-order effects. One is the aforementioned fiscal externality, which is unambiguously negative. The magnitude of this effect depends on the income levels earned before and after the switch—which together determine the resulting difference in tax paid. Each of these income levels will be different for different switchers. Some may switch from a very high income to a very low one, and others from an income only slightly above the targeted income level to one only modestly below it. One must know not only the total number of switchers but also the before and after incomes of each to compute the magnitude of this negative fiscal externality.[48]

---

[46] Nonpecuniary disutility of being an entrepreneur could motivate the analysis in Scheuer (2014) that posits a nondeductible expense of that occupational choice and hence may justify reduced income tax rates on entrepreneurs, or at levels of income where there are more entrepreneurs. But if entrepreneurs experience nonpecuniary utility, the opposite adjustment would tend to be optimal.

[47] This additional dimension of heterogeneity is examined further in section 6.2, which discusses some literature that examines this extension in addressing income transfers effect on the participation margin.

[48] Consider an example noted in the literature: a Wall Street financier or lawyer who is induced by higher top income tax rates to become a teacher. It is easy to imagine that the negative fiscal externality substantially exceeds

The effect of these occupational switchers on the overall magnitude of conventional externalities involves related considerations. For each switcher, one must know the externality for each of the two occupations. A negative-externality financier or lawyer may become a positive-externality teacher. A very-high-positive-externality researcher may become a moderate positive-externality teacher. A positive-externality round-the-clock entrepreneur may become a no-externality manager or worker.[49] Note further that each switcher's optimal labor effort may change, and the individual's externality for both the before and after job choices depends on that adjustment as well.

We can see that changes in both the fiscal externality and the conventional externality associated with these occupational shifts will be challenging to estimate. These effects, combined with all of the foregoing factors (direct, intensive margin effects, income effects, and general equilibrium effects on wages), together determine the optimal adjustment to the income tax schedule. It is thus unsurprising that Lockwood, Nathanson, and Weyl (2017a) as well as Rothschild and Scheuer (2014, 2016) warn us of the difficulty of determining optimal policy in this regard. Returning to where we began, Lockwood, Nathanson, and Weyl (2017a, 2017b) conclude that it probably makes the most sense to employ conventional, targeted instruments rather than adjustments to "an untargeted income tax [that is] struggling to precisely reallocate individuals."[50] Suggestive simulations (that, recall, abstract from redistributive considerations) suggest that the income tax alone can achieve only a sliver of the corrective gains whereas direct subsidies (notably, for research) can achieve in excess of forty times as much.[51]

Additional investigation has focused on a particular form of negative externality from labor effort: extracting rent from others. Piketty, Saez, and Stantcheva (2014) analyze a case in which top executives at public companies extract rents through bargaining, making higher top income tax rates beneficial through discouraging this activity. In their analysis, the extraction is in essence from owners (shareholders), taken for simplicity to be pro rata (all individuals own the same amount of equity in public firms). As the authors explain, it is as if the executives are taking funds directly from the treasury, which can be appreciated by noting that this rent extraction is equivalent to a reduction in the magnitude in the grant in the income tax schedule (involving an equal contribution from everyone in the population) that is then transferred to the executives. Note that if one instead models shareholdings as proportional to wealth, the compensatory tax equivalent would be a reduction in the rate of capital income taxation (the extraction by executives is akin to raising the capital income tax on equity and giving them the proceeds). If one sticks to a model of bargaining for higher wages that comes directly from the fisc, perhaps the closest fit is public-sector unions, in which case higher tax rates toward the middle of the income distribution may be optimal.

---

the full salary of the teacher. Nevertheless, if there were a large negative externality associated with the high-paid job or a sufficiently high positive externality from teaching, this switch would raise social welfare.

[49] Reflection on the complexity of the problem brings to mind a core challenge of multidimensional screening: in a given equilibrium, the matching of traits to occupations reflects selection and hence is correlated even if the underlying distributions are not. For example, those with atypically high nonpecuniary costs of the higher-paying occupation will accordingly also tend to have atypically high ability in that occupation if they are currently choosing it.

[50] By contrast, Jones (forthcoming) considers a setting in which research subsidies are infeasible, making it optimal to employ lower top marginal income tax rates because of research spillovers that contribute to economic growth.

[51] It is notable that they reach this conclusion even abstracting from distribution, for when the income tax schedule employs high marginal tax rates on that account—rather than optimally being zero throughout in the absence of conventional externalities—any gains from adjusting marginal income tax rates will have to trade off distributive effects. Relatedly, fiscal externalities are large to begin with under a significantly redistributive income tax rather than arising only to the extent that nonzero marginal income tax rates are introduced for externality correction.

A complication with such prescriptions arises directly from the previous analysis regarding targeting. On one hand, public company executives' incomes are not at the very top of the income distribution and in any event are mixed with others, notably, much larger numbers of founder-entrepreneurs (Smith et al. 2019), many of whom might generate positive externalities, favoring perhaps lower income tax rates at the top. Moreover, executives of public companies already have their pay directly targeted by the U.S. income tax code (more aggressively after recent tax reforms), and others suggest that improvements to corporate governance are even better targeted than that.

Rothschild and Scheuer (2014, 2016) provide a particularly subtle analysis of rent extraction that focuses on which other workers are victims of negative-externality-producing activities. Motivating examples are credit-claiming among workers and high-speed traders, who may impose most of their negative externalities on each other. To that extent, there is some self-correction—their efforts implicitly tax each other—and correspondingly, an income tax increase, by reducing these efforts, has an offsetting effect of raising the returns to those very rent-seeking actions. Once again, some of these externalities, such as those in finance, might best be targeted directly. And, stepping back, one must be cautious about wholesale condemnation of credit claiming because some credit is actually due. When principals cannot observe ability and effort very well, the resulting pooling diminishes incentives. Although signaling is costly, zero signaling is hardly optimal in most settings with imperfect information because it undermines incentives to invest in human capital, the efficient matching of workers to jobs, and contemporaneous labor effort.

The occupations examined in the literature on rent extraction—whether public company executives, high-speed traders, or various credit-stealing members of teams—are, as in the other examples, just some of the many occupations represented at various income levels. These analyses indicate that determination of the negative externalities associated with such activities is more difficult than may first appear. The broader challenge, of course, is that even if such effects could be estimated for certain individuals, one would still need to determine the relevant weighted averages at different income levels for all occupations, for occupation-switchers at other income levels, for related fiscal externalities, and so forth in order to know what adjustments to the income tax schedule are appropriate on account of these externalities.

4.2. *Externalities through Individuals' Utility Functions*

A longstanding literature on externalities through individuals' utility functions has developed independently of the more recent literature on production externalities caused by individuals' labor effort.[52] Nevertheless, it is worth recognizing at the outset that these subjects have much in common, at least in certain cases associated with each. Put simply, with production externalities, some individuals' incremental labor effort may raise or lower other individuals' wages (other than through general equilibrium effects, examined in section 3.2). With externalities through utility functions, some individuals' incremental labor effort—perhaps as a consequence of raising their own consumption—may raise or lower other individuals' utilities. In the former case, the effect is on others' effective w's and thus on others' budget constraints, whereas in the latter case, some effects on actors' own circumstances are taken to be arguments in other individuals' utility functions.

[52] Earlier work on the former, which mostly focuses on envy, includes Brennan (1973), Boskin and Sheshinski (1978), Layard (1980), and Oswald (1983), and is surveyed and extended in Tuomala (2016). Hochman and Rodgers (1969), by contrast, focused on sympathy toward the poor.

If one aligns the externalities, including the relevant functional forms, and accounts for the existing income tax schedule, the results may likewise be aligned, at least roughly. The effect of earning a lower wage (because of a negative production externality from others' increased labor effort) is to enjoy less utility from consumption from a given level of one's own labor effort. This is also what happens if a change in some arguments in one's own utility function (because of the utility externality from others' increased labor effort) reduces the consumption utility one can achieve from a given level of disposable income. This claim encompasses not only mechanical effects (consisting of direct reductions in others' utilities, holding their labor supply constant) but also labor supply effects and, accordingly, related fiscal externalities. Regarding effects on behavior, reducing the net-of-tax wage, $w\big(1 - T'(wl)\big)$, and shifting downwards individuals' marginal utility of consumption, u_c (when those effects are aligned in magnitude, at different levels of earnings and consumption), each has the same effect on individuals' first-order condition for labor effort (expression (2) in section 2.1).[53] Consequently, there are potential research gains from cross-fertilization, wherein findings or methods in one realm might usefully be applied in the other. These connections will not be examined explicitly here (and they were not in section 4.1 either), but in principle they often seem present and may sometimes be worth pursuing.[54]

Focusing explicitly on externalities through individuals' utility functions, the central theme of this section is that this possibility involves a quite heterogeneous set of problems (Kaplow 2008a). Hence, we should not be seeking the correct or best model, or even a handful, for these diverse phenomena. Empirical exploration may well justify attention to some particular sets of assumptions, but as will be noted at the end of this section, this is a setting in which reliable evidence of the operative channels and relevant magnitudes is hard to come by.

To begin, note the variety of pertinent dimensions. First, these externalities can readily be positive, such as when individuals are altruistic, or negative, when they are envious. Second, these concerns may involve others' before-tax income, consumption (after-tax income), leisure, or utility—with qualitatively different implications, including as to the sign (envy of others' consumption and of their leisure call for opposite Pigouvian corrections aimed to influence labor effort). Third, individuals' may care about levels and magnitudes or about their rankings, with only the former motivating corrections. Fourth, individuals may care fairly similarly about the full distribution in society, about those near themselves in the distribution, especially about the rich (who may be envied) or the poor (why may generate empathy), or about those who are more local, which might refer to geographic neighbors, friends, family, co-workers, co-congregants, or other groupings.

It is apparent that the number of possible configurations is large. Rather than considering them all or selecting a favored few, each dimension will be examined briefly, drawing out some

[53] It might appear that the two phenomena have different fiscal externalities from differences in the transmission mechanism. For example, lowering others' wages mechanically reduces income tax revenue whereas lowering others' utility from the consumption of after-tax income does not. However, in the latter case, if one considers settings in which the utility effects are uniform and all-encompassing, the effective utility purchased with nominal government dollars is reduced even though the dollars themselves are not, which is to say that there is a reduction in the value of government dollars rather than in their amount.

[54] For example, it will be discussed below how externalities through individuals' utility functions may sometimes be local in various ways rather than global. Likewise, production externalities involving the marginal product of labor may sometimes be local rather than global, depending on the labor markets involved and the importance of trade. In a fairly straightforward case, one's increased labor effort may steal credit and thus income from others on a team, and likewise the higher earnings by one team member may be envied by and thus reduce the utility of other team members.

of the implications for optimal income taxation. In all, this section aims to broaden the research agenda, only some of which has been explored. Closing remarks will be offered on empirical challenges and normative questions raised by crediting these sorts of preferences.

Suppose that every individual's utility depends on others' utilities and, moreover, that this dependence is global via a utilitarian SWF: that is, the sum of all individuals' utilities enters into each individual's utility function. If individuals are altruistic, this additional source of utility is additively separable, and the planner's SWF is also utilitarian, the externality has no effect on the social optimization. The planner maximizes the sum of the (narrow) utilitarian SWF plus the same utilitarian function weighted by the total of individuals' altruistic weights.[55] If instead the sum of others' utilities enters each individual's utility function with a negative sign (envy), the result is the same, short of cases in which the total of individuals' negative weights exceeds one, implying that social welfare thus defined would be maximized by minimizing the sum of utilities.

These results are offered merely as a conceptual starting point. To see how fragile they are, note initially that, if the altruistic (or envy) component is not additively separable, the optimal income tax would in general differ and perhaps in subtle ways. Likewise, if the planner's SWF is instead strictly concave in individuals' utilities, our base case with altruism (envy) may favor less (more) redistribution than otherwise. Consider altruism: if we add a common constant (some weight on the sum of utilities) to everyone's utility, utility levels rise but with no change in the absolute dispersion, so there is a fall in relative dispersion, which in turn will tend to reduce the difference in welfare weights between the rich and the poor. Finally, the "social" welfare function that is part of individuals' preferences may have a functional form different from the planner's or it may depend on different objects. Indeed, it could be any function whatsoever, so anything is possible.

Suppose next that the externality is measured in others' consumption rather than their utility, and consider for concreteness the case of envy, as is commonly done in this setting (Frank 1999, Ireland 2001). Here, each individual's consumption creates a negative atmospheric externality, making it optimal to impose a Pigouvian correction—a tax—justifying higher marginal tax rates than otherwise. The analysis is similar if it is before-tax income that creates the externality, as suggested by Atkinson (1983), although the optimal correction may have a very different magnitude and pattern across the income distribution. Taking the special case in which everyone's uncompensated labor supply elasticity is zero, no correction would be appropriate. Raising marginal tax rates, aside from its behavioral effect (which is absent in this special case), has a mechanical effect on the magnitude of the externality when the externality depends on other individuals' consumption, but not when it depends on their before-tax income. Furthermore, with redistributive taxes, the difference between before- and after-tax income (the latter being consumption) varies greatly across the income distribution. Notably, the sign of the gap reverses, from negative to positive, when the break-even point is crossed. Hence, the nature of the Pigouvian correction to the optimal nonlinear income tax schedule depends, both qualitatively and quantitatively, on how the target of individuals' envy is internally calibrated.

This sort of negative atmospheric externality is familiar. Consumption can, of course, cause conventional negative externalities as well, such as through pollution and congestion. Likewise, there may be positive externalities through altruism or warm-glow effects as well as more familiar ones, such as agglomeration externalities. The relevant consideration in our present context is that aggregate consumption, rather than particular activities or expenditures, is what causes the externality.

---

[55] The recursion is left implicit, and it is assumed that the total is finite.

Consider finally the mostly neglected case in which the (still assumed to be negative) externality is caused by others' leisure. Veblen (1899) and Duesenberry (1949) helped popularize the focus on conspicuous consumption, but often leisure is more conspicuous, and at least it may also be conspicuous. It may be easier to know how long one's neighbors were away on vacation than how much they spent. It may be obvious that a neighbor does not engage in market employment, and visible signs of physical fitness (noisily) signal significant time available for exercise. To the extent that leisure is what is envied, the analysis is reversed from the case with consumption: lower taxes are optimal because they reduce leisure, which is the source of the negative externality. More broadly, if both leisure and consumption cause the externality, and they do so to similar degrees, then we may be back to a case in which other individuals' utilities should be the focus, where we saw that, in a basic case, the externality through utility functions has no effect on the optimal income tax schedule.[56]

With respect to both conspicuous consumption and conspicuous leisure, it may be that optimal income tax rates should be changed little in any event, with correction achieved instead through the selective adjustment of commodity taxes and subsidies. For example, if it is particular luxuries that create envy from others—perhaps automobiles or certain home improvements—those should be taxed. Maybe personal swimming pools should be subject to the highest tax of all if the pools themselves make neighbors jealous and the poolside is a particularly glaring way to observe neighbors' leisure, which is a source of jealousy.

Regarding each of these possibilities, there are other dimensions as well. Individuals may envy the rich but empathize with the poor, in which event greater redistribution would be favored in many of the above cases.[57] Consider as well that preferences about others' circumstances can depend on rankings rather than magnitudes. For example, individuals may care a great deal more about what portion of the population is ahead of them, but less so about how far they are head. In that case, there may be no implication for the extent of redistributive income taxation because adjustments preserve rankings (and, in extensions with additional dimensions of heterogeneity, where rankings may change, that reshuffling may well be largely orthogonal to income and thus to the optimal income tax schedule).

Turn now to another important dimension of this problem: both positive and negative preferences about others' circumstances may well be local, where the applicable notion of locality can take many forms. Individuals may care about "keeping up with the Joneses," who are often imagined to be neighbors rather than stand-ins for the vast, anonymous mass of fellow citizens.[58] The pertinent localized group may be geographic neighbors (the swimming pool example), siblings, former classmates, co-congregants, or co-workers (who themselves may be those in one's area of an office, at one's entire facility, at one's firm, in one's job classification, in one's cohort, or in some weighted intersection of those categories). For many of these cases,

---

[56] For example, Boskin and Sheshinski (1978) show how the argument that envy favors greater redistribution assumes that it is relative income or consumption rather than utility that is the object of envy, and they further warn that the evidence for the former is "virtually nonexistent, let alone convincing" (1978, p. 599).

[57] Not all—for example, if leisure is important. If individuals get disutility from observing the poor working hard at unpleasant jobs, this would favor not only higher transfers but also higher marginal tax rates, both of which reduce labor supply, taken to be the source of this negative externality. By contrast, Wane (2001) supposes that individuals suffer disutility from the shortfalls of the poor's earnings (before-tax incomes) from the poverty line—rather than consumption shortfalls—which can favor negative marginal tax rates. And if individuals are envious of the leisure time enjoyed by the rich, then lower marginal tax rates at the top end would be corrective.

[58] The analysis in the text of this case (like that of others in this section) sweeps aside many subtleties and hence further variation in the phenomena. For example, Dupor and Liu (2003) describe the need to distinguish the effect of the Joneses' consumption on one's utility level and on one's marginal utility of additional consumption (which helps one catch up with the Joneses).

the implications for optimal income taxation would be similar. Suppose that every individual envies some localized set of others' consumption levels, creating a local rather than atmospheric externality but one that involves, very roughly, each individual's consumption imposing a similar total negative impact on others' utilities. In that case, higher marginal tax rates throughout the income distribution may be optimal. If absolute or relative effects vary by income, so would the corrections because most of these "local" groupings tend to involve individuals with similar income levels. But those are just some of the possibilities.

Another form of localized externality is illustrated by individuals' concentrated concerns about homeless populations located near their own residences or places of work because it is the direct observation of others' raw poverty that generates disutility. The analysis is similar for more tangible negative externalities associated with poverty, such as when there are costs associated with public health or safety. In all of these instances, assistance to the poor may be a local public good, which helps to explain local income redistribution and other local expenditures on the poor—whereas if these concerns were entirely society-wide, the free-rider problem predicts little local effort (Pauly 1973).

By contrast to such localized concerns, these same envious or empathetic individuals may care little about the rich or the poor who are out of sight and out of mind.[59] Or they may even have opposite-signed preferences. For example, some may find it entertaining to view the lifestyles of the rich and famous, including royalty, even while they envy their neighbor who just bought a fancy new car that is quite basic compared to a Rolls Royce or royal carriage. This particular configuration might favor generally higher marginal tax rates—supposing that neighbors' consumption rather than utility or leisure causes the externality—but lower rates at the very top.[60] Or, as is often associated with NIMBYism in housing markets, the same individuals may have negative preferences regarding poor individuals who are nearby but positive regard for those living afar.[61]

The breadth and nature of externalities that may arise through individuals' utility functions raise empirical and normative questions. Of the myriad psychological reactions that individuals might have to the contemplation of others' prospects, it is all too easy to privilege one or another that aligns with one's predilections or reinforces one's policy preferences. John Stuart Mill famously claimed that poetry was superior to pushpin based on his own experience of the two activities, on the implicit assumption that projection provided insight into others' subconscious experiences, apparently not appreciating that his own mind and background were highly atypical.

Economists favor an empirical approach, but empirical assessment of these features of individuals' utility functions is difficult to undertake. Modern research uses surveys,

---

[59] Perhaps the richest individuals in society are more aware of each other, in part from media coverage, Forbes listings, and the like. In that case, if envy within this group is of others' consumption levels (and perhaps much less so if it is of others' incomes), they would mutually gain from all of them working less, favoring higher marginal tax rates. However, given the fiscal externalities associated with their labor effort, the overall effect of their additional labor supply on total welfare is likely positive. If one focuses on the revenue-maximizing top marginal tax rate, this phenomenon would matter primarily to the extent it leads to different estimates of the elasticity of taxable income for these individuals. Perhaps they "keep score" based on pre-tax income, and marginal tax rates thus have little disincentive effect.

[60] State finance of royal extravagance would illustrate the alternative strategy of financing what is here imagined to be a public good.

[61] Instead, the externalities associated with others' circumstances may be the same, but there is a separate, negative impact from experiencing the poor nearby. Hence, individuals may get positive utility from helping the local poor—taking as given that they are local—but oppose local aid because, holding these individuals' utilities constant, the local well-to-do would prefer that current and prospective future poor individuals live elsewhere.

experiments, and field work to address such questions, but problems of elicitation and external validity can be particularly great in this context. After all, these preferences have a social nature, where context can be all-important. Information is also relevant but influenced by the attempt to measure these phenomena.[62] For example, when prompted, individuals may be willing to pay for the rich to be taken down a notch or for the poor to be helped, but during most of their lives these considerations may, as mentioned, be largely out of sight and out of mind. There is often support for policies that keep the homeless off the streets or at least confined to streets in other neighborhoods, and it is familiar that many prefer to cross the street rather than confront the homeless, which might also afford the opportunity to lend assistance, the anticipation of the inclination to do so perhaps adding to the desire to avoid the interactions.

Furthermore, the above discussion emphasizes many dimensions involving subtle differences that could be consequential, including functional form assumptions.[63] In principle, much of this could be measured, but we might not be confident in the results. Taking an analogy, many people have concerns about the fate of animals, particularly species with certain qualities. However, attempts to elicit contingent valuations can produce results that differ by orders of magnitude depending on subtle differences in how scenarios are described or the sequence in which they are presented. Perhaps the relatively low magnitude of charitable donations—particularly to aid distant beneficiaries with whom one has little direct affiliation—suggests that many of these possible external influences on individuals' utilities, whether regarding animals, humans, and other causes, are fairly small, although the free-rider problem and limited information provide alternative explanations. Also, much evidence supports warm-glow rather than altruistic motivations. To the extent that it is one's own act of giving rather than the effects on beneficiaries that matters, there may be little impact from centrally administered redistributive tax and transfer programs, by contrast to local homeless shelters that one might personally support financially or with one's time. However, if most individuals suffer disutility from neighbors' conspicuous consumption expenditures, those expenditures can indeed be reduced by upward adjustments to marginal tax rates, so the posited benefits would materialize. Of course, we would still need to measure those benefits' magnitude and just what they depend on, for example, by examining individuals' revealed preferences to leave near individuals with higher or lower consumption levels than their own.

This article is not an appropriate place to survey, even superficially, the literatures just referenced or the many more that are pertinent. The preceding discussion is selective and intentionally argumentative, meant to raise questions rather than to answer them—which, as stated, would require evidence. An optimal income taxation agenda focused on externalities through individuals' utility functions greatly needs empirical guidance that, unfortunately, will be hard to come by.

Normative questions are also raised by the prospect of giving weight to such preferences about other individuals' circumstances. Economists generally tend to use an individualistic SWF—that is, one that assesses policies based solely on how they affect individuals' utilities—and, moreover, to adopt a subjectivist view of individuals' utility functions that credits whatever

[62] Relatedly, empirical work in this realm poses challenges confronted in the literature that attempts to measure happiness or well-being. The importance of, say, envy versus neoclassical measures of utility that limit consideration to one's own consumption mirror some of the debate about whether rising national incomes improve well-being. See the competing perspectives in Easterlin (1974) and Stevenson and Wolfers (2008).

[63] Even fairly subtle features can be highly consequential, as suggested by section 8.2's discussion of heterogeneous preferences.

preferences individuals have.[64] As discussed further in section 8, this approach is supported by the Pareto principle, which is to say that using a nonwelfarist SWF or failing to respect individuals' actual utility functions entail accepting Pareto-dominated policies in some settings.

The most commonly expressed concern in the present context is with the crediting of negative other-regarding preferences such as envy.[65] Taken as an intrinsic objection, the position is either difficult or too easy to sustain. Why are neighbors' noise, fumes, congestion, or even lack of taste (such as with zoning restrictions on aspects of homes' appearances) all externalities that should count, but not their swimming pool or Jaguar, which may bother others even more? To illustrate the point about Pareto conflicts, suppose that the latter preferences do not count and hence Pigouvian corrections are not made. From that baseline, it might be possible to allow wealthy neighbors to make more noise or have uglier homes, but fewer swimming pools, making both themselves and their onlooking neighbors better off. Simply put, the externality-causers may like noisy parties more than swimming pools, and their neighbors may be bothered more by swimming pool envy than by noise.

One can simply stipulate lists of the preferences that count and those that do not, as many have done through the ages, but how are we to resolve disagreements?[66] Indeed, the objecting neighbors disagree and may well resent elites or governments telling them that their preferences are too vile to be credited. There are also conceptual conundrums with the concept of preference censoring because it is then necessary to choose some alternative utility function (that differs from the pertinent individuals' actual utility functions) from among myriad possibilities, and these choices may have implications opposite to what is contemplated when objecting to the preferences being censored.[67]

Consider briefly a functionalist, welfarist take on objectionable preferences, in the spirit of the analysis in section 8.1. For example, social and governmental opposition to, rather than crediting of, racist and sexist preferences is believed—independently of its immediate, direct effects—to usefully shape preferences over time. To the extent that individuals are socially induced to develop positive rather than negative attitudes toward others, society can subsequently achieve higher overall welfare, understood as higher levels of experienced utility, for any given level of resources. Perhaps the envied rich or envied neighbors should be viewed similarly. A society that loves its neighbors more may be happier overall and also interact more effectively. Similar prescriptions may apply, but in reverse, to positive other-regarding preferences, the fostering of which both may raise total welfare for given resources and also

[64] The most systematic deviation in modern research agendas involves behavioral economics, which mostly raises different issues. There are some overlaps, such as the possibility that altering available information or an individual's focus can itself affect their utility, as well as normative questions about how such situations should then be assessed.

[65] The discussion that follows draws on the more extensive critical exploration in Kaplow and Shavell (2002) of the mostly philosophical literature that argues against crediting various preferences deemed to be objectionable.

[66] An analogy more familiar to economists involves "merit" goods. If the term is not a stand-in for goods whose consumption imposes actual positive externalities—or for goods that, due to information and self-control problems, would otherwise be under-consumed—just what is the basis for deeming some goods more meritorious than others and thereby adopting welfare-reducing policies to favor those goods?

[67] Suppose that censoring a preference means subtracting (treating as if zero) the satisfaction it generates. Does one then implicitly assume that any resources that had been spent on the now-censored preference are removed from the individual's budget constraint, or are they treated as if they had been spent elsewhere, and if so, on what? If the SWF is concave, in which event utility levels and not just marginal utilities matter, does society now seek to direct more resources to such individuals—who are now deemed to have lower utility because of the censoring—some of which will then be spent satisfying the objectionable preferences to an even greater degree? As a further note, some advocate censoring all other-regarding preferences, including positive ones. This would seem to entail social policy ignoring affection for others, including spouses and children, with all of the implications that would entail.

enhance trust, which itself is productive.

Although the foregoing extends well beyond the scope of this article, and some would say beyond the domain of economics, it is offered because these questions are in the air and sometimes on the page when exploring the issues considered in this section. This presentation, like the comments on empirical challenges, is offered primarily to encourage engagement rather than to promote particular claims. Some regard positional preferences or other preference independencies as among the key mechanisms through which inequality affects society and therefore as important forces that should guide society's responses, including through redistributive taxation. Accordingly, addressing these issues should be part of the optimal income taxation research agenda.

4.3. *Market Power and Rents*

Concerns about inequality are sometimes associated with the level of market power in the economy and increases in margins in recent decades.[68] This connection raises two sets of questions. First, if much income at the upper end of the distribution arises from monopoly rents, might greater redistribution than otherwise be optimal?[69] On its face, this question cannot readily be answered in the standard optimal income tax framework that assumes perfect product and labor markets, does not envision profits, and accordingly does not attend to the distribution of ownership interests.

Second, if market power itself involves a labor wedge, might less redistribution be optimal? After all, markups in product markets raise prices, thereby reducing the real wage, and quantity reductions by sellers with market power are associated with reduced input demands, notably, for labor.[70] Suppose, for example, that market power resulted in a labor wedge of the same magnitude as that created by labor income taxation. Then, following the familiar rule of thumb, total distortion would be four times higher than in a world with just one of these sources of distortion. Taking the monopoly wedge as given, income taxation would cause three times the total distortion ordinarily assumed, with the marginal distortion being twice as large.[71] This point is dramatized by Bilbiie, Ghironi, and Melitz (2019), who, in their analysis of endogenous product variety, suppose that, in the background, the government imposes a linear income tax at a *negative* rate (financed by a uniform lump-sum tax) to offset the labor wedge due to market

---

[68] Studies finding large average markups include Domowitz, Hubbard, and Petersen (1988), Rognlie (2015), Feenstra and Weinstein (2017), Hall (2018), and De Loecker, Eeckhout, and Unger (2020). See also the more skeptical discussions in Basu (2019) and Syverson (2019).

[69] A corresponding question (answered in the negative at the end of this section) is whether the broad range of policies influencing the state of competition should be toughened (relative to a benchmark of maximizing efficiency) and perhaps also tilted toward the maximization of consumer surplus rather than total surplus. Such suggestions have long been advanced—Robinson (1933) and Comanor and Smiley (1975)—and have recently received increasing attention (OECD 2017, World Bank Group and OECD 2017). For example, the U.S. Horizontal Merger Guidelines (2010) and EU Guidelines on Horizontal Mergers (2004) seem to embrace an exclusive focus on consumer surplus, which in some policy discussions is motivated by the difference in the distributive incidence of markups versus that of profits.

[70] Discussions of market power involving a labor wedge include Lipsey and Lancaster (1956), Hart (1982), and World Bank Group and OECD (2017).

[71] This view is suggested in varying ways by Browning (1994), Kaplow (1998b), and Jonsson (2007). That work, however, is partly informal and uses representative-agent models, an approach associated with Ramsey (1927) that can be misleading for reasons elaborated in section 7. A variation on this theme, with heterogeneous abilities, is explored in Eeckhout et al. (2021).

power.[72] Of course, that income tax—which seems benign in their representative-agent model—is the opposite of what is done in practice in a world with distributive concerns, and the difference between their hypothesized income tax and either actual or optimal income taxes indeed involves (crudely) a wedge on labor income that is twice what is usually contemplated.

The model and analysis in Kaplow (2021), however, shows that both views are seriously incomplete. Indeed, in a simple, benchmark case, market power has no effect on the optimal income tax, properly interpreted in real terms. To explore this question, we can first supplement the basic Mirrlees framework by introducing multiple of goods or sectors, across which individuals allocate their disposable (after-income-tax) incomes. Suppose next that, in each sector, instead of price equaling marginal cost, there is an exogenously given markup, with the difference between price and marginal cost generating profits to owners of firms producing in that sector. Finally, assume that there is a common portfolio of all of the economy's firms and that individuals' ownership shares are a function of their income (without further restriction). This ownership assumption encompasses the special case in which ownership is pro rata but also captures more realistic and interesting cases in which higher-income individuals own (perhaps substantially) greater shares. Recalling that the Mirrlees framework is often interpreted as a collapsed dynamic model, one might suppose that individuals' portfolios consist of their savings, the magnitude of which rises with income. This formulation also captures the possibility that those with larger portfolios have access to higher-return investments.[73]

To analyze optimal income taxation in this model, begin with the special case in which markups in the economy are proportional across sectors. As a consequence, market power will not distort individuals' consumption choices, leaving only our core question of how market power affects the labor wedge and, through this channel, optimal income taxation. In this case, the result is that there is no effect whatsoever. An economy with proportional markups is in an important sense equivalent to an economy with no markups. Moreover, the optimal income tax in the economy with no markups—the solution to the standard Mirrlees problem—corresponds (mechanically) to the optimal income tax in the economy with proportional markups, resulting in individuals of every type supplying the same labor effort, choosing the same consumption bundles, and achieving the same utility. Furthermore, the government's budget continues to balance under this corresponding income tax.

To see this, begin in the economy with proportional markups and suppose that there is in place any nonlinear income tax that satisfies the government's budget constraint (expression (4) in section 2.1). From there, we can construct a corresponding income tax for the economy with no markups under which individuals' budget sets are the same and hence individuals' behavior and utility are the same and the government's budget still balances. This construction has two components. First, raise the income tax at each income level so as to tax away all of the income that was spent paying markups (which we now imagine are no longer present). Second, lower the income tax at each income level so as to rebate the profits received from individuals'

[72] Their analysis (and some work in international trade) is partly inspired by Lerner's (1934) argument that proportional markups do not involve inefficiency. Lerner insightfully and (albeit parenthetically) noted that these uniform markups had to include leisure for the claim to be valid. That, in turn, motivates a tax on leisure, which can be implemented by imposing a subsidy on labor income. Given the era, it is unsurprising that Lerner's analysis neither addressed inequality nor considered redistributive income taxation.

[73] Although ownership may rise with income, it is taken to be independent of ability as such. Relaxing that assumption would have familiar implications for the possible optimality of a tax on capital income because that income would then be correlated with unobserved ability (Golosov et al. 2013). That and some other extensions would not fundamentally alter the core logic developed in the text, although they may disrupt the strong equivalence result.

ownership shares (which profits vanish as well). Note that, if an individual (of any type, who earned any level of income) now chooses the original labor supply and consumption bundle in the transformed economy with no markups, that individual's budget constraint will hold. The goods are cheaper, but there is less disposable income to that extent, and profits are no longer available to spend on goods, but that loss has been rebated. Individuals' budget sets are therefore the same in these two economies for every level of income that they may choose to earn. Finally, the economy's total resource constraint is unchanged because the same labor is supplied under this tax and the same goods are demanded, which implies that the government's budget continues to balance as well.

Because this exercise works for any nonlinear income tax in the economy with proportional markups and, moreover, the construction can be run in reverse, starting with any nonlinear income tax in the economy with no markups, it follows that the optimal income tax problem is unchanged in real terms. Whatever tax is optimal on the economy with no markups—a Mirrlees economy—the corresponding income tax will be optimal in the economy with proportional markups. Note that these two income tax schedules, although equivalent to each other in real terms when imposed in their respective economies, are not nominally the same. In moving from the optimal income tax schedule in the no-markup economy to the constructed income tax schedule in the economy with proportional markups, income tax rates will be lower at all income levels to compensate for the markups and higher at all levels to, in essence, tax away the profits associated with ownership interests in firms that earn markups. If the distributive incidence of profits is more upward skewed than the distributive incidence of markups, as normally supposed, then the corresponding income tax schedule in the economy with proportional markups will be one that is relatively higher on the rich, compared to the income tax schedule in the standard Mirrlees economy with no markups. In that sense, the income tax in the economy with markups is more redistributive, but it should be kept in mind that, in real terms, we have after-tax equivalence.

This result, it should be apparent, is general in some respects but restrictive in others. Consider first that, in this special case, firms' profits are implicitly taken to be pure profits rather than quasi-rents that reflect the recovery of real resource costs. Investments are typically necessary to create the opportunities that generate price-cost margins, whether the investments are in facilities, research and development, or rent-seeking. Regardless of the source—which may be relevant to the optimal design of many other government policies—if we take these rent-generating processes as given, then we must extend the analysis to account for the resource use that underlies those markups.

To keep the analysis focused on our core concern with distribution and the labor wedge (rather than with intersectoral efficiency), assume that some common fraction of the proportional markups in each sector corresponds to real resource use, with the remainder constituting true profits. A natural case to contemplate is where the former fraction is one, which is to say that ex ante, all investments earn a common, risk-adjusted market return equal to a cost of capital. A variety of familiar models with free entry, including some models of rent-seeking, have this property.[74] In any event, the analysis to follow holds regardless of what this fraction happens to be.

We can analyze optimal income taxation in this economy in two steps. First, undertake

---

[74] Illustrative of the range of such work are Dixit and Stiglitz (1977), Mankiw and Whinston (1986), Aghion and Howitt (1992), Hopenhayn (1992), Ericson and Pakes (1995), and Melitz (2003). This result is also suggested by Hall and Woodward's (2010) calculations, noted in section 3.3, that founders of venture-financed firms roughly break even on an ex ante risk-adjusted basis.

precisely the transformation contemplated previously to eliminate the portion of the markups (if any) that corresponds to true profits. This move preserves equivalence and hence has no real effect on the optimal income taxation problem. The result is an intermediate economy in which all remaining markups correspond to real resource use.

Second, let us now eliminate these markups as well. This transformation requires accounting for the resource use that generated the remaining markups. Recall from above that, in our original analysis, the economy's real resource constraint implied that the government's budget continued to balance, something that will no longer be true here unless further modifications are made. There are two ways of viewing this difference between these two economies. One, developed in Kaplow (2021), shows that the resulting equivalent economy with no markups is one in which the distribution of individuals' abilities, $f(w)$, is shifted downward from that in our original economy. Another noted there is to suppose instead that the marginal costs of producing goods is shifted upwards. Either way, we can view the production possibility frontier in these economies as reflecting firms' production functions, including the investments that may generate markups that yield quasi-rents, and the real productivity of individuals' labor effort. Regarding the latter, recall section 3.1's discussion of how individuals' abilities are not a Platonic concept but rather a set of traits whose productivity and ultimate value depend on an economy's technology, individuals' preferences, and various policies.

The result of this analysis is that this two-step transformation places us in a standard Mirrlees economy—albeit with a different distribution of abilities or of marginal costs—that is equivalent to our original economy. Therefore, we again are able to solve for the optimal income tax schedule in the usual manner. To determine the corresponding income tax schedule for our economy with markups and resource dissipation, we need to reverse all of the steps in this set of constructions. That process is again mechanical, although more involved. Hence, we can use familiar tools to determine the optimal income tax schedule for our economy with proportional markups and resource dissipation. Many extensions of the Mirrlees framework could likewise be applied, mutatis mutandis.

Reflection on these corresponding income tax schedules suggests that there is no single, obviously correct way to compare the degree of redistribution entailed by the optimal income tax schedule in the economy under consideration, with markups and rent dissipation, to that in an economy without that rent dissipation but that had the same distribution of abilities and level of marginal costs. These are different economies with different production possibility frontiers. Taking the interpretation in which equivalence is generated by a downward shift in $f(w)$, that economy has less (absolute) dispersion in abilities but also a lower mean, which may have differing and subtle effects on the optimal extent of redistribution. The broader point is that, to understand how market power affects optimal income taxation in real terms, it is necessary to ascertain the underlying forces that generate the market power, specifically, the nature of investments, firms' production functions, and the translation of underlying ability into productivity.

Before leaving this subject, briefly consider relaxation of the assumption that markups and the resource use in generating them are proportional. In this more realistic setting, it is of interest to analyze policies aimed at influencing markups (notably, reducing high markups), but we will need to assess how such changes interact with the income tax.[75] Under weak

[75] Earlier work, surveyed in Myles (1995) and Auerbach and Hines (2002), considers the use of corrective commodity taxes and subsidies to offset markups in a representative-agent setting without distributive concerns or income taxation. Kushnir and Zubrickas (2021) examine a model in which the extent of redistribution has a feedback on markups.

separability of labor in individuals' utility functions, the basic technique developed more fully in section 7 can be used to show that the additional issues raised are essentially orthogonal to optimal income taxation analysis and hence policies that influence markups can be assessed independently.

To sketch the analysis, begin with the original construction of corresponding income tax schedules for the proportional case (with no resource use that dissipates profits), where we were concerned with equivalences. The income tax adjustments that left all individuals (at all income levels) with the same budget sets had two components: taxing away the income spent on markups and rebating back for the erased profits from ownership shares. When markups are not proportional and we are adjusting some policy that influences markups, we can undertake a similar neutralizing income tax adjustment that again results in individuals supplying the same labor effort and achieving the same utility. This tax schedule adjustment, however, is not the same as the one we had before because we no longer have equivalence. Specifically, this policy experiment changes price ratios, so individuals' consumption bundles will now change accordingly. Policy experiments that move markups toward proportionality (uniformity, such as with commodity taxes in the Atkinson and Stiglitz (1976) framework), tend to increase allocative efficiency in a manner that (before the tax adjustment) inures to the benefit of individuals. As a consequence, income tax rates need to be raised more in order to keep individuals on the same indifference curve for each level of income that they might earn.

As a consequence of this difference, the government's budget will, in general, no longer balance after a policy reform that is accompanied by this adjustment to the income tax schedule. The net change in the government's budget position is given by the sum of the two components in the income tax adjustments. The first taxes away the increase in consumer surplus (which rises, ceteris paribus, when allocative efficiency increases in the manner just suggested), and the second rebates the fall in producers' surplus (profits). Hence, the government will have a budget surplus if and only if the change in the sum of consumer and producer surplus is positive, which is to say, when efficiency, conventionally viewed, increases. In our suggestive policy experiment in which allocative efficiency rises, if we further suppose that productive efficiency is unchanged, then total surplus is higher and the government will therefore have a budget surplus under the income tax adjustment that holds everyone's utility constant. That budget surplus can be rebated, say, pro rata, so as to generate a strict Pareto improvement.

The conclusion is that, regardless of the distributive incidence of markups on consumers and of ownership interests, and regardless of how markups and profits themselves, as well as the income tax itself, affect labor supply, policies that reduce the inefficiency associated with markups can be implemented so as to generate Pareto improvements. Because this notion of inefficiency encompasses both consumers' and producers' surplus, this logic holds when full account is taken of how policies that affect markups may influence productive efficiency, although that feedback must be assessed to know whether a particular reform raises or reduces total surplus.[76]

5. *Behavioral Optimal Income Taxation*

Behavioral economics has increasingly been applied across a number of fields, including

[76] Kaplow (2020b) offers a complementary exploration of these efficiency effects that abstracts from redistribution and labor supply but makes endogenous prices, the number of firms, and markups in every sector. Policies influencing markups in a given sector now affect—directly and through general equilibrium effects—entry and exit in every sector as well as prices and markups in all sectors.

optimal taxation.[77] Much of this work focuses on differential commodity taxation employed to correct externalities and internalities.[78] But there are also important strands on optimal income taxation. Mirrlees (1971) identified "rational calculation" as a key assumption, which was only to be explored much later.[79] This section emphasizes three diverse lines of work that offer suggestive preliminary results and indicate promising areas for additional theoretical and empirical investigation.[80]

One set of work begins with the familiar, longstanding concern that income tax schedules—particularly when considering that, in the Mirrlees framework, these are taken to incorporate a range of tax and transfer programs—are too complex for most to understand. How, then, do individuals behave in light of their inevitably imperfect information about their true net-of-tax wages, which is compounded by their inability to process whatever information they possess? And what are the implications of any behavioral regularities that can be identified for the design of the optimal income tax schedule?

Other work considers respects in which even informed individuals may not be rational maximizers of their own utility, with an emphasis on present bias. One application involves myopic labor supply, including investments in human capital, where the delay between investment and ultimate reward is substantial. Another line of work begins with behavioral economics investigations of individuals' savings decisions and considers the implications—particularly of corrective policies such as forced savings funded by payroll taxes, savings subsidies, and automatic retirement contributions—for labor supply and hence optimal income taxation.

5.1. *Perceived Income Tax Schedule*

Actual, proposed, and optimal income tax schedules are nonlinear. There are often additional complexities involving deductions, the treatment of dependents, additional tax schedules (such as payroll taxes and, in federal systems, other jurisdictions' income taxes), and transfer programs (each with their own implicit tax schedules and other rules such as asset tests and time limits). Consequently, it is important to consider deviations between the actual, all-inclusive income tax schedule and the tax schedules that individuals perceive.

As a starting point, suppose that this gap consists of noise: individuals receive inaccurate but unbiased signals of their actual marginal tax rates. Even here, particularly if the noise is substantial, there may be important effects to consider.[81] Total distortion is nonlinear in tax rates

---

[77] For an overview and survey, see Congdon, Kling, and Mullainathan (2009) and Bernheim and Taubinsky (2018).
[78] Illustrative investigations include O'Donoghue and Rabin (2006), Allcott, Lockwood, and Taubinsky (2019), Farhi and Gabaix (2020). See also Goldin (2015) on using less salient commodity taxes to reduce distortion.
[79] Mirrlees (1971, p. 176) noted specifically that, especially at high incomes, individuals' motivation for supplying labor may involve sources of utility other than conventional consumption. This possibility has received some attention but has not been a focus of modern developments in behavioral optimal income taxation.
[80] Empirical investigation receives additional emphasis in this section for two interrelated reasons: there is little empirical evidence regarding key assumptions highlighted in these theoretical literatures, and to the extent that some of the posited assumptions have force, much prior work (such as on the elasticity of labor supply and of taxable income, and involving calibrated simulations) is misspecified. Regarding the latter, section 5.1 considers whether individuals take their average tax rate as their marginal tax rate (the latter featuring in existing empirical work), and section 5.3 asks whether payroll taxes that fund retirement savings (and for much of the population are of similar magnitude to income taxes) are perceived as tantamount to income taxes or as savings contributions that fund consumption in retirement.
[81] Compare Kaplow (1998a), who analyzes government errors in the assessment of individuals' taxable income rather than individuals' errors in estimating accurate government assessments.

and hence may be greater when some overestimate their marginal tax rates and others underestimate them. Errors can also reduce welfare through budget misallocation: those overestimating their taxes may underspend during a relevant period whereas those underestimating taxes may overspend; both mistakes reduce welfare (the errors do not cancel) to an extent that rises nonlinearly with the magnitude of the errors because utility is concave. Note further that heterogeneity in the systematic biases that are considered in the remainder of this section raises similar issues because total distortion and misallocation costs differ between cases in which, say, half the population exhibits the full bias versus when all of the population exhibits half the bias. Some of the most recent work attends to this difference.

Setting aside noise and heterogeneity within otherwise similar types, consider next the case in which the perceived income tax schedule differs from the actual schedule as some arbitrary function of income. To analyze optimal income taxation, it is necessary to specify further how this difference responds to changes in the tax schedule. One can in principle undertake a familiar form of normative behavioral economic analysis: the planner sets the tax schedule to maximize social welfare, taken to be a function of individuals' experienced utility, but where individuals' choices (the incentive constraints) are determined by their behavioral utility functions, here taken to be standard except for the misperceived tax schedule.[82] The challenge is to determine the perceived income tax schedule as a function of the actual one. Empirical analysis faces the identification problem wherein different behavioral responses at different income levels will embody both differences in underlying utility functions and differences in perceptions. For example, a low measured elasticity may reflect a low underlying ("true") elasticity or a low responsiveness of perceptions (particularly in the relevant estimation window) to changes in the actual schedule. Such empirical challenges and thoughts on how they might sometimes be addressed are considered later in this section.

Before turning to the main strand of modern literature on specific heuristics that may characterize some individuals' misperceptions of their income tax schedules, consider briefly the implications of individuals' misconstruing the income tax schedule for their budget constraints.[83] For example, if there were no withholding, not only may individuals fail to learn about their true net-of-tax wage from their paychecks (on which, more below), but they may also lack the funds to pay their taxes at the end of the year. Similarly, if the EITC were routinely paid each month, those with rising incomes may not only lose eligibility but have to return benefits at a later time after they have spent them. Here, setting aside the effect of misperceptions on labor supply, we can see that there may be significant welfare effects through budget misallocation. Indeed, these concerns motivate withholding—and typical over-withholding in the U.S. as well as delayed EITC payments, both of which involve some forced savings. Most of the literature discussed below makes subtle assumptions so that this problem does not arise, but further research on this front is important. If real forces do induce individuals to effectively use their true income tax schedules for budgeting purposes, perhaps they implicitly learn to use them as well when choosing labor effort. And if these assumptions that solve individuals' budgeting problems do not correspond to actual behavior, the omitted welfare consequences regarding budgeting could significantly alter prescriptions that reflect labor effort alone, which is the focus in the analyses of optimal income taxation that are examined next.

A notable line of literature on behavioral optimal income taxation addresses a particular

[82] From this perspective, one might also wish to account for how setting the schedule and other aspects of tax design may affect the extent of taxpayers' misperceptions (Moore and Slemrod 2020, Craig and Slemrod 2022).
[83] This question is highlighted in Chetty, Looney, and Kroft (2009) and is the subject of Reck (2016), neither focusing on the optimal income tax context (although Reck briefly addresses it).

case of systematic bias: the possibility that individuals behave as if their actual average income tax rate is their marginal income tax rate. Highlights from this research will be presented, followed by further discussion of how individuals may learn about income tax schedules over time and suggestions for reconstruction of prior empirical research to better choose among behavioral theories and identify behavioral responses to income taxation.[84]

Exploration of the possibility that some individuals use their average tax rate as if it were their marginal tax rate begins with de Bartolome (1995). As motivation, he drew on early work by Simon (1978) on procedural rationality, research by behavioral psychologists, and work by economists on electricity demand and firms' pricing decisions, each variously suggesting that some individuals may use average tax rates as marginal ones. He further noted that individuals examining their pay stubs might readily calculate their average but not their marginal tax rate.

To assess this possibility further, de Bartolome (1995) performed an experiment using MBA students, noting that this sample would tend to bias his results in favor of the correct use of marginal tax rates. Among those presented with a format like the tax brackets displayed in instructions for upper-income U.S. taxpayers and in some European countries, a substantial majority made choices reflecting use of the marginal tax rate, with only a few seeming to use the average tax rate. But when presented with a tax table, like that in instructions for low- and middle-income taxpayers in the U.S., almost half exhibited behavior associated with use of the average tax rate and only thirty percent with the marginal tax rate.[85] He noted that such behavioral biases likely exhibit significant heterogeneity. And he also emphasized from the outset a key implication regarding empirical work: if indeed many individuals use average tax rates as their marginal rates, empirical models presuming correct knowledge are misspecified. Relatedly, conventional predictions regarding tax reforms that change marginal rates much more than average rates could be quite misleading.

De Bartolome (1995) also examined the implications of the use of average tax rates as marginal ones for optimal income taxation in light of the key point that most actual, proposed, and optimal income tax schedules feature marginal tax rates that exceed average tax rates throughout the income distribution. In his calculation for a middle-income household, those using their average tax rate supply 5% more labor, generate 6% more tax revenue, and are subject to 43% less distortion than those using their correct marginal tax rate. Although each individual, viewed in isolation, would obviously be better off knowing the true income tax schedule, society is better off if everyone perceives the tax schedule using average tax rates. More revenue can be raised with less distortion, enabling higher levels of social welfare to be achieved. Indeed, if everyone so behaved (and were otherwise identical, except for income-earning ability), a Pareto superior regime would be feasible, although for a given SWF such may not be optimal.

Liebman and Zeckhauser (2004), in a widely cited but unpublished working paper, extend this exploration in a number of ways. Their greatest contribution is to address a range of

[84] The concern that empirical results would be different if individuals used their average tax rates as their marginal tax rates was noted in the survey by Hausman (1985) but not pursued.

[85] This finding led him to conjecture that upper-income taxpayers may behave in accord with their marginal tax rates while lower- and middle-income taxpayers use their average tax rates. This would imply that higher marginal tax rates may be optimal for (only) the latter groups. For a long time, however, probably few taxpayers consult either type of schedule directly, instead relying on tax preparation software and tax planners that may display neither the average nor marginal tax rates, just one of them, or both (although more sophisticated tax planners and software presumably include marginal tax rates). If one conducts a Google search for "What is my tax rate?"—without any modifier like "marginal" or "average"—as this author did in October 2021 when drafting this segment, the top hit (actually, a direct display) showed income tax brackets (marginal rates) rather than a tax table.

applications—nonlinear pricing, utility regulation, and optimal taxation—with a unified approach that contemplates a wide variety of factors that contribute to what they call "schmeduling." They also present two empirical analyses to illustrate the phenomena at hand. Most relevant for present purposes is their explicit analysis of optimal income taxation.

Employing familiar assumptions from prior work (notably, quasilinear preferences and a Pareto distribution at the top), Liebman and Zeckhauser (2004) show that the optimal asymptotic top rate is higher when individuals take their average tax rate to be their marginal tax rate (a phenomenon that they and subsequent authors call "ironing"). This result may seem surprising since the average and marginal rates are essentially equal in the limit, but that overlooks the subtleties of the relevant perturbation experiments. Suppose, for example, that it was contemplated to raise the top marginal rate by 5% (five percentage points) on all incomes starting at $1,000,000. In the standard analysis described in section 2, this causes all taxpayers with income above that point to reduce their labor supply. However, if labor supply is determined by the average tax rate—which is to say that high-income individuals do not react to the reformed tax schedule as such but instead notice only the change in their average tax payments—there is no effect at $1,000,000 and merely a slight effect just above that point. Only in the limit does the perceived marginal tax rate rise by the full 5%. As a consequence, raising the top marginal tax rate causes a smaller reduction in labor effort and thus a smaller fiscal externality (integrating over all individuals subject to the rate increase) when it is misperceived in the posited manner.

Rees-Jones and Taubinsky (2020) offer additional insight into these issues. They go further than the prior literature in incorporating the phenomenon of some taxpayers using their average tax rates as marginal rates, reaffirming that this raises achievable welfare and favors higher marginal tax rates, including at the top of the income distribution (but they caution that their analysis abstracts from possible welfare costs due to mistakes in budgeting). Farhi and Gabaix (2020), although mostly addressing other issues, include a section on optimal income taxation that, like other work, shows that higher marginal tax rates tend to be optimal when individuals underestimate them, such as by taking their average tax rate to be their marginal tax rate. They also emphasize that a lower top marginal tax rate may instead be optimal, such as when the top marginal tax rate is particularly salient and thereby contaminates (revises upwards) others' estimates of their marginal tax rates.

Rees-Jones and Taubinsky's (2020) main emphasis, however, is not on the analytics but instead on obtaining evidence to calibrate the behavioral optimal income taxation model. They undertake a survey and an experiment that are designed to more fully elicit the nature of possible misperceptions. Both of their studies suggest that a significant fraction of individuals use average tax rates, and they also find little support for any notable portion relying on other particular heuristics that they examine. This motivates their simulations in which a fraction of individuals exhibit the bias of taking average tax rates as marginal. As the authors make clear, one should be cautious regarding whether such empirical results accurately describe actual taxpayers' behavior.[86]

[86] As will be discussed in the text below, the manner in which taxpayers learn about how taxes affect disposable income probably differs qualitatively from framings that are readily produced in the lab. Moreover, individuals who volunteer to take computer surveys for modest compensation and their behavior under such conditions may be unrepresentative in important ways. Finally, as the authors explain, in their experiment a third of responses had to be ignored because they were variously implausible. One accordingly wonders what this suggests about others' responses, including that about half of those remaining seemed to ignore tax rates. In addition, the fact that many did not click to obtain further information might be interpreted, on one hand, as participants being happy to use average rates as marginal rates or, on the other hand, as an unwillingness to obtain more data that would only be

Further reflection on both the theoretical and empirical analysis teaches additional lessons and suggests important qualifications. Beginning with the theory and focusing on the phenomenon of individuals' understating their marginal tax rates by using their average tax rates, a number of important subtleties should be noted. (Analogous issues will be elaborated in section 5.2's exploration of myopic labor supply.) Most directly, when individuals misperceive their tax rates, this creates a welfare-relevant internality. In derivations of the first-order condition for optimal marginal tax rates, it is conventional to employ individuals' own first-order condition for labor effort and, in particular, to make use of their envelope condition, which implies that the impact of a tax rate change is given by its direct effect on individuals' utilities, their adjustments of labor effort having only a second-order effect. By contrast, with misperceptions this is no longer true; here, higher tax rates tend to correct individuals' internalities, which is a force for higher marginal tax rates.[87] There are additional effects: the fact that a given income tax schedule raises more revenue reduces the shadow value of funds, a force for lower tax rates; a given income level is now earned (on average) by a lower type (ability) than before, which changes the value of the density function, the cumulative distribution function, and the pertinent wage rate, all of which appear in the first-order condition. Finally, as the literature already emphasizes, the elasticity of labor supply differs. In sum, because essentially every term in the first-order condition for marginal income tax rates changes, the full effect on the optimal income tax schedule (including the intercept) is more difficult to determine than may initially appear.

Before considering further the core empirical questions, consider briefly another particular behavioral phenomenon considered in the literature. In addition to "ironing," Liebman and Zeckhauser (2004) introduce the term "spotlighting," which they defined as treating one's current marginal incentive on a nonlinear schedule as if it applied globally. For example, if one's phone plan provided free minutes or data units each month up to some point and a high marginal charge thereafter, individuals' usage would be as if it were free until they learned (perhaps by a text message) that they had used all of their prepaid allotment and would be subject to charges thereafter. Liebman and Zeckhauser (2004) do not claim that such behavior would arise under a nonlinear income tax schedule. Note, for example, that with the EITC, a middle-income individual acting in this manner would behave in the first couple months of the year as if the marginal rate were significantly negative, during the next months as if it were zero or small, then for a few months as if it were quite high, reflecting the phaseout, and for the rest of the year at a lower rate (unless, later still, moving into a somewhat higher income tax bracket). Such behavior is implausible because, among other reasons, it would require at each point in time knowledge of where one was on the schedule, all the while ignoring the rest of the schedule, repeating this cycle each year.

Rees-Jones and Taubinsky (2020) considers a different notion of spotlighting that might be relevant for income taxation: taking the marginal income tax rate on one's full, annual income to be the applicable marginal rate at all income levels. For local changes in effort or other income-influencing activity, behavior would be essentially the same as under full information. By contrast, for an extensive-margin decision such as joining the labor force full time or quitting, one would be off substantially. Note that this version of spotlighting involves computation and projection that is opposite in spirit to that considered by Liebman and Zeckhauser (2004). Here,

---

useful (and they could not know how useful) if they intended to engage in further mental calculations that might possibly lead to different choices in the experiment.

[87] However, if the welfare weight on individuals' utilities goes to zero as income approaches infinity, this factor has no effect on the optimal asymptotic top marginal tax rate.

one observes one's marginal tax rate and, for the alternative decision imputes the line with that slope and, for example, determines the intercept if one is contemplating quitting one's current job. Under this formulation, single individuals earning $200,000 in the U.S. in 2020 would expect to receive a rebate of nearly $20,000 if they quit. Or part-time workers near the end of the EITC phase-in range might expect a large additional subsidy rather than the reverse if they switched to full time. In Rees-Jones and Taubinsky's (2020) studies, which involve rather different choices, the authors did not find evidence of spotlighting behavior defined in this manner.

Turn now to the empirical question of how individuals are most likely to learn tax information relevant to their choices of labor effort. Although de Bartolome (1995) experimentally investigates and Rees-Jones and Taubinsky (2020) contemplate how different presentations of the tax schedule (for example, tables versus brackets) may influence individuals' ability to absorb pertinent information, one suspects that few individuals consult tax schedules, perform calculations, make projections, or use tax software to analyze counterfactuals when making labor supply decisions. Moreover, however well the use of thoughtful, subtle survey elicitations like that in Rees-Jones and Taubinsky (2020) may illuminate individuals' understanding of the tax system at that moment, it is difficult to know how such elicited opinions relate to actual choices involving labor effort. Some speculations regarding how individuals may learn in context and react (which would be difficult to emulate in online elicitations) are offered next, followed by discussion of how empirical testing of some hypotheses may be possible through reanalysis of existing data on individuals' responses to past tax reforms.

Among the most important labor supply decisions are investments in human capital, many of which are made early in life before the individual is actually working, receiving a paycheck, or completing a tax form (with software or other assistance). Indeed, such individuals may know fairly little even about before-tax compensation as a function of the efforts they now contemplate undertaking. A conjecture is that most have some rough sense of the standards of living associated with different occupations, which may correspond to after-tax income and hence induce decisions that would crudely reflect the actual tax schedule. On the other hand, perhaps many choices are guided by perceived before-tax income, such as when students hear about starting salaries for different jobs, in which case taxes might be ignored altogether—along with the subsequent career earnings profile associated with each occupation. An additional challenge in ascertaining how such behavior may respond to tax changes is that responses may take years or even a generation to materialize. For example, students may take some time to perceive changes in living standards associated with different occupations. Unfortunately, some of these highly consequential impacts of income taxation may be among the most difficult for taxpayers to perceive and react to and, in turn, the hardest for researchers to measure.

For those in the workforce, one suspects that relevant understandings of the income tax depend greatly on income levels and the types of choices under consideration. At very high incomes, the current, proposed, and ultimately implemented reforms to the top income tax rate are highly salient political facts, particularly for high-income taxpayers who will be subject to them. These individuals are also the most likely to receive sophisticated tax advice. Hence, it seems unlikely that they make decisions based instead on their average tax rate.[88] One might

---

[88] Reconsider the above-described perturbation used in the literature to determine the optimal top asymptotic marginal tax rate under the maintained hypothesis that these individuals use their average tax rate as their marginal rate. If the top bracket were, after extensive political debate, raised by 5% starting at $1,000,000, would an individual earning $1,300,000 hear that news (or consult a tax advisor)—and thus react to the 5% increase as such—or instead ignore the tax change entirely until noticing the consequent fall in take-home pay, and then compute that

still suspect that many might underestimate their effective marginal tax rates because they ignore state income taxes and hidden add-ons, or overstate their effective marginal tax rate because they do not take into account deductions or the alternative minimum tax (which often raises the average tax rate but reduces the marginal rate).

For the large portion of taxpayers in the middle of the income distribution, tax rules are simpler but inference may be harder. Intermediate tax brackets are less salient, and sophisticated advice is not likely to be obtained. Again, state income taxes may be ignored but so may deductions. But these may be the wrong questions. Instead we might ask how such individuals decide whether to work overtime or how much effort they should undertake to try to get a larger bonus or a promotion. For these and other decisions, they may consult their paystubs, specifically, their take-home pay that is regularly deposited into their checking accounts. Those net payments reflect not only federal and state income taxes (in manners that reflect imperfections in withholding rules) but also payroll taxes and retirement savings deductions (see section 5.3) as well as other subtractions for various expenditures (Liebman and Zeckhauser 2004). Hence, individuals may behave as if their tax rates are higher than they are.

We can also consider whether this implicit estimation using paystubs involves the use of average or marginal rates. If employees engage in rough mental math to divide their take-home pay by their stated salary, they would identify their average tax rate, albeit one that may exceed their true average tax rate and thus perhaps even their actual marginal rate. It is not clear how a taxpayer would use this information in making real decisions. On the other hand, if a worker earns an additional $100 through overtime and immediately thereafter receives a paycheck that is only $65 higher, the (perhaps overstated) marginal tax rate may be more apparent than the average rate. Likewise when individuals receive end-of-the-year paychecks with bonuses or their first paycheck after a raise. Middle-income individuals may also exchange stories with co-workers, friends, relatives, and neighbors, many being in similar tax situations but others not (because they differ on dimensions that the observer may not even realize affect income tax payments or withholding). That information may involve overtime, bonuses, and raises, and also the effects of changing jobs or entering and leaving the labor force, which would help inform (or misinform) analogous decisions.

The bottom of the income distribution is also important because correct decision-making may be highly consequential and important tax provisions—including relevant transfer programs—can be quite complex due to differing nonlinear (implicit) tax schedules, definitions of dependents, asset tests, and interactions across programs. Some low-income taxpayers benefit from social workers or volunteers who assist in tax preparation and many EITC recipients use paid tax preparers, but most do not receive meaningful tax counseling that would illuminate the overall, phaseout-inclusive marginal effective tax rate. Nor does one suspect that tax tables, eligibility rules, and calculators are their main source of information. Instead, like middle-income taxpayers, these individuals probably rely on personal experience and their networks, the latter suggested by the neighborhood effects identified in Chetty, Friedman, and Saez's (2013) study of the EITC. If one enters the labor force or changes from half- to full-time employment, one will soon see one's take-home pay and may quickly discover, for example, large losses of benefits. If the results are substantially worse than expectations, such decisions may be reversed. By contrast, because EITC refunds do not arrive until much later, they may not be factored into the decisions of many individuals, leading them to understate their net-of-tax wages. Low-income individuals will also learn from relatives, friends, and neighbors who, for example, lost

the average tax rate has only risen, say, only by 1.15%, and react to that as if it were the increase in the marginal tax rate that was front-page news for months?

many transfer program benefits or experienced the lagged receipt of EITC refunds, aiding decision-making on pertinent decision margins even if not producing refined judgments. Note that individuals may learn a great deal about likely impacts on their standard of living without specifically contemplating or being able to communicate the mechanisms by which average or marginal tax rates or other eligibility rules combine to generate the bottom line that they or others experience.[89]

Despite the significant heterogeneity and possible inscrutability of individuals' knowledge and decision-making processes, both across income levels and among those with similar earnings abilities, much empirical illumination may be possible using existing data. As de Bartolome (1995) emphasized, if individuals use their average income tax rates as if they were their marginal rates, prior regressions are misspecified. Regarding his particular point and viewed more broadly, this criticism also points toward important opportunities for further research. For concreteness, this can be illustrated by reconsidering the question whether individuals' use average or marginal income tax rates.

Most straightforward would be to rerun myriad past regressions on labor supply elasticities or taxable income elasticities using average as well as marginal tax rates to determine changes in (perceived) net-of-tax wages. (Important qualifications of how this should properly be done are ignored; the discussion here is merely suggestive.) One could run a standard horserace or attempt to determine what portion of taxpayers behaves as if they are using each tax rate. The results also may vary over time and across contexts (extensive and intensive margins may differ; the EITC and changes in ordinary tax rates may yield different outcomes). Note that in a conventional difference-in-differences analysis, those whose relative marginal rates rise may have relative average rates that change barely at all or even fall; this too would vary greatly across natural experiments. Such variation presents opportunities to disentangle reactions to marginal and average rates by comparing responses across reforms or between different groups that have differential changes in marginal versus average tax rates.[90]

It is also important to consider possible systematic biases that may be present in prior investigations. For example, when those with relatively higher marginal tax rates experience little change in average rates, measured responses would be attenuated if many individuals responded to average rather than marginal rates. In calibrating an optimal income tax exercise in which many use average rates as marginal rates, differently estimated elasticities would need to be used for each group.

Another complexity concerns the time frame. As discussed in section 2.4, it is already known that short-run elasticities may be underestimated for a variety of reasons. In particular, it

---

[89] This perspective may justify the practice of ignoring VATs and sales taxes which (taking the case in which they are uniform) are equivalent to upward shifts in the income tax and transfer schedule. If such taxes are unchanged for a significant period of time, individuals would come to associate various levels of disposable income with real purchasing power that implicitly reflects these taxes. If the rates increase, one might suspect that, over time, individuals would come to associate somewhat lower purchasing power with given levels of disposable income, suggesting in turn that labor effort would react, although perhaps very little in the short run. The tendency for behavior to reflect these taxes does not depend on individuals ever being aware that the taxes exist any more than they need to understand the costs incurred by firms in generating the goods and services that individuals subsequently purchase.

[90] In addition, as illustrated by the earlier example of a 5% increase in the marginal tax rate above $1,000,000, when the marginal tax rate rises in a segment of the income distribution, the average tax rate rises not at all at the beginning of the interval but by increasing amounts at higher levels of income (and also at income levels above the interval, where no increase in marginal tax rates may be experienced). Hence, there is often a significant source of variation that may illuminate reactions to average tax rates that differ from neoclassical predictions that depend on marginal tax rates.

has long been understood that narrow windows necessary for plausible identification suffer from the problem that many taxpayers may not yet have learned of the changes. If only a fraction of individuals are even aware that any change has occurred, measured elasticities may be only a corresponding fraction of actual, long-run elasticities, even when there are no other frictions. It may also be particularly difficult to estimate responses to changes in average tax rates, which often are quite small. Moreover, different learning channels imply different time frames: Individuals may notice and respond almost immediately (perhaps on overtime decisions) to changes in paychecks (supposing that mandated withholding formulas adjust quickly to reforms). But if individuals learn from impacts on relatives, friends, and neighbors who change jobs—and who, after a reform, experience different deltas in their standards of living—reactions could take a long time. The results in Chetty, Friedman, and Saez (2013) on the diffusion of information about the EITC are suggestive of gradual reactions to changes in the program.

Abstracting from the foregoing particulars, taking a more behavioral perspective not just on the design of experiments and the development of new research strategies but also on the analysis of existing data could be very instructive. If there are substantial regularities, such as suggestions that perhaps half of taxpayers behave as if their average tax rates are their marginal tax rates, these could be assessed in many ways. One might look to already-analyzed natural experiments but compare different groups, consider the same groups but use supplemental tests, or exploit other natural experiments that seem particularly relevant for testing new hypotheses even if they offered little predictive power for neoclassical ones. As a thought exercise, consider the hypothesis that many low-income individuals react to average rather than marginal tax rates. Taken literally, this implies that large increases in transfers to those earning no income should raise labor supply at the bottom because such a reform entails significantly lower (more negative) average tax rates even though marginal rates are unaffected—until income entered the phaseout range, where marginal rates are higher even though average rates are lower. This and other (more plausible) conjectures could be tested using existing data on past reforms.

5.2. *Myopic Labor Supply*

Section 5.1 examines implications of individuals misperceiving the income tax schedule. Here, we instead relax a different aspect of Mirrlees' (1971) "rational calculation" assumption: individuals understand everything but have myopic behavioral utility functions wherein they underweight the future relative to the present when making labor supply decisions (Kaur, Kremer, and Mullainathan 2015). Following Lockwood (2020), suppose that individuals exert labor effort one period before they consume the disposable income thereby produced and that they have preferences that exhibit β-δ quasi-hyperbolic discounting. For example, investments in human capital entail effort that significantly precedes consumption. More simply and broadly, labor effort in many settings precedes the receipt of a paycheck. Taking the latter case for concreteness, assume further that individuals promptly consume the proceeds, with the paycheck accurately withholding income taxes (so that taxes are imposed at the same moment as payment and consumption), and consider the implications for the optimal income tax schedule.[91]

Lockwood (2020) focuses on how the formula for optimal marginal income tax rates now reflects an internality correction. When the social planner fully weights individuals' future consumption even though their behavioral utility functions do not, individuals' labor effort

[91] Note that, under appropriate assumptions, we can analogize the case in which individuals underweight subsequent consumption to one like that analyzed in section 5.1 in which individuals overestimate income taxes and hence underestimate the consumption that their labor effort generates.

adjustments in response to marginal tax rate changes no longer satisfy the standard envelope condition regarding effects on individuals' utility. Instead, individuals' reductions in labor effort cause a negative internality on their future selves (with a weight of $1 - \beta$). Because increases in labor effort generate positive internalities, they should be subsidized. The result is easiest to see in a world of fully homogenous individuals (all have the same ability and degree of myopia), so there is no redistributive motive. The optimal income tax schedule—rather than exhibiting zero marginal tax rates throughout—would be negative, with the deficit funded by a uniform lump sum tax. When this phenomenon is instead embedded in the standard optimal income tax problem, internality correction is a force for lower marginal income tax rates (and a smaller grant) than otherwise would be optimal. This result is much the same as in section 4.1's analysis of conventional externalities: if all labor effort conveyed a positive externality, the Pigouvian correction would lower optimal marginal tax rates.[92]

Determination of the full impact of myopia on the optimal income tax schedule, however, requires further and often subtle analysis for the familiar reason—already noted in section 5.1—that most of the factors on the right side of the standard first-order condition for optimal marginal tax rates (expression (8) in section 2.3) are endogenous. Here, in contrast to the case with simple externalities (but like that with misperceptions of income tax rates), we are introducing a change in how individuals choose their labor effort. This modification directly influences many other determinates of optimal income tax rates.

Begin with the income tax schedule that would be optimal in a world without myopia ($\beta = 1$) and consider the effects of increasing myopia (reducing β). Because myopic individuals reduce labor effort, tax revenue falls. If one thinks heuristically in terms of a revenue requirement, tax rates would have to rise to make up the difference, a force in the opposite direction from the internality correction. More precisely, one can think of the deficit as being funded by a uniform lump-sum tax, which in the Mirrlees framework corresponds to a reduction in the magnitude of the uniform grant. At this point, the government's budget balances but we now have a force for greater redistribution because the lowest-income individuals are relatively worse off. Put another way, the required reduction in the grant has decreased the resulting extent of redistribution given the marginal tax rate schedule. (The full tax schedule consists of not just the marginal rates but also the intercept.) Under standard SWFs, this tends to favor higher marginal tax rates the more concave are individuals' utility functions and the SWF.[93]

Additional factors, some cutting against the foregoing, are relevant as well. Raising marginal tax rates tends to produce less revenue at the margin due to myopia. The inframarginal (mechanical) benefit of a higher marginal tax rate at income y is greater revenue collected from all those earning above y. But when all individuals exhibit myopia, every type earns less, favoring lower marginal tax rates. In addition, marginal distortion depends on the density of types at y, which is now different because we are at a different point in the ability distribution. At higher incomes, being at a higher w means that $f(w)$ is lower, so marginal distortion is less, whereas at lower incomes this effect is reversed. In addition, marginal distortion is proportional to w, which is higher at the posited y. Finally, the elasticity of labor supply is a property of individuals' behavioral utility functions, which are now taken to differ. To summarize, essentially every endogenous element in the first-order condition for the optimal marginal tax

[92] Lockwood (2020) explores some variations, including that myopia is plausibly stronger at lower incomes, which generates a relatively larger corrective force at the bottom of the tax schedule.

[93] This tendency is suggested by comparing two of the simulations in Lockwood (2020): they show that moving from low to moderate social welfare weights on low-income individuals lessens the degree to which optimal income tax rates are reduced by myopia.

rate (8) is directly affected by the introduction of myopia, so it is difficult to know a priori how the optimal income tax schedule changes.[94]

To suggest some lines of further research, it is useful to reflect on the timing assumptions in the foregoing analysis, variations across different applications, and relevant policy instruments, including in the administration of income taxation. As will be seen, much can be learned by considering alternative timelines. The foregoing analysis assumes that labor effort is undertaken in a first period while payment, taxation, and consumption all occur simultaneously in a second period. More broadly, these four elements can be arranged in many sequences, including ones in which some of the events are concurrent but not necessarily those combined in the foregoing model.

A key implicit assumption concerns employers' behavior. Suppose that paying employees at the end of the month, or even every couple weeks—which is the practice of most U.S. employers—significantly reduces labor effort at a given wage, as the foregoing analysis assumes. Then it would appear that these employers are leaving a significant amount of money on the table. Changing pay to weekly or even daily (which has little administrative cost in today's world) would perhaps boost effort substantially, a benefit mostly captured by the employers themselves, supposing that they hold wages constant.[95] The government captures the fiscal externality, but the rest benefits employers. Interestingly, Uber and Lyft both currently offer drivers optional forms of essentially instant payment. Hence, in understanding implications of the timing of labor effort and the payment of wages, it seems important to consider what timing structures would emerge in a labor market equilibrium in which some employees may be significantly myopic and to use this information in making inferences about the nature and extent of employees' myopia in different settings.

Another important assumption is that individuals cannot borrow. If they can, reversed cases may arise in which consumption precedes labor supply. When individuals borrow to fund current consumption, paying off their loans with future labor effort, might the impact of myopia on optimal marginal income tax rates then reverse? To elaborate, consider individuals whom we might have imagined live paycheck to paycheck, with each paycheck arriving after their choice of labor effort. But now suppose that they are able to borrow with their credit cards to consume in the present, and they use their next period's earnings to pay these credit card bills.[96] Myopia now leads to overconsumption—a more familiar result in the behavioral economics literature—rather than underconsumption, and this requires additional subsequent labor supply, the disutility of which is excessively discounted by our myopic individuals. Hence, the internality is reversed, creating a force for higher marginal income tax rates. Likewise, revenue is now higher, reducing the need for high tax rates, and all of the additional effects noted above are reversed as well.

In addition to this example with credit card debt, one can imagine longer-term applications, such as when individuals buy more expensive homes, with mortgages and other carrying costs to be paid with future labor effort. Juxtaposing these possibilities with the range of situations in which labor effort precedes consumption decisions (including those considered below, notably, investments in human capital) raises empirical questions concerning which

---

[94] In considering the optimal asymptotic top marginal rate under the assumption that top-income individuals receive a zero marginal social welfare weight (so we are interested in the revenue-maximizing tax rate) and the distribution is Pareto (so the ratio of the inframarginal and marginal effects is constant), all of these features—including the internality correction—vanish, except for any effect on the labor supply elasticity.

[95] Short periods having large effects are suggested by the experiment with data entry workers in India by Kaur, Kremer, and Mullainathan (2015). Employees may resist, however, preferring to contract for deferred pay as a commitment device to mitigate overconsumption.

[96] Payday lending raises similar questions.

formulations best match actual behavior. It seems plausible that, even at a given income level, there may be substantial heterogeneity, not only in the magnitude of myopia but in the relevant timing of decisions and hence the direction of any desired adjustments to marginal income tax rates. This heterogeneity also reinforces the potential benefits of policies targeted at specific settings in which significant myopia may be present rather than relying primarily on adjustments to the income tax schedule to correct for weighted-average myopia on a wholesale basis.

Returning to our original setting in which labor effort precedes consumption, consider the range of situations in which this timing may be important. In addition to paychecks, there are myriad other decisions, particularly involving investments in human capital, where effort or other investment significantly precedes returns in the form of greater consumption. Indeed, because of the much longer durations, myopia may have a greater effect on these margins.

For human capital investments that precede employment, individuals exert effort and forgo earnings resulting in additional deferment of consumption in order to raise future wages to enable work in occupations associated with less disutility from labor effort. It is natural to consider whether the internality due to myopia in these settings is best addressed by adjustments to far-future tax rates or by more targeted policies that might offer direct, present subsidies for human capital investments.[97] Such policies are often motivated to address liquidity constraints, but unconstrained myopic individuals may similarly underinvest. Likewise, offering loans has the additional effect of moving payments to the future, which matches when income is subsequently earned, reducing the need to defer present consumption in order to make human capital investments. Myopia may also lead some individuals to underinvest in human capital in earlier years, such as a failure to study in secondary school or a decision to drop out to begin earning and consuming sooner. Again, more targeted policies seem more promising than lower future marginal income tax rates.

Important human capital investments are also made on the job, and these are often rewarded by deferred compensation, promotions, and the prospect of higher subsequent wages or bonuses. Here, the aforementioned discussion of employers' timing of paychecks is apt, with potentially greater consequences. If employees' myopia leads them to underinvest, particularly in firm-specific human capital, because rewards are deferred, employers have incentives to restructure payments to address this problem. Signing bonuses may in part involve inducing myopic individuals to accept jobs involving significant subsequent effort. Likewise, offering lucrative compensation to new hires who initially have low marginal products may reflect that more precise matching of pay to productivity would fail to attract myopic but talented employees. Of course, employers also face an important constraint because they are generally unable to recoup payments from employees who fail to deliver on implicit or explicit promises of future effort. To the extent that deferred pay is the only instrument, we may then return to the question of how optimally to adjust income tax rates in light of employees' myopia.[98]

Turning to tax administration, the rules most relevant to individuals' myopia may well be those governing income tax withholding and refunds because they directly determine the timing

[97] This focus also suggests consideration of more fine-tuned interventions, for example, targeting individuals' choices of what to study in light of the fact that different concentrations involve different tradeoffs between present consumption and effort associated with the study itself and future earnings associated with different courses of study.

[98] The present discussion of how best to design income taxation in the presence of the myopia imperfection that is partly addressed by private contracting has some elements in common with that in section 3.3 of optimal income taxation when private contracting can only partially address moral hazard and asymmetric information. In both settings, more sophisticated modeling and evidence of market arrangements are important inputs to understanding how best to adjust income taxation to address the same imperfections.

of tax payments. Such rules may influence both consumption and labor effort. If income taxes were only paid after the end of the year, myopic individuals might overconsume from their earnings, creating welfare losses from budget misallocations. Present formulas in the United States entail overwithholding for many individuals, and the EITC even more powerfully provides significant, delayed income tax refunds. Both entail forced savings, which may aid myopic individuals. However, the analysis here suggests that, by significantly deferring some of the consumption associated with current labor effort, the tax system may be generating a large internality with respect to individuals' labor effort decisions. Because these rules directly target the timing of tax payments and refunds relative to the timing of labor supply, their design seems likely to be particularly important in addressing various consequences of individuals' myopia.

5.3. *Behavioral Savings, Savings Policies, and Labor Supply*

Perhaps the most significant application of behavioral economics research has been to individuals' savings decisions and the design of savings policies (Bernheim 2002, Thaler and Benartzi 2004, Bernheim and Rangel 2007, Bernheim and Taubinsky 2018). Central features of fiscal systems in advanced economies—notably, social insurance, but also retirement savings incentives and other features of capital income taxation—are substantially motivated by the view that many individuals, in making their own choices, suboptimally and, in particular, inadequately save for their retirement.[99] Yet, most theoretical work on optimal income taxation, empirical research on relevant elasticities, and calibrated simulations assume that these decisions are neoclassical. Nor do assessments of savings policies aimed at behavioral infirmities consider their implications for labor effort.

A central motivation for exploring this intersection—beyond the a priori conceptual interest in integrating these lines of work—is that labor supply effects have first-order welfare consequences in the presence of income taxation, which is employed in the same developed economies that engage in substantial social insurance for retirement that is funded by payroll taxes, which is to say, labor income taxes. As a simple motivation, suppose that, for the bottom and middle of the income distribution, payroll and income taxes are of similar magnitude. Moreover, payroll taxes fund retirement savings that it is feared myopic individuals would not otherwise undertake. Assume further that myopia (or misperception) is such that the present taxes are taken fully into account in individuals' labor supply decisions but the future benefits are ignored entirely. Using the rule of thumb that distortion rises with the square of the tax rate, might the total distortion be four times that of the nominally described "income tax" alone? And what are the implications for optimal income taxation and the optimal design of social insurance and other savings policies?

The focus of work to date has been on how labor supply effects in the presence of income taxation feeds back on optimal savings policies, which is where the discussion here begins. Next, drawing on the lessons of the preceding two subsections, consequences of behavioral savings and corrective savings policies for the design of the optimal income tax schedule are examined. Finally, this section considers implications for the interpretation of empirical work and calibrated simulations and, relatedly, identifies questions for further research in light of the

[99] The analysis here focuses on myopia and other phenomena that involve internalities. Another motivation for encouraging savings involves the Samaritan's dilemma, in which individuals save too little in the expectation that others (relatives or the government) will come to their aid (Buchanan 1975). In this case, raising savings reduces a negative externality rather than correcting an internality. Kaplow (2008a) examines the effects of forced savings to combat the Samaritan's dilemma in a model with no behavioral infirmities.

fact that past analyses assume that the parameters identified and imputed are determined by neoclassical behavior.

Kaplow (2008a, 2011, 2015a, 2015b) considers how savings policies influence labor supply under a variety of assumptions about savings decisions and how those decisions feed back on individuals' choices of labor effort.[100] Consider the familiar two-period model in which individuals supply labor only in the first period (their working years) and divide their disposable income between consumption in that period and savings that funds consumption in the second period (their retirement years).

Suppose initially that all individuals exhibit identical β-δ quasi-hyperbolic discounting in allocating their disposable income between the two periods, so they overconsume in the present, saving too little for retirement (Laibson 1996, 1997).[101] For concreteness, assume further that the government is setting the stringency of a forced-savings requirement, which is one way to understand an actuarially fair social insurance program funded by payroll taxes and providing benefits in retirement.[102] (Analysis of capital income subsidies and other retirement policies would be similar, which as we will see carries implications for the optimal taxation of capital income that are quite different from those in the literature.[103]) As this requirement begins to bind, individuals' experienced utility—taken as the maximand for social welfare—rises, and it continues to do so (at a declining rate) until optimal savings are required. This standard depiction, however, is only correct if labor supply is exogenous when we are in a world with an income tax and hence fiscal externalities from changes in labor effort.

To examine labor effort explicitly requires a further assumption regarding how myopic savers make their labor supply decisions. Two cases will be considered.[104] First, suppose that the same myopic behavioral utility function that determines consumption and savings also governs labor supply. (Consider, for example, a decision whether to work overtime on some day when the alternative is an evening with family and friends or some favorite diversion.) Here, as the forced-savings constraint begins to bind, there is no first-order effect on individuals'

---

[100] In each of the cases that follow, individuals' labor supply decisions also affect their consumption in the present and future and accordingly their realized utility. The discussion in the text focuses on the fiscal externality from the identified effects on labor effort, but a full welfare analysis—including the subsequent discussion of optimal income taxation in light of behavioral savings—needs to incorporate these effects as well.

[101] Quasi-hyperbolic discounting need not generate overconsumption if savings plans are structured to become effective in the future ("saving more tomorrow") (Thaler and Benartzi 2004). That formulation would nevertheless raise some of the questions regarding labor supply that are explored here.

[102] Additional issues, set to the side here, would arise to the extent that individuals suffer from a further informational problem of being unable to accurately appreciate the tax-benefit linkage.

[103] Optimal capital income taxation in connection with labor income taxation has been a subject of extensive study, largely using models in which individuals are fully rational. In many of these settings that feature uncertainty about future earnings, individuals tend to *over*save from a social perspective because their levels of precautionary savings ignore resulting future fiscal externalities from their consequently reduced labor effort. By contrast, much government policy toward savings, including the income taxation thereof, is predicated on the opposite view and hence is subject to the present analysis. The large differences in optimal capital income taxation relate not only to the differences in savings as such but also, the focus here, in qualitatively different implications for how capital income taxation affects labor effort. Both aspects suggest a new research agenda for the optimal taxation of capital income, which is not the focus here.

[104] Under another variation of sophistication, wherein individuals are not myopic when choosing labor supply but fail to foresee their own myopia in making savings decisions, a binding forced-savings requirement will have no effect on labor supply (assuming that the constraint is not so tight as to require more savings than would be chosen by a nonmyopic, optimizing individual) for the simple reason that such individuals would not expect the forced-saving requirement to be binding on themselves (it requires less savings than they anticipate choosing on their own account).

behavioral utility levels, but thereafter in standard cases the marginal utility of after-tax income falls because individuals cannot allocate that income as they (think that they) prefer.[105] When there is an income tax, this generates a fiscal externality equal to the labor supply effect times the marginal income tax rate, a factor ordinarily ignored when assessing savings policies. Therefore, the optimal degree of forced savings in this model is likely to fall significantly short of the ideal savings target. Even so, we can see that a payroll tax funding retirement savings is quite different from just an additional tax on labor effort, as imagined in this section's opening example: it is only relevant when it is binding and, even when it is, its negative impact is mitigated by the fact that even myopic individuals value retirement consumption in their behavioral utility functions, albeit less than fully.

Second, assume that our myopic savers are sophisticated in their labor supply decisions. Perhaps they choose investments in human capital and accept jobs (with fixed hours) with the recognition that, when they receive their paychecks, they will save insufficiently. Now, when the forced-savings constraint begins to bind, individuals appreciate the first-order utility gain; that is, they value the government's imposition of a commitment device. Hence, in standard cases, they have a higher marginal utility of after-tax income, which encourages labor effort and thereby generates a positive fiscal externality. By contrast to the first case, there is a first-order effect the moment the constraint begins to bind, and the magnitude of the effect falls, reaching zero at the ideal level of savings. Under this assumption, the effects have the opposite sign of imposing a tax on labor income in the presence of a pre-existing tax and have properties that otherwise differ qualitatively (notably, the effects fall as the stringency of the constraint and hence the tax rate rise).

The foregoing analysis contemplates myopic individuals who otherwise fully calculate relevant effects. But in the lifetime savings context, it also seems plausible that behavioral infirmities more like those examined in section 5.1 might also be operative (Bernheim 1994; Johnson, Kotlikoff, and Samuelson 2001; Diamond 2004). The complexity of the problem may give rise to errors, and the fact that important effects are in the future may inhibit learning. (Individuals only live once, and learning upon retirement that savings are too low, or too high, does not enable much correction.) As with complex income tax systems, this raises the question of whether there are systematic tendencies in individuals' behavior and what are their implications for labor supply.

Suppose now that there is no myopia, that individuals cannot figure out how optimally to save, and that they accordingly engage in what may be called targeted savings decisions. Perhaps individuals simply stick with whatever degree of savings results from social security and employers' retirement contributions. Perhaps employer defaults on employee contributions simply stick (Beshears et al. 2008). Or perhaps individuals follow advice from family, friends, or financial advisors, not understanding the calculations that may (or may not) underlie these recommendations. Such an assumption answers our question of how much individuals save, but it does not indicate how they make their labor supply decisions. Moreover, this understanding of individuals' savings decisions is inconsistent with standard neoclassical assumptions so, as with myopia, it is necessary to contemplate other formulations. Consider three possibilities.

First, suppose that individuals when making their labor supply decisions treat savings as

[105] This point has an important subtlety. The direct effect is as stated in the text. However, there is an opposite-signed indirect effect: forcing an individual's consumption into the future raises the marginal utility of present consumption (which encourages labor effort) and reduces the marginal utility of future consumption (which reduces labor effort). For standard utility functions with upward-sloping labor supply curves, the net of this indirect effect (which is to raise labor supply) tends to be smaller than the direct effect, in which case labor supply falls as stated.

if it vanishes. Perhaps they look at their disposable income as reflected in their paychecks and take that as their return to labor effort. In this case, forced savings is akin to a tax in the presence of a preexisting tax, with the reduction in labor effort generating the associated fiscal externality. Optimal savings levels would then be much lower than ideal targets.

Second, assume that individuals realize that their savings does not vanish but, by assumption, they have difficulty knowing how to value this savings. Suppose further that they address this conundrum by treating their savings as having a value akin to if they had consumed it currently. In this case, changing the savings target has no effect on the perceived marginal utility of consumption, so labor supply decisions would be unaffected.

Third, suppose that individuals make a more sophisticated imputation. Specifically, although they cannot themselves determine the optimal level of savings, they assume (behave as if) the target-setter solved this problem for them, at least approximately. The government, their employer, or their advisor is taken to have acted paternalistically, in their own best interest.[106] The implication is that individuals take the marginal utility of any increment to their savings to equal the marginal utility of their current consumption, which is precisely what the relationship would be if they had in fact determined their optimal allocations between present and future consumption and saved accordingly. In this case, when the target is set higher, current consumption is mechanically lower, which raises the marginal utility of current consumption and, along with it, the perceived marginal utility of savings. Hence, a stronger forced savings requirement raises labor effort, producing a positive fiscal externality, thereby favoring a higher savings level than the otherwise-ideal target.[107]

Considering together the two cases with myopia and the three cases with targeted savings—which undoubtedly do not exhaust the possibilities—we can see that there are a variety of possible effects on labor effort, both positive and negative, and that these may well be first-order considerations in setting optimal savings policies in light of the large income tax wedge that implies significant fiscal externalities. More subtle analyses of savings policies are required in light of the fact that behavioral infirmities are undoubtedly heterogeneous. For example, if only some individuals are myopic, forced savings may be preferred to capital subsidies because both boost savings of the myopic but the former avoids inducing excessive savings by those who are not myopic (Kaplow 2008a, 2015a, 2015b; Farhi and Gabaix 2020). In addition, similar analysis can be applied to other forms of social and private insurance, notably health, unemployment, disability, and life insurance—all of which individuals with behavioral infirmities may under- (or over-) consume (abstracting from adverse selection).[108] Many of these are provided in ways similar to social insurance for retirement, funded by payroll taxation or payroll deductions by employers; hence, they too may have labor supply effects that have first-order welfare implications on account of fiscal externalities. These observations suggest an even broader research agenda at the intersection of behavioral infirmities, social insurance,

[106] Although much analysis of employer retirement policies, as influenced by government regulation thereof, seems predicated on employers being motivated to maximize employees' interests, taking employers instead to be profit-maximizers who hire employees in labor markets may well suggest otherwise, with employers catering to rather than correcting employees' biases (Bubb and Warren 2020).

[107] Put another way, when the paternalistic target-setter chooses a higher savings target, the individual gets the message that consumption in the future is regarded to be more valuable. Therefore, after the resulting reallocation, incremental disposable income—consumed some in the present and some in the future—is perceived to be more valuable overall.

[108] Each of these forms of insurance, like retirement savings, involves (in the simple, first-best case) equating marginal utilities of consumption across time or states; each likewise involves present payments that fund future benefits; and each involves complexity, often involving uncertainty and related behavioral infirmities.

corrective policies, and income taxation.

The foregoing analysis focuses on how behavioral savings and corrective savings policies affect labor supply. Turn now to the question that is the focus in the earlier sections: how do these behavioral infirmities affect the optimal income tax schedule? Most of the issues already considered arise here as well, although sometimes in a different form. To begin, note that the social welfare weight on a dollar to an individual now differs. First, at a given disposable income, marginal (and, for a strictly concave SWF, total) utility differs as a consequence of different labor supply and misallocation between current and future consumption. Second, the marginal dollar itself may be misallocated or allocated in light of a constraint that differs from the standard first-order condition. These factors depend on the behavioral assumption about savings, the behavioral assumption about labor supply in light of how savings is determined, and the prevailing policy. Hence, the determination of welfare weights is qualitatively different and notably more complex in this setting.

The mechanical and incentive effects of adjusting marginal tax rates differ as well. Like in the preceding sections, it will generally be true that the type (ability) associated with a given level of before-tax income will differ. For example, if the interactions among behavioral savings, labor supply decisions, and the prevailing savings policy lead individuals to work less, then individuals earning a given income will be of a higher type. Accordingly, raising the marginal tax rate at that income level collects less revenue from higher-earning types because there are fewer of them. And the marginal distortion depends on the density function at that income, which in this instance would be higher at the lower end of the income distribution and lower at the higher end. It also depends on the wage itself, which is higher. Finally, the marginal distortion depends on the labor supply elasticity, which also will differ.

Taken together, we can see that each combination of behavioral assumptions and prevailing savings policies will be associated with different optimal income tax schedules. Relatedly, changes in savings policies in general change what income tax is optimal in a manner that depends on which set of assumptions about savings and labor supply is applicable.[109] Note further that these behavioral assumptions and savings policies have important implications for the interpretation of empirical work and the calibration of simulations because both types of work as currently conducted usually involve the measurement and imputation of parameters based on the assumption that the data was generated by neoclassical behavior.

As a consequence, an important part of a broadened research agenda is empirical. Results are qualitatively different if observed savings are generated neoclassically, myopically, or by targeting as a satisficing reaction to complexity and uncertainty. Regarding the latter possibilities, labor supply decisions may be determined in qualitatively different ways that generate labor supply effects of different signs and with different comparative statics. Moreover, when one adds asymmetries (for example, a tightened savings constraint or target may be binding at the margin for some but not others) as well as the fact that many labor responses exhibit nonlinearities, we can see that the undoubted presence of significant heterogeneity across individuals—and even within individuals but across decisions (choice of job versus overtime decisions)—means that average responses at a given income level are not sufficient statistics for welfare analysis.[110] Research on this front is challenging because many natural experiments

[109] For a preliminary exploration of some of these complex interactions in a model with myopia, see Moser and Silva (2019).

[110] Epper et al. (2020) show that significant heterogeneity in time discounting explains much of the observed heterogeneity in wealth inequality. Although participants' responses to different time frames in preference elicitation experiments are argued to show that the documented preference heterogeneity involves true discount rates

generated by reforms do not directly bear on these margins, although some do. In addition, some of these theories suggest that changes in employers' behavior regarding retirement savings may have labor supply effects, which opens new channels of investigation.

Another promising avenue involves reanalysis using existing data. Notably, a variety of calibrated simulations implicitly take strong stands on many of these questions, but variations could be explored. Most obvious is the decision whether to treat payroll taxes that fund social insurance purely as taxes on labor income (see Fullerton's 1991 survey of estimates of the marginal welfare cost of taxation), to include them only to the extent of gaps in tax-benefit linkage (which may be much smaller or even of opposite sign for many individuals), or to treat the tax and benefit pieces separately in ways suggested by some of the foregoing analysis. For example, one could examine the extent to which different formulations of individuals' behavioral utility functions for labor supply regarding retirement savings affect the analysis. Recall in particular the theoretical predictions of how different assumptions can even reverse the sign of the labor supply effects of these taxes that are a very large part of what is conventionally taken to be the labor wedge for all but high-earning individuals.

Finally, as suggested at the outset, such exploration could also be highly consequential for analyses of optimal capital income taxation. It is notable that the extensive modern literature in that field has yet to be integrated with the substantial behavioral economics research on individuals' savings behavior.

6. *Optimal Income Transfers*

Optimal income taxation at the lower end of the income distribution is very consequential given high marginal utilities of consumption and possibly additional weight if the SWF is strictly concave. Indeed, in the Mirrlees framework, the central objective is to redistribute toward the bottom. Although optimal marginal tax rates at the tip-top of the income distribution have received disproportionate attention from the beginning, there also is important work focusing on the bottom.

This subject encompasses assessments of optimal income transfer programs. In many fiscal systems there are many of these, most being separate from the officially designated income tax, but some are incorporated within it, such as the EITC in the United States. In the Mirrlees framework, however, all of these essentially separate income tax schedules—many consisting of a grant and a phaseout schedule—are summed together and considered in a unified manner. One consequence is that there is no such thing as an "optimal EITC," for only the aggregate grant and composite marginal tax rates matter. Moreover, as Mirrlees (1971) stated, one cannot focus on transfer programs in isolation of the entire income tax schedule, not only because the grant must be funded but also we know from the standard first-order condition (8) that a key feature of the optimal marginal tax rate at any income level is the revenue it collects from all those earning higher levels of income.[111]

A natural question to ask about the optimal design of income transfers for the poor is why we are asking this question at all. Why not just look toward the left end of the optimal nonlinear income tax schedule, derived in the usual way? Whatever we see constitutes the optimal income

---

rather than present bias, the fact that the measured average annual discount rates fall in the range of 39% to 51% suggests that much of the variation may be attributable to differential myopia.

[111] An interesting aspect of the historical evolution of the broader field is that, as Mirrlees (1971) notes, many of the central ideas about optimal income taxation were first advanced in Diamond's (1968) review of Green's (1967) book on the negative income tax.

transfer scheme. What we typically do observe in a range of simulations are a substantial grant, fairly high marginal tax rates at the bottom (often falling thereafter), and the lowest-ability individuals not entering the labor force. The high marginal rates near the bottom reflect the modest revenue loss from marginal distortions because those distorted have low productivity and the large inframarginal revenue gain because most of the population earns higher incomes. Regarding the latter, suppose that one contemplates lowering the marginal tax rate by 10% (p.p.) on all incomes from $0 to $10,000, which has an inframarginal revenue cost of $1000 for each individual earning income above $10,000. If there were 100 million such individuals, the revenue cost of this modest boost in work incentives at the bottom would be $100 billion per year. Reducing, say, a 60% marginal tax rate that one sees in some simulations to the negative 40% of the EITC phase-in, in order to have a truly large impact on work incentives at the bottom, would cost $1 trillion per year on this oversimplified accounting.

The analysis in this section takes the standard Mirrlees framework as the starting point and considers a number of factors that may influence the results (Kaplow 2007a, 2008a). Most proposed modifications are applicable at all income levels and in any event affect the entire tax schedule, yet they are often framed as addressing optimal transfers because they are thought to be particularly important at the bottom of the income distribution. Section 6.1 examines categorical assistance, which is most familiar when more generous treatment is offered to families with young children, individuals with disabilities, and the elderly. These traits are, of course, present throughout the income distribution, and in any event we must consider the entire schedule and indeed the entire system. The natural extension is to allow there to be separate income tax schedules for different groups, with the applicable schedule determined by some signal of individuals' types. The schedules are related by a common shadow value of government funds, with the implication that deficits in one schedule (for a needier group) can be funded by surpluses from another. Section 6.2 considers two additional directions that have received significant attention: forms of workfare and extensive margin responses.[112]

6.1. *Separate Income Tax Schedules*

Many transfer programs are categorical, providing more generous treatment for the disabled, the elderly, or families with children, often depending on their numbers and ages. To analyze differential treatment, one can state the planner's problem as optimizing a multiplicity of nonlinear income tax and transfer schedules that are linked by a common shadow value of funds. See, for example, the depictions in Werning (2007) and Kaplow (2007a, 2008a, 2008b) and the application in Blundell and Shephard (2012). This approach constitutes a generalization of Akerlof (1978) in the Mirrleesian setting and is yet another extension originally suggested in Mirrlees (1971).

In this formulation, there may be two or more discrete schedules or a continuum of schedules. These schedules may optimally differ both because of differences in the underlying distribution of ability and because of differences in need (really, in the pertinent utility

---

[112] A range of topics—some addressed in other sections of this article—are omitted here, including: the effect of transfers and work inducements on wages (3.2); externalities (4.1), which here might be associated with children; internalities (5), which may be relatively greater at the bottom; in-kind provision, which may be motivated by the presence of externalities, internalities, and as tagging (free medical care is used more by those with unobservably high medical needs); two-earner families (8.2), in light of marriage penalties often being high at the bottom given the design of many transfer programs; human capital, where free public education is particularly significant at the bottom; minimum wages; and the criminal justice system, because crime is an externality of low income and punishment affects future earning ability.

functions).[113] The applicable schedule, $T(y, \theta)$—with its own grant ($-T(0, \theta)$) and marginal tax rates—is determined by a signal (or signal vector) θ.[114] This formalization applies whether the categories perfectly or imperfectly indicate underlying differences in ability or need and whether the signal for the categories itself is perfect or noisy. Nevertheless, for ease of exposition and conceptual clarity, the discussion here will rule out (deferring until section 8.2) any residual heterogeneity in utility functions within types having the same signal θ.

The analysis here will take θ to be exogenous. For some important categories in use—notably involving disability, children, and marriage—this assumption importantly omits moral hazard and fairly deliberate life choices. If one introduced endogenous categorization, then changes in the grant (intercept) or in marginal tax rates of any group's schedule would add another term to the optimization, the integral of social welfare differentials weighted by pertinent elasticities. The former factor is not only complex but normatively contentious when it involves differences in the number of children.

Using this formulation, we can restate the first-order condition (8) for these income tax schedules as follows:

$$(9) \qquad \frac{T'(wl(w,\theta),\theta)}{1-T'(wl(w,\theta),\theta)} = \frac{1-F(w,\theta)}{\xi(w,\theta)wf(w,\theta)} \frac{\int_w^\infty \left(1-\frac{W'(u(\omega,\theta))u_c(\omega,\theta)}{\lambda}\right) f(\omega,\theta)d\omega}{1-F(w,\theta)}.$$

Two observations are in order. First, in the standard formulation with a single income tax schedule $T(y)$, the grant—the value of $-T(0)$—is not separately stated because it is implied by knowledge of the full schedule and any revenue requirement. Here this is not the case. In general, it will be optimal to have different grants for each schedule, as will be discussed.

Second, these schedules are linked by the common shadow value of revenue, λ. These two observations are related in that, for example, a higher grant for those of a type θ having low abilities or high needs compared to other groups can and often would optimally be financed by higher taxes on the other groups. This point is implicit in restating the tax instrument as $T(y, \theta)$ but is worth emphasizing. Among other things, these observations explain why the widely used concept of phaseouts is incoherent, not only because all transfer programs are integrated but also because, as just emphasized, it is incorrect to think of any group in isolation. Of course, one always can, in an accounting sense, deem all benefits to be phased out for any group upon reaching the level of income $y(\theta)$ such that $T(y(\theta), \theta) = 0$. But this in no way suggests any sort of phaseout target. Instead, that break-even point will emerge implicitly as a result of an optimization involving all θ, which determines each group's grant (intercept) and schedule of marginal tax rates.

One can also see that there is no simple relationship across groups' optimal income tax

[113] As explored in section 8.2 on heterogeneity, some cases overlap. For example, a "disability" might be a lower "ability" (equivalent to a lower w), or a greater disutility of effort (u_l having a greater negative magnitude). Certain functional forms of the latter are equivalent to the former, but regardless of the formulation, using different tax schedules that depend on a signal of the disability raises achievable social welfare.

[114] The signal technology and its calibration are taken to be exogenous. The formulation below would in principle allow one to derive the value of improved accuracy in a signal technology (such as refinements to disability assessments). The standard of proof is different. Often dichotomous methods are used (an individual is deemed disabled if the signal exceeds a threshold), whereas continuous implementations are generally optimal, abstracting from administrative costs. Hence, we can take θ as the raw signal rather than a dichotomous indicator determined by some rule applied to the raw signal.

schedules. For example, it is commonly suggested that if some identifiable group has a higher need, that group should optimally be compensated in that amount. Here, that would correspond to a higher grant but an otherwise identical schedule. But because, in general, θ will indicate different distributions of ability and different utility functions (which are the underlying source of different "needs"), and those in turn imply that different types will earn a given income level y across the groups, the entire schedules will be different and the grant differences will not in general equal any difference in need, which will not typically be uniform across the income distribution in any event.[115]

Consider now more systematically some of the determinants of how income tax schedules should differ across groups. For concreteness, the exposition will refer to a high group, H, taken to have systematically higher income-earning ability than another, low group, L. The discussion will be heuristic, suggesting implications for optimal grants, g_H and g_L, and for the optimal marginal income tax rates in $T(w,H)$ and $T(w,L)$ making use of the first-order condition, a fraught exercise due to the endogeneity of many of the parameters in that expression, a number of which will be elaborated here.

We are assuming here that the only difference between the groups is in the densities of abilities, $f(w,H)$ and $f(w,L)$, and their corresponding cumulative distribution functions, $F(w,H)$ and $F(w,L)$. Suppose that the L group has some observable disability, such that the density is particularly concentrated at the bottom and thin at the top relative to the unconditional density, implying that the H group has the opposite characterization.[116] The most obvious implication is that the optimal income tax system will tend to have $g_L > g_H$ (with the optimal grant, g, under a uniform schedule being at an intermediate level). Those in the L group on average have higher marginal utilities of consumption and lower levels of utility, so redistributing toward them tends to raise social welfare. Moreover, reductions in labor effort due to income effects of the higher grant will tend to be less costly in terms of lost revenue (the fiscal externality) because this group is less productive.

Regarding optimal marginal tax rates, consider initially the bottom of the income distribution, which is where analysis on transfer program design focuses. Because of the leftward shifted density function, $f(w,L)$, the marginal distortion from higher marginal income tax rates on the L group will be larger because those rates apply to relatively more individuals, by comparison to the H group or to when there is a common income tax schedule. Moreover, the inframarginal revenue benefit from higher marginal rates is smaller both because there are relatively fewer inframarginal individuals earning higher incomes and because those individuals will tend to be concentrated at lower levels of income and hence have higher welfare weights. Both effects favor lower marginal tax rates in the L group at the bottom. Per the above warning about interpreting the first-order condition, however, there are further adjustments. Notably, although the inframarginal individuals in the L group will also be worse off before redistribution, they may, at the optimum, be better off if g_L is higher by a sufficient amount.

Taken together, it seems plausible that the L group should optimally receive a larger

[115] As explored in Kaplow (2008b) and noted briefly in section 8.2, there is a special case in which a simple, compensatory grant adjustment would be optimal: when the only difference between the groups is in their utility functions and moreover the difference involves subtraction from consumption of a common constant in one of the groups. For example, if everyone in one group needs to spend an additional \$100 per year on eyeglasses, at which point their utility for a given level of other consumption and labor effort is identical to that of individuals in the other group, and moreover there are no differences in the distribution of abilities between the groups, then a grant differential of \$100 would be the only difference between the two groups' optimal income tax schedules.

[116] No attempt is made to be precise. As is familiar, subtle differences in the shapes of the density function can have important implications for the optimal schedule of marginal tax rates.

grant—which by conventional thinking would call for a more aggressive phaseout—and yet also be subject to lower marginal tax rates, just the opposite. As explained previously, we can think of the higher g_L—and now lower marginal tax rates at the bottom as well—as being financed by higher taxes (a lower g_H and higher marginal tax rates) on the H group.[117] This point also casts the introduction's provocative illustration of the cost of reducing marginal tax rates at the bottom of the income distribution in a different light. Here it is contemplated that there would be lower marginal tax rates not on the entire population but only in the L group, so the cost is lower in absolute terms due to this limitation and may be much lower if the targeted group is a small portion of the population. Moreover, because within this group there are relatively fewer individuals at higher levels of earnings for whom the marginal rate reductions are inframarginal, there is less revenue loss on that account as well. Indeed, this was a key part of the explanation for why optimal marginal tax rates may be lower at the bottom for the L group. By contrast, in the H group, where relatively more are inframarginal, even higher marginal income tax rates at the bottom tend to be optimal. Relatedly, when calibrating optimal income tax simulations, particularly for the purpose of assessing particular transfer programs, it is important to take into account whether eligibility for some of them is limited to certain groups. Average transfers and average phase-in and phase-out rates across the population can be a highly misleading guide to the optimal design of the grants and marginal tax rates for each group.

Consider next another manner in which our two groups may differ. Set any differences in ability to the side and assume instead that our two groups, H and L, differ only in "needs," which here means in the applicable utility functions.[118] As mentioned, the L group will be taken to have a lower level of utility for a given ability w and level of labor effort l. As explored further in section 8.2 and Kaplow (2008b), such differences can arise in a number of ways with qualitatively different implications for optimal income taxation.

Suppose initially that these individuals' marginal utilities of consumption are higher. (The case in which their marginal utilities are lower will be qualitatively the same, with signs reversed.) The direct effect is that redistribution from group H to group L will be favorable. This might be implemented by raising g_L relative to g_H or by lowering the relative marginal income tax rates imposed on group L. The case for the former is intuitive. For the latter, note that raising any marginal income tax rate at a given level of income, ceteris paribus, generates revenue from the higher-income inframarginal types. When they are taken to have higher marginal utilities of consumption, the benefit from that will be less—so long as their grants are not raised enough to erase or reverse this across-group difference in marginal utilities. Additional effects arise if the SWF is strictly concave. In the current setting, the effects will be reinforcing because those with higher marginal utilities also have lower total utility. (But in the reverse case in which the marginal utilities of consumption are lower in the group with lower utility levels, there would be opposed effects.)

As usual, however, all else will not be equal. A higher marginal utility of consumption itself will encourage labor effort, so on that account the type w earning a given income y will be lower in group L, so the values of the density and cumulative distribution functions will differ

[117] Differences at the upper end of the two income tax schedules are less clear. Notably, the L group will have both a lower density in the denominator and also a lower inframarginal mass in the numerator of the first-order condition, so whether this ratio is higher or lower will depend on differences in the shape of the two groups' density functions toward the top. And, as mentioned in the text, there are other differences as well, notably in the size of the grants, although at sufficiently high incomes this would be of little consequence.

[118] The restated first-order condition (9) does not allow the social welfare function to depend *directly* on θ; if it did, further implications would be apparent.

and, for the inframarginal types, we will be integrating over a different (broader) range of individuals. Moreover, the elasticity of labor effort is itself a feature of the utility function; hence, in general, it too will differ. Regarding all of these differences, one can make the problem more tractable by positing that individuals' utility functions take a particular functional form and make further cardinalization assumptions in order to specify all of the effects in the different groups. Whether such assumptions plausibly correspond to the particular disabilities, family configurations, or other group differences under consideration is another matter.[119]

When utility functions differ across groups, it tends to be optimal to employ not only different grants but also income tax schedules that differ throughout the income distribution. For example, more generous treatment of families with children, particularly young children, is a common feature of categorical assistance. In addition, such benefits in actual tax and transfer systems are often phased out as income rises, so that perhaps by the middle of the income distribution no difference remains. Optimal income tax and transfer schedules, by contrast, tend to be different. Suppose that a more generous grant and lower marginal income tax rates at the bottom reflect that, for a given level of household consumption, the marginal utility of an additional dollar to the household is higher when children are present.[120] This relationship holds throughout the income distribution, even though the magnitude of the differences in marginal utility may well decline with consumption. Indeed, even that decline in differences in marginal utility may not carry the normally supposed implication regarding magnitudes of expenditure: perhaps an additional \$5000 per child at the very bottom equates marginal utilities, whereas in the middle of the income distribution it may take \$10,000 to do so even if the magnitude of the difference in marginal utilities is smaller. This latter point makes it ambiguous whether the magnitude of optimal differences in treatment between the two groups rises or falls with income because social welfare depends on differences in marginal utilities across the groups (and the other factors noted above).[121] It is important to keep in mind when contemplating this question that, for example, higher grants (that are not phased out) or lower marginal tax rates on the L group are not in any sense financed by the bottom of the income distribution within the L group but by higher taxes (including lower grants) in the H group.[122]

To close this section, reflect briefly on the meaning of the grant levels for different groups, which inquiry overlaps with the government's provision of public and private goods as well as choices of cash versus in-kind provision of assistance, which is otherwise abstracted from here (Kaplow 2006b, 2008a). For concreteness, continue to consider households with different

[119] For further discussion, see section 8.2. See also the discussion (and warnings) on the use of welfare weights in section 8.1.

[120] The exposition in the text assumes that the presence of children raises the (internal) marginal utility of consumption, but similar results would obtain if spending on children generated positive externalities. Further subtleties with different implications would arise if parents' time at home with children generated positive externalities.

[121] Optimal asymptotic rates at the top of the income distribution may differ because, even if the distribution of abilities w is the same in the two groups, those at a given income, as mentioned, will be of a different type (the type being lower in the higher-marginal-utility group). This effect would vanish in the limit with (common-parameter) Pareto distributions in all the groups.

[122] Put another way that is familiar but even less rigorous: one can contemplate, say, how net income taxes paid by households earning \$75,000 should vary with the number and ages of children. Solving the optimal income tax problem with multiple schedules answers this question, but not by asking it directly. Rather, standard variational methods would separately consider adjusting the marginal tax rate at each level of income in each group (and the grants). Even if one compared these perturbation experiments at \$75,000 between the two groups, we would need to know the grant levels and marginal tax rates at all income levels below \$75,000 to know how much more or less one household type should pay compared to another.

numbers and ages of children. In many economies, the greatest differentials in government expenditures throughout the income distribution are attributable to education, health care, and other social services. Should the values of these activities to a given household be included in their grants when, say, calibrating an optimal income tax simulation? As long as the levels of these services are being held constant, it would seem that they should not, but it is necessary to ensure that the relevant utility functions reflect these provisions, which influence (among other things) marginal utilities of consumption. Accordingly, if we wish to analyze changes in the level of public provision, account should be taken of how those changes affect marginal utilities of consumption—and, for example, how government support of child care or public transportation may affect the marginal disutility of labor as well. Some of these points connect to the discussion just below regarding the foundations of extensive margin labor supply responses to changes in the income tax schedule, particularly at the bottom of the income distribution.

6.2. *Additional Explorations*

Much of the work on the optimal design of transfer programs, both that which has evolved independently of the literature on optimal income taxation and that which is part of it, focuses on work incentives at the bottom of the income distribution. As noted at the outset, many optimal income tax simulations feature substantial grants, high marginal tax rates, and as a result many very-low-ability individuals choosing not to work. Because the latter feature in particular is unappealing from various perspectives—ranging from concerns about cycles of poverty to nonwelfarist objections to paying individuals not to work—policy advocates and theorists have tried to identify ways around this feature.[123] The straightforward solution of lowering marginal tax rates substantially at the bottom is extremely expensive—$100 billion for each 10% in the earlier toy example. If one sticks with a welfarist framework, as will be done here, then the optimal work incentives at the bottom, all things considered, are whatever is reflected in the optimal income tax schedule that has already been derived. If we are to obtain different results, we must consider plausible modifications of the standard assumptions. Two domains will be explored here.

One set of ideas focuses on work requirements. Because the large work disincentive from high marginal tax rates at the bottom of the income distribution involves a distortion, it would indeed be efficient if these individuals could somehow be induced to supply more labor effort. It should be kept in mind that this may not be optimal for everyone, such as those with severe disabilities or with young children. Moreover, when separate schedules are employed, we saw that the optimal schedule for the more able or less needy group may well feature a smaller grant, reducing the work disincentive, but it also plausibly has higher marginal tax rates at the bottom as well. Conversely, the optimal schedule for the less able or needier group tends to have a higher grant that discourages work but also lower marginal income tax rates at the bottom.

There is a central challenge to designing policies that induce more work effort, namely, that many schemes assume that labor effort is observable. If it actually is, as Mirrlees (1971) explained, the first-best could be achieved throughout the income distribution because each

[123] For example, Besley and Coate (1995) analyze an objective that involves, instead of welfare maximization, the minimization of the cost of bringing all individuals up to a target level of consumption without regard to the utility they thereby achieve. They find that a form of workfare, under which the poorest individuals perform unproductive public service jobs with high disutility, is optimal. When they instead consider a goal of providing at least a minimum level of utility rather than of consumption, such workfare is no longer optimal.

person's tax schedule could be customized to their ability. Type-specific lump-sum taxes would be feasible and optimal. Relatedly, the appeal of observing labor effort in combatting distortion holds for all types, not just those at the bottom, and the social benefits of boosting labor effort are actually greatest for the most productive because the fiscal externalities are the largest.

Perhaps it is possible to observe only whether individuals work at all.[124] In that case, one could reduce the grant for those who do not work (or equivalently provide the grant in the form of highly negative marginal tax rates over a small, initial amount of earnings), which would boost participation (the extensive margin) but not hours (the intensive margin). Analysis of the extensive margin—perhaps the greatest focus in recent literature—is deferred until later in this section. For the present, observe that (absent further restrictions) such a scheme might readily induce a small amount of work, real or artificial, by everyone. And if some could not muster even that degree of effort, perhaps because they are truly in extreme need, the only effect of such a tax schedule revision would be to leave them destitute.

Another possibility is that hours may be more observable for low-skilled occupations.[125] In general, if taxes are lower for a given income if the hours generating that income are higher and the wage concomitantly lower, the optimal employer-employee response is to inflate hours and reduce the stated wage.[126] In many occupations this may be difficult to police, but perhaps for basic, manual labor (and some other jobs), this would be feasible. A binding minimum wage may also play a role because it would prevent employers from reducing the stated wage below that level. For this reason as well, observing hours at the bottom of the income distribution may be easier than at higher levels of income.

When hours can be measured without error by the tax authority, the optimum for such individuals would feature an individualized lump-sum tax or transfer, with a zero marginal tax rate in the relevant range over which such individuals would actually be working. As will be discussed further below when analyzing the extensive margin, one must also address whether higher types would wish to mimic downwards—which might not be easy to prevent unless one could also reliably observe their hours as well. Some prior work, mostly outside the optimal income tax tradition, has addressed various forms of work requirements, mostly exploring particular designs—such as imposing an hours minimum in order to receive some bonus—rather than asking, in the spirit of Mirrlees, what would be the optimal mechanism if one could

[124] Dasgupta and Hammond (1980) consider the case in which ability (or, equivalently, hours) can be observed for all who work but not for those who do not. The optimal scheme features zero marginal tax rates and full type-specific lump-sum extraction of all incremental earning ability for those above a certain ability who work, but the scheme undertakes incomplete extraction for the lower types who work in order to avoid their mimicking even lower types who do not.

[125] The observability of hours and of the wage (ability) are two sides of the same coin when income (the product of the two) is assumed to be observable. The preceding section's discussion of allowing the income tax schedule to be a function of a signal of ability, therefore, already encompasses the case in which one instead has a signal of hours. Nevertheless, existing and proposed schemes focus on hours, which are often regarded to be what may in fact be observed rather than inferred. The point in the text to follow about manipulating both wages and hours while keeping income and actual work effort constant reinforces this duality.

[126] Moffitt (2002) suggests that these difficulties explain why most actual programs rewarding work focus on earnings rather than hours, although it is familiar that if rewards rise in earnings there may be the opposite problem of artificially inflated earnings. All such manipulations are easier by self-employed individuals, who might be induced to take on such work (or purport to do so) to capture earnings subsidies. To combat this problem, Brett (1998) considers how tying benefits to public employment may be selectively advantageous to low-ability individuals because they forgo less market income than do high-ability individuals when they switch from private to public employment.

accurately observe labor effort.[127]

Another set of ideas focuses on the extensive margin of labor supply, which has received significant attention in modern empirical work as well as in modifications of the Mirrlees framework. When marginal tax rates are raised slightly in some band of income, all individuals in that band who continue to work reduce labor effort a bit and perhaps a few of those individuals stop working altogether. Even if the latter group is vastly smaller, the negative fiscal externality they impose is much larger, assuming that positive marginal tax rates are imposed throughout. For both types of responses, individuals' envelope conditions indicate that effects on their own utility can be ignored, so the combined revenue loss is sufficient information to assess the welfare effects of these behavioral responses.

Extensive margin responses do not arise in basic versions of the Mirrlees formulation under certain assumptions. At the core, the marginal disutility of labor effort is rising throughout and a type's wage rate is constant, so slightly lower marginal tax rates at the bottom would at most induce an individual who was not working at all in the market to exert only a very small amount of labor effort. The core exception in the canonical model involves nonconvex budget sets that arise when there are ranges over which marginal income tax rates are falling, as discussed in section 2.1. This concern has received some attention at the upper end of the income tax schedule.[128] It also can be important at the bottom. Many optimal income tax simulations feature falling marginal tax rates there. Moreover, the combined effect of phaseouts across transfer programs has, in certain time periods and for certain groups (many of the programs are categorical), produced very high aggregate marginal income tax rates in certain bands of income near the bottom, followed by much lower marginal rates on somewhat higher incomes.[129] These large nonconvexities in budget sets could explain substantial extensive margin responses. A further implication is that empirical evidence on these responses, drawn from different time periods and different populations (who are often subject to very different all-in marginal tax rates from transfer programs outside the nominal income tax), must be analyzed appropriately and treated carefully when applying it in different contexts.

Much theoretical and empirical work has explored other modifications of the standard model that might microfound extensive margin responses, particularly at the bottom of the income distribution. Not surprisingly, these alternative explanations mostly focus on reasons that individuals' budget sets may be nonconvex.

One possibility is that individuals' before-tax incomes are nonlinear, specifically, that their wage rates w rise with hours. Scheuer and Werning (2017) consider this possibility with a focus on high-wage individuals, like managers, who may have higher marginal productivity when they work longer hours.[130] Similar logic may apply to many shift-workers involving certain skills. For example, due to knowledge losses between shifts, it may be optimal for nurses in many settings to work long hours, or administrative assistants for important executives may

[127]Some of these alternatives are examined in Kaplow (2007a). For example, Blundell and Walker (2002) and Michalopoulos, Robins, and Card (2005) examine, respectively, an existing and an experimental program that reward workers who meet an hours target. This is also one of the features explored in Blundell and Shephard (2012).

[128] For example, Slemrod et al. (1994) study the optimal two-bracket income tax. In their simulations, they find that the optimal marginal tax rate is lower in the upper bracket, which has the consequence that there is an intermediate range of income that no individuals choose to earn. Here, the jumpers do not move between strictly positive labor supply and none but rather between high and low levels of labor supply, skipping levels in between.

[129] For recent analysis in the United States, see Kosar and Moffitt (2017) and Altig et al. (2020).

[130] In their presentation, longer hours raise marginal productivity because managers can augment their scope so as to generate greater complementarities.

need to match their hours or to work overtime to catch up at the end of the day. Splitting such jobs may sacrifice productivity.

This possibility, which may help to explain the prevalence of full-time work, does not seem to be particularly important for most low-skilled occupations. Even nurses might work long shifts but for only a few days a week, and many jobs requiring coverage (cashiers, various attendants, shelf-stockers) often need to span much more than the hours of nine to five, five days a week. That, in turn, requires employees having different shifts of different lengths, with the result that employers may be eager to hire many part-timers to fill in gaps. In fact, over half of minimum-wage jobs in the United States are part time, and many low-skilled individuals work multiple part-time jobs, which is inconsistent with this nonconvexity explanation—and others that follow. See also Blundell et al. (2000) on the prevalence of part-time employment of single mothers in the United Kingdom.[131]

Another natural source of nonconvexity in before-tax income is the existence of fixed costs of employment (Blundell and Shephard 2012; Eissa, Kleven, and Kreiner 2008).[132] Childcare costs are most often mentioned, along with commuting costs, work clothes, and ex ante investments in human capital that are optimal only if the higher wages they enable are going to be earned for a sufficient number of hours.[133] One suspects, however, that many of these factors are more important for higher-skilled occupations, although fixed costs may be a greater fraction of earnings at the bottom of the income distribution.[134] Most fixed costs, however, like nonlinear w's, probably lead to working full days (or even longer shifts), but not necessarily full work weeks. In addition, childcare costs may have the opposite curvature, at least initially: for part-time work, other family members or informal sharing arrangements may be used, whereas full-time work may require turning to more expensive, market-provided childcare (Blundell and Walker 2002). Once again, the fact that so many lower-income earners, particularly those with young children, do work part time suggests that fixed costs may not play a large role for many of

[131] Nevertheless, a longstanding line of theoretical work microfounds extensive margin responses, particularly at the bottom of the income distribution, by assuming that only full-time positions exist in the economy (see Diamond 1980, Saez 2002a, and variations considered by Boone and Bovenberg 2004 and Choné and Laroque 2005). This assumption relaxes the binding incentive compatibility constraint on downward mimicking because the higher type must continue to work full time rather than enjoying the utility gain from reducing labor effort. Some of these models employ an additional assumption that further relaxes the constraint on downward mimicking: higher types are assumed to be incapable of performing lower-skilled jobs. For much of the economy, however, this restriction is implausible, particularly at the lower end of the ability distribution. A fast-food or floor-cleaning shift leader, who earns a bit more than do others on the shift, is obviously capable of performing the others' work and, indeed, the shift leader probably used to work in that lesser role. Unskilled workers who are more prompt or careful can earn somewhat more at higher-end establishments, but it is hard to believe that their higher quality renders them unemployable in less demanding environments.

[132] Employers' fixed costs in hiring and training are also relevant, although they tend to be smaller for unskilled workers. Such costs probably explain why many occupations, even among the unskilled, have minimum hours even if they are far short of full time. The discussion here focuses on pecuniary costs. Nonpecuniary fixed costs (disutility of labor) are examined in Jacquet, Lehmann, and Van der Linden (2013), but it is hard to imagine that these would be significant for many individuals—i.e., that many would face a large disutility from supplying even a tiny amount of market labor, with the marginal disutility steeply falling after, say, the first hour and then rising gradually thereafter.

[133] Those hours may still involve part-time work over extended numbers of years. Moreover, empirical estimates of extensive margin responses over short time periods that result from tax or transfer program changes would be unlikely to reflect responses on this margin.

[134] Interestingly, there are arguments for allowing work-related costs of these types to be tax deductible. If that is not done, perhaps because some of these costs are hard to measure, then these costs (whether fixed or variable) can be understood to implicitly reduce pre-tax income even before the income tax is applied, which may justify lower marginal tax rates.

them.

Much of the theoretical work on optimal income transfers in recent decades has been motivated by empirical evidence of significant extensive-margin labor supply responses to changes in the EITC in the United States (see Eissa and Liebman 1996, Meyer and Rosenbaum 2001, Meyer 2010, Bastian 2020, Schanzenbach and Strain 2020, and the reappraisal by Kleven 2021). Furthermore, there is an interest in rationalizing the EITC's significant negative marginal tax rates at very low incomes when, in the basic Mirrlees framework, negative marginal tax rates are never optimal and, in a variety of simulations, optimal marginal tax rates in this range are high.

Regarding the former, it is important both for the empirical work and theoretical applications to sort out the microfoundations of extensive margin responses. Notably, nonconvexities in tax and transfer programs may offer an explanation. But the degree of these nonconvexities varies greatly across the time periods of different studies—not only because of changes in the EITC but, even more, because of welfare reform—and the groups studied—because different individuals are eligible for different programs that provide different grants (or equivalents) and have different phaseout rates and ranges. In addition, it is important to assess whether many of the individuals may face rising wage rates or be subject to nontrivial fixed costs of employment.

To calibrate an optimal income tax simulation, particularly to illuminate optimal treatment at the bottom end of the income distribution, it is necessary to consider categorical schedules of the sort examined in section 6.1 that match both the composite existing regime and that can be separately calibrated for the different groups using the corresponding selection of data. Furthermore, extensive margin responses are not sufficient statistics. To state this more precisely, different microfoundations for different groups imply (unless due to falling inclusive marginal tax rates) differences in utility functions that are themselves welfare relevant, so that aspect of the optimization needs to be appropriate for each category as well. Although there may exist common reduced forms that one could calibrate to the data, any consequent findings would indicate what was optimal only if these reduced forms were a valid approximation of actual, heterogeneous utility functions of individuals—or of households, since we are often imagining different family configurations, a subject considered in section 8.2.

All of these points are also relevant to the enterprise of rationalization of the EITC. First, as a purely descriptive matter, at different points in time and for different groups, the familiar EITC trapezoid does not even approximately describe many individuals' actual tax and transfer schedules, which is what is required to apply the Mirrlees framework. When other transfer programs provide large grants and also have aggressive phaseouts over different income ranges (and some cliffs), it is that aggregate tax and transfer schedule—which itself varies across household configurations—that must be rationalized. Recall that there is no such thing as an "optimal EITC" in a vacuum. Moreover, much of the theoretical literature's exploration of nonconvexities (or other subjects, such as myopia in section 5.2) has focused on the tantalizing question of whether negative marginal income tax rates at the bottom of the income distribution can ever be optimal. Even when such results are generated, it is usually difficult to likewise generate the EITC's 60% jump up over a fairly narrow income band—something that seems more a product of informal phaseout thinking that, as section 6.1 explains, is fundamentally mistaken as a matter of optimal program design. In addition, an income tax schedule that looked something like the nominal schedule in the United States, including a universal EITC but no other transfers, provides little assistance to the destitute because of the low (or nonexistent) grant, with large associated social costs (Aizer, Hoynes, and Lleras-Muney 2022; Garfinkel et al. 2022). Moreover, as noted at the opening of this section, this involves the imposition of marginal tax rates on the first $10,000 or so of income that are on the order of 100 percentage points below what many simulations suggest to be optimal, at a staggering revenue cost. The most convincing positive explanations of such features—which in respects do not well describe the full system—probably lie elsewhere.[135]

7. *Optimal Income Taxation and Other Instruments*

In principle, we should undertake an integrated assessment of different policy objectives and associated instruments. Whether assessing other forms of taxation, such as differential commodity taxation, or other types of policies, such as regulation or the provision of public goods, there are important interactions with optimal income taxation that run in both directions. Other policies affect the distribution of income and (relatedly) incentives for labor effort, which are at the heart of optimal income taxation analysis. Conversely, because optimal income taxation cannot achieve the first best, it is natural to ask whether some of the shortfall can be redressed with other instruments (adjusting them so as to relax incentive constraints) or otherwise affects how those instruments should be set. Aspects of these interactions have been the subject of many literatures in public economics for half a century.

This section sketches a methodology that has been developed in Kaplow (2008a, 2020a) and other work to bring greater conceptual clarity to this set of problems in a variety of settings that prove, upon analysis, to have much in common. Specifically, it shows how the two spheres of redistributive income taxation and other policy instruments (including other tax instruments) are substantially modular, in the sense used in computer programming and complexity theory (Simon 1962). As we will see, modularity is broader than and somewhat different from particular functional form assumptions, notably, separability, although they are often related. Modularity does not require a lack of interaction among the modules; if it did, it would be unhelpful here because interactions are the focus. Instead, it can be employed whenever it is possible to compartmentalize analysis, even if each compartment may use outputs from others or if the outputs of multiple compartments will be combined at some higher level.

The method does not involve placing the income tax in a separate module from other instruments. Instead, it takes the other instruments and policies under consideration—say, public goods provision—and puts them in a module with a *particular* adjustment to the income tax schedule, one that involves overall distribution-neutrality within the module. The other module, as we will see, involves a purely redistributive adjustment to the income tax schedule, which is to say, standard Mirrlees analysis. That is, the income tax is at play in both modules, but in a particular manner that, as will be explained, enables modular analysis.

After presenting the general framework, this section elaborates the many payoffs from this form of modular analysis as a standard research tool for both theoretical and applied work. Then it offers a range of applications—commodity taxation, externality correction, public goods provision, estate and gift taxation—emphasizing how the approach highlights strong similarities across these subjects.

[135] A potentially important explanation for the EITC is optics: Reducing the rate of welfare program phaseouts at the bottom raises the cost of "welfare" programs and the number "on welfare," whereas an identical change to the nominal income tax schedule of the sort embodied in the EITC does neither of these things and instead is both a "tax cut" and a "reward for work." Relatedly, opposition to welfare may render the otherwise-optimal grant infeasible, and when the grant is suppressed, lower and even negative marginal tax rates at the bottom may become optimal (Boone and Bovenberg 2004). Importantly, this leaves the lowest-ability types who cannot work or must incur large disutility to do so in dire straits.

7.1. *Modular Analysis*

Suppose we wish to analyze some marginal or discrete policy change. For concreteness, think of a change in the level of public goods, but it could also be a regulatory reform, adjustment of a commodity tax vector, or revision of other policy instruments. This policy change, denoted ΔP, will be understood to indicate its incidence across the income distribution. (That is, like income tax schedules, it is a function of y, which is suppressed throughout in this section.) Assume further that ΔP is accompanied by an adjustment to the income tax schedule, ΔT^P. The only restriction is that this adjustment, when combined with the policy change—that is, $\Delta P + \Delta T^P$—is revenue neutral. If ΔP is an increase in the level of a public good, the income tax adjustment must raise the revenue required to pay for that increase, taking into account the behavioral effects of both the policy change and the income tax adjustment. In general, there exists an infinite variety of adjustments to the income tax schedule that have this property. (For policies that themselves have no impact on tax revenue, this set would include a null adjustment.)

Our task is to undertake an overall assessment of the reform package, $\Delta P + \Delta T^P$. The proposed modular approach entails a two-step decomposition (Kaplow 1996b, 2004, 2008a, 2020a). To implement this decomposition, we can construct a different, distribution-neutral adjustment to the income tax schedule, ΔT^{DN}. To be more precise, this schedule is defined such that the overall reform package consisting of $\Delta P + \Delta T^{DN}$ holds the utility of all individuals constant. That is, ΔT^{DN} is the schedule of compensating variations associated with the policy change ΔP. For example, if the policy is a marginal increase in a public good, the income tax adjustment equals the marginal rate of substitution at each level of income.

Some further observations about ΔT^{DN} are in order. First, it is *not* assumed (and in general will not be true) that $\Delta P + \Delta T^{DN}$ is budget neutral, a point that will be elaborated below. Second, the main assumption required to construct a schedule ΔT^{DN} that has the stated property is that there be a single dimension of heterogeneity, as in the standard Mirrlees problem.[136] Separability—notably, weak separability of individuals' utility functions in labor—is not required, as will be evident.

Before proceeding, some additional features of this set-up should be noted because it departs significantly from much analysis in optimal income taxation and other branches of policy analysis, including the important extension of the Mirrlees framework in Atkinson and Stiglitz (1976) that incorporates differential commodity taxation. Notably, it is not required that either the initial income tax schedule or the proposed income tax adjustment be optimal. Relatedly, it also is not required that either the initial policy setting (say, the level of a public good) or the proposed adjustment be optimal. Moreover, the framework is applicable to both marginal and discrete changes in the instruments. To be sure, we will often be interested in the characterization of the optimum, which then will involve analysis of perturbations local to the optimum. But much headway is possible in broader settings that often are of interest. Moreover, many of these more general results can be obtained simply and intuitively using this modular approach rather than conventional optimization techniques.

The actual construction is remarkably simple. To evaluate our original reform package, $\Delta P + \Delta T^P$, we simply decompose it as follows:

[136] Multidimensional heterogeneity is discussed in section 8.2.

$$(10) \qquad \Delta P + \Delta T^P = \underbrace{(\Delta P + \Delta T^{DN})}_{\text{Step 1}} + \underbrace{(\Delta T^P - \Delta T^{DN})}_{\text{Step 2}}.$$

As we will see, step 1, the distribution-neutral module, can be evaluated entirely on efficiency grounds because, by construction, there is no redistribution. In the example of funding a public good, efficiency is determined by the Samuelson (1954) rule, with no adjustments either for distribution or for a marginal cost of public funds—although if labor effort is not weakly separable in individuals' utility functions, there will be a fiscal externality adjustment to that rule. The module comprised by step 2, by contrast, is a purely redistributive change in the income tax schedule (it holds provision of the public good constant), so its welfare impact is determined by standard Mirrlees analysis.

To elaborate step 1, recall that ΔT^{DN} is constructed such that the combination $\Delta P + \Delta T^{DN}$ holds all individuals' utilities constant. However, that is not the end of the analysis because, in doing so, no attention was paid to the overall impact on the government's budget. If there is a budget surplus, it would be possible to undertake a pro rata (or other) rebate scheme to generate a strict Pareto improvement. If there is a budget deficit and if, say, the reform was a marginal one, then reform in the opposite direction would yield a surplus that could be rebated so as to generate a Pareto improvement. Hence, regarding step 1, the impact on the government's budget is a sufficient statistic for policy analysis and, moreover, we have a pure efficiency test. However, unlike familiar efficiency tests (notably, the Kaldor-Hicks test), this efficiency test constitutes a complete welfare analysis. Distributive effects are not ignored; instead, there are none in fact. Hence, all assessments can be made using the Pareto principle.[137]

To make this abstraction more concrete, consider our example of funding a public good, denoted G, and begin with the case in which labor effort enters individuals' utility functions in a weakly separable fashion (and, for ease of exposition, confine attention to the case in which individuals' ability does not affect utility directly). That is, we can write $u(c, l, G)$ as $u(v(c, G), l)$ for some subutility function v. In that case, it can be demonstrated that step 1's distribution-neutral package implies that no type's labor effort will change. As explained in Kaplow (1996b) and elsewhere, the posited tax adjustment, ΔT^{DN}, implies that each type's total utility as a (reduced-form) function of labor effort, l, is unchanged by step 1's policy reform. Therefore, whatever level of labor effort previously maximized utility will continue to do so.

We can now determine the effect of $\Delta P + \Delta T^{DN}$ on the government's budget in a straightforward manner. The revenue raised by ΔT^{DN} is, by construction, the integral of individuals' compensating variations, which for a public good is the integral of their marginal rates of substitution (for a marginal change; if the change is discrete, we would integrate that integral over the change in the public good). Hence, there is a budget surplus if and only if individuals' total (unweighted, measured in dollars) value of the increase in the public good exceeds the cost of increasing provision of that public good, which is the pure Samuelson rule. There is no question of whether one should be weighting individuals' valuations for distributive effects because there are none. Note further that the implicit cost of government funds in this exercise is 1.0, and that this is so regardless of what the preexisting income tax schedule is and what distortion it involves. The reason none of this matters is that the experiment holds distribution constant, and it is redistribution that causes any labor supply distortion in this setting.[138]

---

[137] Not surprisingly, this strong claim is where the assumption of a single dimension of heterogeneity is important.

[138] Note that the finance mechanism, ΔT^{DN}, is *not* a uniform or a type-specific lump-sum tax. Rather, as stated, it is an adjustment to the entire income tax *schedule*. For example, if the public good had the same value *in utils* to all

Suppose now that labor is not weakly separable in individuals' utility functions. Then, we would have an additional efficiency term associated with step 1. For example, if the public good was a leisure complement—say, improved parks—there would be a negative fiscal externality that would reduce any efficiency gain (or increase any efficiency loss) associated with increasing expenditures on the public good. But if the public good was a leisure substitute—perhaps improved urban transit—there would be a positive fiscal externality. Note that, in either case, because of the construction of ΔT^{DN}, the relevant fiscal externality *for step 1* is that directly associated with achieving a given level of utility by providing more of the public good in lieu of consumption of private goods. The total budgetary impact in the absence of weak separability is any surplus or deficit associated with the pure Samuelson rule plus or minus the fiscal externality due to any labor supply adjustments associated with greater provision of the public good.

This latter adjustment is related to the longstanding injunction to tax leisure complements and subsidize leisure substitutes. It begins with Corlett and Hague's (1953) analysis of differential commodity taxation in a Ramsey framework and was famously integrated into optimal income tax analysis by Atkinson and Stiglitz (1976), discussed further in section 7.3. More broadly, this idea should be understood in the context of Lipsey and Lancaster's (1956) general theory of the second best, which holds that, in the presence of other distortions in the economy, standard first-best prescriptions no longer govern. The central preexisting distortion here is the labor-leisure distortion that is inherent in redistributive income taxation when individuals' abilities differ but only income can be observed by the tax authority. In a wide range of contexts examined in section 7.3, first-best principles continue to be applicable when labor is weakly separable in individuals' utility functions because, when that is true, distorting other margins cannot improve (or worsen) the labor-leisure distortion. Note further that, when ability rather than labor effort interacts with other margins (notably, when individuals' utility functions depend on ability in ways that bear on their preferences associated with such margins), it has been understood since Mirrlees (1976) that corresponding deviations from otherwise first-best principles may improve (or worsen) the labor-leisure distortion (see also Saez 2002b). In intuitive terms, these possible interactions with labor effort or with ability are the central answers to this section's opening query regarding the possibility that the use of other instruments may relax incentive constraints in the optimal income tax problem.

Return now to step 2, $\Delta T^P - \Delta T^{DN}$. Because both terms solely involve changes to the income tax schedule, this module is the pure Mirrlees problem. Specifically, we need to assess the difference between the actually contemplated income tax adjustment, ΔT^P, and our constructed, distribution-neutral tax adjustment, ΔT^{DN}. If our original reform, $\Delta P + \Delta T^P$, entails an increase in redistribution, for example, there will be associated welfare gains, whose magnitude will depend on the SWF, and on the increase in labor supply distortion with associated negative fiscal externalities, whose magnitude will depend on the usual elasticities and the density function of individuals' types.

This two-step decomposition and, in particular, our second module highlight that many analyses focusing on policy instruments other than the income tax entail latent redistribution. This will be true whenever $\Delta T^P \neq \Delta T^{DN}$. Indeed, this is often true even for a pure regulatory change that has no budgetary impact and for which $\Delta T^P = 0$ because, fairly broadly, ΔP will not be distribution neutral. Step 2's $\Delta T^P - \Delta T^{DN}$ indicates how redistribution may be greater

individuals, then this income tax adjustment would be rising with income at the reciprocal of the rate the marginal utility of consumption was falling with income because individuals' marginal rates of substitutions are their ratios of the (here, constant) marginal utility of the public good to their (not constant) marginal utilities of consumption.

overall, less overall, or different in ways that cannot so readily be characterized—for example, when a reform benefits the middle class at the expense of both the rich and the poor. Regardless, the policy analysis of step 2 is just that of the standard Mirrlees problem, no more and no less. And, because the two-step decomposition is modular in the manner described, the Mirrlees analysis can be conducted independently of any analysis of the distinctive features of ΔP, the analysis of which (stripped of its distributive effects) is fully contained in step 1.[139]

7.2. *Elaboration*

The particular form of modular analysis associated with the two-step decomposition offers a number of benefits in a wide range of applications. Perhaps most important, it advances conceptual understanding and enhances clarity, including in the communication of results within the field and to a broader policy audience.

Consider two separate analyses of a carbon permit trading scheme that are conducted in conventional ways, which is to say that each, in the background, employs some income tax adjustment to balance the budget. The first study finds an overall welfare gain and the second a welfare loss. But why do their conclusions differ?

We can apply the two-step decomposition to each analysis. Consider the following possibility. The first study, at step 1, actually finds an efficiency loss but, because the permits were taken to be auctioned and the proceeds used to reduce redistributive taxes, step 2 had a large enough positive impact on labor effort to generate an overall efficiency gain. Moreover, that overall efficiency gain was described as a social welfare gain due to the use of a representative-agent model wherein there is no welfare loss associated with any implied reduction in redistribution. The second study, let us imagine, has the opposite features: there is an efficiency gain at step 1 but a larger loss at step 2 because the proceeds were used to reduce taxes in a highly redistributive fashion (for example, by raising the grant). And again, no welfare consequence was attributed to the latter because of the use of a representative-agent model.

This illustration, in part based on how important work in the field has actually been undertaken—see the discussion in Kaplow (2012)—highlights a number of problems of failing to employ the two-step composition wherein step 1 is, in aggregate, distribution neutral. First, regarding the analysis of the permit scheme, it appeared that the first study found it desirable and the second undesirable, but the decomposition reveals that it is the other way around. More broadly, if research is to progress in a focused manner, it is critical that not just the signs but the underlying sources of results—and actual, apples-to-apples disagreements—be understood. Was there a difference between how industrial sectors were modeled? In adjustment costs? In the rate of technological change? In the data sets used to calibrate key parameters? In functional form assumptions? In the strengths of other policies aimed at reducing carbon emissions?

Untangling all of these questions and advancing understanding on all fronts is difficult enough. We hope to build knowledge as research progresses, with subsequent efforts refining methods and applying them to richer data sets. But when the results from each study along the way are entangled with different distributive effects arising from different assumptions about

[139] That said, it is often forgotten that the standard Mirrlees problem abstracts from many features of real economies and government actions. Most obviously, whenever other taxes are present, changes in the distribution of income—through effects on labor supply and from income effects on consumption—generally involve fiscal externalities. Not only that: redistribution can, for example, alter the usage of publically funded roads, changing the degree of wear and tear that in turn requires repairs that likewise involve a fiscal externality. For present purposes, the central point is that these additional effects arise even without any change in policy, although different policies will often affect these background conditions.

income tax schedule adjustments—and perhaps different assumptions and calibrations related to that part of the analysis—the problem is needlessly confounded. The latter difficulty, however, is entirely avoidable if researchers do one of two things. First, researchers can employ the two-step decomposition, reporting the results for step 1 separately. Second, they can simply eschew the analysis of redistribution altogether by stipulating that the income tax adjustment to be analyzed is the distribution-neutral one, that is, $\Delta T^P = \Delta T^{DN}$, so that step 2 is null. As discussed in section 7.1, the policy can then be assessed using the Pareto principle based on whether it, combined with the distribution-neutral income tax adjustment, generates a government surplus or deficit.

Before proceeding, it is worth reflecting further on the use of representative-individual models, something that many researchers (including this author) find helpful in a wide range of settings. The aforementioned difficulties arise precisely when this simplification is mixed, often for purposes of greater realism (and to calibrate models to data), with redistributive instruments. Notably, an income tax is often employed in representative-agent models because of the recognition that uniform lump-sum finance of government operations (the optimal tool when everyone is identical) is highly regressive and that type-specific lump-sum taxation is infeasible. But if the income tax is going to be modeled for such reasons, then one faces two choices. One can take on the distributive analysis explicitly, rendering the representative-agent model inapt and thus losing the benefits of simplification. Or one can take an often clearer and easier route by sticking with the representative-agent setup but, to avoid step-2 contamination that is not going to be analyzed, be sure to hold distribution constant which entails setting to zero the implicit (or, one might say, virtual) redistribution associated with the reform. The latter can be done by using a distribution-neutral adjustment to the income tax. Then, confining analysis to step 1 of the decomposition is legitimate and avoids misleading results that entangle but do not identify distributive effects. For this distribution-neutral adjustment, one raises or rebates revenue and otherwise accounts for interactions by constructing an income tax schedule adjustment that holds utility constant throughout the relevant range. When this is done, the representative individual who contemplates raising or lowering labor effort after reform of the permit scheme will find that the achieved utility for each choice matches that in the original regime. (With weak separability of labor in the utility function, labor effort would not change; with nonseparability, there will be a fiscal externality, but this modeling approach would limit it to that associated with how the environmental policy affects the marginal disutility of labor, excluding further effects from changes in the degree of implicit redistribution.)

A second advantage of the modular approach is that it facilitates specialization. This is clearest when those studying, say, public goods or permit schemes, in fact set $\Delta T^P = \Delta T^{DN}$, rendering step 2 null. In addition to the clarity and comparability of results associated with step 1, the researcher's specialty, there is no need to analyze the Mirrlees problem at all. This avoids having to take a stand on elasticities, the distribution of abilities, and the SWF—and to perform the associated analysis. It may be valuable to include, or even for some research to focus on, the distributive incidence of all manner of policy changes, that is, to determine what ΔT^{DN} is for a given set of reforms. But if one then sets $\Delta T^P = \Delta T^{DN}$, that completes the distributive analysis.

When seeking to analyze some overall package, $\Delta P + \Delta T^P$, any researcher or policy-maker can combine the best step 1 results with whatever that analyst or policy-maker deems to be the best step 2 analysis, using an SWF of their own choosing. Returning to our two studies of permit schemes, policy-makers would like to know, regarding step 1, what is the truth (or best understanding) of the matter. And if multiple ΔP's are on the table, policy-makers would like to know which are best as a matter of environmental policy. If distribution-neutral implementation of a given option or of the best option is efficient—which, as explained in section 7.1, means it can be implemented so as to generate a Pareto improvement—there is good reason to favor it. If step 1 is inefficient, it should be eschewed. If a policy-maker likes the distributive effects of some $\Delta P + \Delta T^P$ even though step 1 is inefficient, it would be superior (indeed, Pareto superior) to implement only step 2, that is $\Delta T^P - \Delta T^{DN}$. And if a policy-maker dislikes the distributive effects of some $\Delta P + \Delta T^P$ even though step 1 is efficient, it would be superior to implement step 1 with some other income tax schedule adjustment—indeed, perhaps a distribution-neutral one.

These points about specialization and a policy-maker's perspective suggest a third advantage of the modular approach, in the realm of political economy. Consider why anyone would wish to analyze $\Delta P + \Delta T^P$ in the first place, rather than just ΔP in isolation. One justification is that ΔP alone may not be budget-neutral. If that is the only reason, one faces the problem (or luxury) that there exists an infinite number of ways to design ΔT^P to meet the government's budget constraint. Hence, the question becomes why one would wish to analyze a particular ΔT^P, and specifically one that differs from ΔT^{DN}. The best answer would seem to be realism: perhaps a pending proposal, $\Delta P + \Delta T^P$, employs the particular ΔT^P that one now seeks to incorporate into the analysis.[140]

This justification, on reflection, is problematic. Usually this is not so in fact, so the researcher's choice of a particular ΔT^P only makes sense if it is a good prediction of what actually will happen. That, in turn, requires the researcher to be an expert in political economy, specifically, skilled in the prediction of the distributive politics of the regime in question. But even experts at this task find such predictions difficult. By the time research is completed and a working paper is posted, the answer may well have changed, and yet again by the time of publication and even again by the time the paper is drawn upon subsequently. When a given ΔT^P is actually proposed, it often will change before enactment. Or it may be a tax reform that would at least partly have been implemented in any event. Or it may be one that would not be long-lasting. The challenge of selecting the true ΔT^P can also be seen in another way: different analysts of the same policies, say a given permit scheme, often make different choices of ΔT^P. They cannot all be right if these particular choices are to be rationalized as predictions.

In sum, the presentation of analyses of $\Delta P + \Delta T^P$ that forgo the two-step decomposition create a Tower of Babel that obscures analysis of the distinctive policies—such as various public goods or permit schemes—as well as that of redistribution. It sacrifices the benefits of specialization and, in particular, implicitly involves researchers embedding political economy assumptions in their analysis. Conceptual clarity, communication, and specialization are aided by using the two-step decomposition or, better still, having most research specialize in the first module by employing what may be regarded as a universal benchmark of analyzing $\Delta P + \Delta T^{DN}$. Note that performing distribution-neutral analysis does not at all downplay the importance of distribution, which is at the heart of the optimal income tax problem that is the whole of step 2. Rather, it highlights rather than hides distributive effects that are often implicit in the ΔP under analysis or embodied in a particular ΔT^P that is not the focus of the study and may be largely submerged when employing representative-agent models.

[140] Another reason is tractability: adjusting just the grant (intercept), moving the tax schedule in a linear or proportional manner, or varying a single parameter in, say, a two-parameter tax schedule to balance the budget may be easier in certain respects. As the analysis throughout this section explains, however, what can then be learned is often hidden and may be quite difficult to extract as a consequence, and in any event all such alternatives entail some mixing of redistribution with other effects of the other policy under analysis. Moreover, as explained, using ΔT^{DN} actually simplifies much analysis because step 2—and all the requisite assumptions and calibrations—can then (legitimately) be skipped entirely.

7.3. *Applications*

This modular approach has been applied to a number of tax and nontax policy instruments. Much of this work employs the further assumption of weak separability of labor in individuals' utility functions to simplify the exposition. When that is done in basic settings, optimal policy is fully dictated by familiar, first-best policy rules, such as those favoring uniform commodity taxation, first-best Pigouvian correction of externalities, and public goods provision according to the Samuelson rule.

The most familiar such conclusion—although, as we shall see, employing a qualitatively different approach—is the result of Atkinson and Stiglitz (1976) that, when there is a nonlinear income tax that is optimally set, weak separability implies that optimal commodity taxes are uniform. This paper further noted results for the nonseparable case. As it happens, their verbal formulation was reversed, although, fortunately, the correct intuition from Corlett and Hague (1953)—that leisure complements should be taxed, not subsidized—was the lesson most economists understood and lived by nonetheless.[141] In any case, familiar Ramsey (1927) results such as the inverse-elasticity rule were not merely qualified but overturned.[142] With separability, commodities' elasticities are entirely irrelevant, and with nonseparability, the sign of optimal differentiation (a relative tax or subsidy) is determined entirely by the cross-elasticity with leisure, not the magnitude of the own-price elasticity relative to that of other commodities.

The potential broader implications of Atkinson and Stiglitz (1976) did not achieve much traction for decades. One indication is the continued focus on Ramsey taxation in leading texts and surveys (Myles 1995, Auerbach and Hines 2002, and Salanié 2011). The failure to pursue Atkinson and Stiglitz (1976) can be explained in many ways. One is the longstanding view that, because their derivation makes use of first-order conditions, the uniformity result holds only when the income tax schedule is in fact set optimally. Many took the optimality of the existing income tax system to be an inapt assumption when offering guidance to policy-makers. Also regarding commodity taxation, Atkinson and Stiglitz (1976) only characterize the optimum and hence could not be applied to reforms not in the neighborhood of the optimum.

On another important dimension, the close connection between commodity taxation and myriad other policy instruments was not adequately appreciated. More concretely, the modular approach with the two-step decomposition was not employed even though in a very rough sense it has long been part of economists' thinking.[143] Finally, the Pareto principle was naturally

[141] Atkinson and Stiglitz's misstatement was also replicated in subsequent texts (Myles 1995, Salanié 2003). For further discussion and a formal treatment that traces the misunderstanding to a misinterpretation of the sign of the costate variable in the Hamiltonian for the optimal income tax problem, see Kaplow (2010b).

[142] Standard formulations of the Ramsey problem, importantly, do not allow for a nonzero intercept (a uniform lump-sum grant). Statements such as that in Salanié (2011) that introduce treatments of the Ramsey approach by observing that a wage or income tax is allowed but must be linear are misleading because they omit the further restriction that there be a zero intercept. See also section 2.2. Actually, that restriction, not linearity, is necessary to generate the core results in the literatures that build on Ramsey. The relationship of Ramsey models to modern work that admits an income tax is discussed further in Atkinson and Stiglitz (1976), Stiglitz (1987), Mirrlees (1994), and Kaplow (2008a). Scheuer and Werning (2018) examine how one can establish a theoretical linkage between Mirrlees (1971) and Diamond and Mirrlees (1971) despite the fact that the former features a nonzero intercept whereas the latter, on the surface, does not. This conceptual point does not, however, restore traditional Ramsey results because standard formulations do not admit the necessary extension introduced by Scheuer and Werning (2018).

[143] Some readers may recall Musgrave's (1959) suggestive distinction between what he termed the Allocation and Distribution Branches of government.

understood to be inapt in assessing nearly any policy that had distributive effects. Economic analyses of various policies often required revenue-neutral tax adjustments, but distribution neutrality was only occasionally examined.

The development of what is here described as the modular approach began with cost-benefit analysis of public goods provision. Hylland and Zeckhauser's (1979) under-appreciated article showed that, in a simple model, no distributive weights should be employed in cost-benefit analysis. They used a distribution-neutral rather than optimal income tax approach and did not relate their work to Atkinson and Stiglitz (1976).[144] The modular approach was made explicit in Kaplow (1996b, 2004) in showing the applicability of the simple, unweighted Samuelson rule when weak separability is imposed.[145] Results were also sketched for Pigouvian taxes and subsidies, and both settings were related to Atkinson and Stiglitz's (1976) analysis of commodity taxation. For public goods, invoking the earlier notion of Lindahl (1919) pricing is one way to make this connection explicit.[146]

The most direct and broadest extension of Atkinson and Stiglitz (1976) is Kaplow (2006a), which uses the modular approach—rather than an integrated optimization that makes use of sometimes-complex first-order conditions that apply only in the neighborhood of the optimum— to derive more general results with less analytical effort. The broadest proposition there states that, with weak separability, a pure efficiency test characterizes *any* change in an arbitrary, initial commodity tax vector, regardless of the initial specification of the income tax. If one employs ΔT^{DN}, rendering step 2 moot, a commodity tax reform yields a Pareto improvement if and only if it reduces resource use in the economy. This result holds regardless of whether changes are marginal or discrete and regardless of how many elements of the vector are adjusted and in which direction. Another, more concrete result is that proportionally reducing any differential commodity tax vector can be implemented with ΔT^{DN} so as to generate a Pareto improvement. Note that such a proportional reduction reduces in a uniform fashion the degree of preexisting differentiation. Of course, the latter entails the special but interesting case of reducing all commodity taxes to zero (moving any revenue-raising from above-mean commodity taxation to the income tax schedule). And a special case of that type of reform involves, as in Atkinson and Stiglitz (1976), reforms in the neighborhood of the optimum.[147]

Kaplow (2012) derives analogous results for environmental taxation, where the zero commodity tax vector (the simplest case of uniform taxation) is replaced by a vector of first-best Pigouvian taxes and subsidies—which, of course, all equal zero when there are no externalities. Results for reforms short of moving to the first best require additional assumptions because of the possibility that incompletely corrected externalities may be exacerbated at some segments along a "straight" path to the first best. Similar results can be derived in other policy domains.[148] Keep in mind that all of these applications make use of Pareto assessments and hold regardless of

---

[144] Shavell (1981) extended Hylland and Zeckhauser (1979) to legal rules.

[145] For similar results in the neighborhood of the optimum in the Atkinson-Stiglitz (1976) tradition, see Christiansen (1981) and Boadway and Keen (1993).

[146] It is also useful to reflect on the connection between the provision of public goods and the correction of atmospheric externalities. It is remarkable that the second-best literatures that related each subject to income taxation developed almost entirely independently of each other. Yet we can consider, for example, how similar it is, on one hand, to directly raise the quality of public parks, and on the other hand, to reduce an externality whose only effect is to worsen the quality of the same parks.

[147] See also Konishi (1995) and Laroque (2005).

[148] Kaplow (2021), discussed in section 4.3 on market power and rents, uses the same techniques to show that, starting with an arbitrary income tax and competition policy, any reform of the latter can generate a Pareto improvement if and only if the reform raises the sum of consumers and producers surplus.

the original income tax, the distributive effects of ΔP, and any effects that it, standing alone, may have on labor effort.

The modular approach also illuminates the understanding of tax instruments that are often regarded as part of the distributive apparatus, notably, capital income taxation, wealth taxation (which in some models is equivalent to a form of capital income taxation), and wealth transfer taxation, that is, estate and gift taxation. A similar two-step decomposition may be employed in these spheres as well (Kaplow 2008a). Hence, we can ask: For two individuals of equally high labor income, to what extent should the taxes imposed on one be higher or lower than those imposed on the other if the former saves more (generating more capital income and wealth) or gives more to children in lieu of increasing own consumption? Atkinson and Stiglitz (1976) famously explained that one could interpret different commodities as consumption in different periods of time, generating basic results on capital income taxation. These now have been extended in subsequent literatures, although that approach has been eschewed in other important work on the subject.[149]

Kaplow (1998c, 2001) introduced the two-step decomposition for the analysis of estate and gift taxation. When distribution (in the donor generation) is held constant by using ΔT^{DN}, there are two remaining efficiency effects: the direct positive externality on the donee and a negative fiscal externality associated with the wealth effect on the donee's labor supply.[150] In addition, there is a distributive effect because the resulting reduction in the donee's marginal utility of consumption is, in general, relevant to the social planner. This framework has been applied and extended by Farhi and Werning (2010), Kopczuk (2013), and others.

A wide range of policies, including public goods, environmental regulation, and most forms of taxation, have important distributive effects. Indeed, many policies are favored in part because of their distributive consequences, and others may be scaled back because of their adverse distributive impacts. A correct assessment necessarily requires that such analyses be integrated with the analysis of optimal income taxation in order to determine the appropriate setting of multiple instruments to hit multiple targets. The modular formulation examined here—mostly aimed at particular applications—suggests a way forward that is rigorous, more general, and often simpler to implement than are other approaches.[151] The two-step decomposition, and in particular the use of ΔT^{DN} that enables analysis of step 1 in isolation, clarifies thinking, brings into focus important commonalities across diverse lines of research that have been pursued separately, facilitates specialization, and enables the analysis of optimal income taxation to proceed in a largely autonomous fashion that nevertheless can be linked to other policy assessments in a straightforward manner.

[149] The dynamic Mirrlees literature examines labor supply over multiple periods in the presence of earnings uncertainty (Golosov, Tsyvinski, and Werning 2007). By contrast, Chamley (1986) and Judd (1985) and subsequent elaborations by Chari, Nicolini, and Teles (2020) and Straub and Werning (2020) have proceeded independently.
[150] Diamond (2006) and some others regard the former as involving double counting, but in that event it must be either that we do not have an individualistic SWF (elaborated in section 8.1)—here, by ignoring the utility parents obtain from their children—or that children (even adult children) are not in the SWF.
[151] Complementary work, such as Slemrod and Yitzhaki (2001) and Hendren (2016, 2020), employs integrated approaches that do not use the two-step decomposition. Alternative methods that seek to address both efficiency and distribution combine in some fashion measures of the marginal cost of public funds and the marginal value of public funds, which requires explicit use of distributive weights. For a recent set of applications, see Hendren and Sprung-Keyser (2020).

8. *Utility and Social Welfare*

Much work on optimal income taxation, particularly regarding applications and simulations, focuses on how the optimal schedule depends on $f(w)$, the density function for abilities, and ε, the elasticity of labor supply. The first-order condition for optimal marginal income tax rates (expression (8) in section 2.3), however, also depends on how changes in consumption influence social welfare, reflected in the expression $W'u_c/\lambda$, where the dependence of W' and u_c on realized utility, consumption, and labor effort of each type is suppressed in this notation. Some aspects of the SWF and individuals' utility functions are longstanding subjects of inquiry in the field, some have received significant attention more recently, and yet others have been largely unrecognized and thus unexplored. This section, building on Kaplow (2008a, chs. 13–15), selectively examines some of these issues. The aim is to illuminate and sometimes recast a range of modern work in the field that mostly focuses on other subjects, where the SWF and utility functions are in the background. The analysis also identifies a number of topics that would benefit from further research. As is often the case, many of the questions were identified in Mirrlees (1971), and, regarding the SWF, were insightfully elaborated in Mirrlees (1982).

8.1. *Social Welfare*

Viewed broadly, an SWF can embody all manner of objectives. In applied welfare economics and much work on optimal income taxation, attention is confined to individualistic SWFs, that is, the arguments of the SWF are individuals' utilities and nothing else. In respects, this choice is unsurprising. On one hand, the effects of policies on individuals' well-being seem patently relevant to how such policies should be evaluated. And on the other hand, it is not immediately apparent why individuals' well-being should be sacrificed to serve objectives independent of anyone's well-being, or at least to serve them to a degree or for reasons beyond any effects on well-being when tradeoffs are involved.

This approach to the social objective, of course, entails value judgments, and alternative approaches to social welfare have been developed (Fleurbaey and Maniquet 2018). Nevertheless, Kaplow and Shavell (2001) prove that, with a modest continuity assumption, all nonwelfarist SWFs violate the Pareto principle. That is, all nonwelfarist SWFs sometimes place a higher value on policies under which every individual is strictly worse off. The nub of the simple proof is as follows: If individuals' utility profiles are not sufficient information under a given SWF, then there exist states of the world in which everyone's utility is the same but the SWF ranks them differently. Starting with the lower-ranked state—that is, the one disfavored on some nonwelfarist ground—construct another state in which everyone's consumption is higher by epsilon, holding everything else equal, including the degree to which the nonwelfarist objective is served. For epsilon sufficiently small, the posited nonwelfarist SWF still ranks the now-modified state lower even though everyone in that state is strictly better off than in the higher-ranked state.

This demonstration has implications for a wide range of normative questions, many of which are explored at length in Kaplow and Shavell (2002). For example, Rawls (1971, 1982) famously advanced that social welfare should be assessed with respect to what he called primary goods rather than utility, and Sen (1985a, 1985b, 1997) advanced individuals' functionings or capabilities in place of utility. As Kaplow (2007b) shows, not only do all such formulations violate the Pareto principle, but for related reasons they in principle favor wide-ranging policies that forbid trades (having no externalities and subject to no information infirmities) and even much private individual activity (such as an individual transforming one good into another, preferred one). After all, when left to themselves, individuals will maximize their utilities, not some other, externally specified maximand that conflicts with their utilities—that is, unless they are compelled to behave otherwise.[152] This conflict with the Pareto principle seems quite problematic for those who advance such alternatives in the name of personal autonomy and freedom, as many proponents do. Of course, no logic requires acceptance of the Pareto principle, and many indeed advance nonwelfarist principles—although usually without acknowledging (and often being unaware) that the identified principles conflict with the Pareto principle.

An important reconciliation, owing to a line of philosophical writing going at least back to Hume (1751), Mill (1861), and Sidgwick (1907) and perhaps first elaborated by an economist in Harrod (1936), distinguishes two levels at which principles may operate.[153] A familiar example is that a social norm commanding truth-telling, even though it sometimes reduces welfare, may be overall best for social welfare for a range of familiar reasons.[154] Under this view, truth-telling is not itself a constituent of welfare but rather instrumental to it. Nevertheless, it often proves useful to take truth-telling as a proxy objective, all the more so given limitations of human nature and social institutions. Indeed, much of Sen's advancement of functionings and capabilities can be understood as instrumental. For example, Sen (1985a) elaborates the practical concern that conventional well-being measures, particularly applied to developing countries, focus excessively on market income at the expense of other indicators. Such concerns gave rise to such constructs as the Human Development Index, the use of which (particularly in place of sole reliance on per capita income) hardly implies a normative rejection of placing individuals' well-being at the center of the social assessment of regimes.

Returning to tax policy in particular, a range of nonwelfarist principles have been advanced, such as concerns for horizontal equity and mobility. In light of the above, it should not be surprising that, if taken as part of the social objective, they generate Pareto conflicts and pose other problems (Kaplow 1989, 1995). However, their allure can readily be understood if they are taken instead as proxies, in many instances as signals of how well institutions are performing. Regarding horizontal equity, many violations of equal treatment entail errors (if two individuals really are identically situated in relevant respects, it is usually optimal to treat them the same way) and, often in practice, violations of equal treatment reflect invidious discrimination, political favoritism, or corruption. Similarly, significant immobility may indicate inefficient roadblocks to success as well as failures to make valuable investments, such as in the human capital of those in difficult circumstances. Mobility is unusual in this regard because scores of "zero" (complete rigidity) and "one" (random assignment of individuals to tasks) both signal serious malfunctions. Moreover, if the mobility measure is 0.47 when all policies are optimized but is only 0.32 currently, it does not follow that all actions that increase mobility (such as may arise from forcing some random swaps) raise social welfare. The fact that mobility as such is problematic as an underlying maximand does not deny its important use as a proxy or

[152] Sen's (1970) famous proof of the impossibility of a Paretian liberal is notable in this regard. Sen advanced an axiom he associated with liberalism over the Pareto principle but did not explain how the requisite implicit bans on trade were consistent with the motivations for his purportedly liberal precept. In Sen's example and corresponding proof, adherence to his stated maxim requires that the two individuals be prohibited from undertaking a latent mutually advantageous trade, which he does not mention. Nor does Sen in his elaborate later writings on capabilities and functionings highlight the requisite rejection of the Pareto principle.

[153] The most extensive elaboration is Hare (1981), who also wrote an interesting review (Hare 1973) of Rawls (1971), with particular attention to maximin, a subject also addressed by Arrow (1973).

[154] For models in this spirit, see Kaplow and Shavell (2007) and Weinzierl (2017).

diagnostic measure for policy analysis in many settings.[155]

The key lesson, which although familiar is often forgotten, is that many objectives are worth pursuing because they are useful institutional or social guidelines or serve as useful proxies that may indicate how social welfare can better be raised, not because they are themselves constitutive of welfare, to be pursued in principle at the expense of (possibly everyone's) well-being. Of course, none of the foregoing explanation for the allure of such objectives on welfarist grounds rules out the possibility of embracing nonwelfarist principles as first principles, to be pursued at the expense of individuals' well-being.[156]

Consider next the increasing use of social welfare weights in optimal income tax investigations and more broadly in applied welfare economics. Certain usages appeared in the earliest optimal income taxation literature in the 1970s, which have been continued and sometimes replaced by others in more recent work. Some points of clarification and caution are in order lest important assumptions or subtleties be overlooked. First, the notion of welfare weights has been used by different authors in different contexts to mean different things, so one must match the right context to the right application and, even then, readers need to be attentive to nuance to make sure they draw the right lessons, particularly when comparing results across papers or transplanting to new applications. Second, sticking to a particular usage in a given context, care must be taken because the weights might be misunderstood as exogenously specified when they are endogenous and hence change when various parameter differences or policy comparisons are contemplated.

It is useful to begin by considering what the weights might represent and, closely related, just what it is that they are weighting. A common usage in optimal income tax analysis, such as in the familiar first-order condition for optimal marginal tax rates, is that they weight the experienced consumption of different types. The expression employed here (in the first-order condition (8) in section 2.3) for that weight is $W'u_c/\lambda$. A marginal dollar consumed by a given individual raises utility by u_c, which receives a social weighting of W' (which with a utilitarian SWF can be taken to be one for everyone; otherwise it depends on the individual's utility *level* because the argument of W is u). This product is divided by λ, the marginal social value of a dollar to the treasury, which can be interpreted as the value of a dollar of additional expenditure on the uniform grant (the 0-intercept of the income tax schedule), which in turn is the average marginal social value of a dollar in the population. Hence, $W'u_c/\lambda$ is the marginal social value of a dollar to a given individual relative to the average marginal social value over all individuals.

First, reflect on why this usage is consistent with the Pareto principle in light of the fact

[155] Measures of inequality and poverty similarly pose difficulties when viewed as ultimate objectives rather than as useful diagnostics (Kaplow 2005). An important consideration in the use of indexes as proxies is that the appropriate index depends on the application. For example, the best summary statistic regarding inequality when trying to predict election outcomes may be the situation of the median voter's distance from the mean, whereas that for predicting crime rates, revolutions, or economic growth may be altogether different. There are obvious virtues to researchers' use of off-the-shelf measures for a variety of purposes, but caution is in order to ensure that the chosen measure is plausible in the context at hand. By contrast, most normative uses of such indexes are problematic on their own terms because they usually require discarding relevant information and giving them any independent weight conflicts with the Pareto principle.

[156] In behavioral welfare economics, significant attention has been devoted to the question of when individuals' preferences (utility functions) for normative purposes can appropriately be taken to differ from their revealed preferences (their behavioral utility functions) (Bernheim and Rangel 2007, 2009). In interesting contrast to some of the literatures discussed in this section, those analyses generally do not contemplate widespread replacement of individuals' behavioral or "underlying" (well-informed, rational, self-controlled) utility functions with some externally stipulated utility function that comports with an outside observer's favored nonindividualist normative principles.

that this term is being used to weight a marginal dollar of consumption rather than utility as such. In the basic formulation, utility depends only on consumption and labor effort. Although labor effort is omitted from this expression, the marginal utility of consumption fully captures the marginal effect on utility because of individuals' envelope condition, as discussed in section 2 and elsewhere in this article. By contrast, section 5 explains how behavioral factors can lead this envelope condition to fail, resulting in a deviation that requires amendment to the first-order condition for optimal marginal income tax rates.

Second, observe that each of the components of $W'u_c/\lambda$ is endogenous. Suppose, for example, that we adjust some parameter of the optimization, as simple as raising the posited labor supply elasticity. For given marginal tax rates, a higher elasticity implies less work effort, lower consumption, different levels of utility, and less revenue. Lower revenue implies a smaller grant, which itself affects utility levels and marginal utilities of consumption, and it likewise raises the shadow value of government funds. Hence, if one began with stipulated welfare weights that implicitly reflected some SWF, and one held those weights fixed when performing this policy experiment, then the implied adjustments in the "optimal" income tax schedule would be comparing tax schedules that implicitly maximized different underlying SWFs. That is, the actual comparison would be between an initial situation with the original SWF and a low elasticity of labor supply and a modified situation with an implicitly altered SWF with a higher elasticity. Note further that, if neither SWF is stated explicitly and hence the differences between them are difficult to discern, it will be challenging to interpret the results from such a comparison.

Again taking the behavioral optimal income tax analysis in section 5 as an illustration, recall that each of the posited behavioral infirmities affected not only whether individuals' choice of labor effort was privately optimal but also the level of revenue raised as a function of the behavioral parameter. In one formulation of schmeduling, individuals taking their lower, average tax rates as their marginal tax rates leads to greater labor effort, and in a model with myopia, labor effort is lower than otherwise. As explained there, the revenue effect itself is an important factor in assessing how optimal income tax rates should adjust, and in plausible cases this factor points in the opposite direction of the internality correction.[157] More broadly, the lesson here is that one can use welfare weights as an expositional and notational stand-in for more complex expressions, but in both analysis and simulations one must be careful to use the underlying utility functions, SWF, and shadow value of funds because the weights are functions of these variables that are each endogenous to the experiment.

In the preceding discussion, it is supposed that what is being weighted by each individual's welfare weight is that individual's realized consumption. Hence, if we perform, say, some local perturbation of the income tax schedule, we can trace through all of the effects on behavior, revenue, the implied grant, and ultimately the marginal change in each individual's consumption. More often, taking advantage of the envelope theorem (when all individuals are rational utility-maximizers), we can confine attention to direct effects and fiscal externalities. Another definition of welfare weights, however, includes the fiscal externality associated with income transfers within the weights themselves. To take a simple illustration, when income

[157] To elaborate, consider a utilitarian SWF so that W' is constant and hence can be ignored. If revenue is, say, lower, both u_c and λ will be higher (supposing that the revenue reduction requires a corresponding decrease in the common lump-sum grant). With standard utility functions, these results will not be offsetting. Instead, u_c will rise by more than average at the bottom of the income distribution and less than average at the top, which raises the marginal social value of greater redistribution. If the welfare weights were taken to be exogenous, this important force would be omitted.

effects are accounted for, giving an individual a dollar (lump-sum) raises that person's own utility by u_c; due to their envelope condition, we can ignore the adjustment to labor effort on utility, but that adjustment has a fiscal externality that is socially relevant. Some usages include the full revenue effects of labor supply adjustments in that type's welfare weight. This alternative usage is sometimes convenient, but note how the resulting welfare weight is qualitatively and quantitatively different from the one discussed previously. Instead of referring to the impact on social welfare associated with the change in that individual's utility, we instead have the impact associated with implicit changes in all individuals' utilities associated with giving that individual a dollar, where effects on other individuals' utilities are captured by λ, the shadow value of government revenue.[158]

Welfare weights also can be employed as a way of embodying non-welfarist considerations, as advanced in Saez and Stantcheva (2016). In theory, weights can be stipulated in any fashion and thus be made functions of anything. Hence, if the weights are taken to be endogenous, they can respond to nonwelfarist features that might be influenced by a reform under analysis. As is clear from the earlier discussion in this section, however, the SWF that implicitly corresponds to such nonwelfarist weights would violate the Pareto principle. Saez and Stantcheva (2016) accordingly emphasize key domain restrictions on their approach that are necessary to avoid Pareto conflicts. First, only local analysis with prespecified (rather than endogenous) weights is possible. Second, any two discrete regimes are, in principle, non-comparable using their approach (Sher 2021). These welfare weights can thus be understood as a construct in which there is no SWF (individualistic or otherwise): an SWF is ordinarily taken to rank all admissible social states, but this method by design cannot do so and must eschew being extended to enable such comparisons because that would create Pareto conflicts.

Furthermore, although such nonwelfarist weights are in principle quite flexible, there are subtleties involved in mapping many nonwelfarist principles into welfare weights. Consider Saez and Stantcheva's (2016) example in which a principle of horizontal equity is understood to prohibit the use of some tag. As they explain, if one wishes to use weights rather than simply prohibit use of the tag as a constraint on the optimization, one has to posit weights exhibiting no differential when the tag is not used at all but weights with an infinite relative differential if the tag is used even infinitesimally in a prohibited manner. In other settings, because the weights likely depend on some or all of the parameters of the system as well as the existing nonlinear income tax schedule, solving the reverse exercise for the weights that locally instantiate a given nonwelfarist principle may not be straightforward. Although this is not difficult for some nonwelfarist principles, for others it is challenging, in which case it would probably be easier to explicitly posit a (non-Paretian) SWF that embodies the nonwelfarist principle and maximize that SWF directly.

A final use of welfare weights is to employ them as a reduced-form way to combine welfarist concerns for equality that have different underlying sources. This practice traces to early work on optimal income taxation, wherein analysts posited different degrees of curvature for the mapping from consumption directly to social welfare. This approach was agnostic about the contribution of curvature in the utility function (diminishing marginal utility of consumption) and curvature in the SWF (notably, social preferences that are more egalitarian than utilitarian

[158] Another feature of this usage—which dates back to some of the earliest work on optimal income taxation—is that often (for example, in the discussion to follow of nonwelfarist weights) it is imagined that the weights are stipulated and have purely normative content, which we can now see involves an endogeneity issue regarding fiscal externalities in addition to the aforementioned endogeneity concerns involving the marginal utility of consumption, the marginal effect of utility levels on social welfare, and the marginal value of revenue.

ones). While convenient, this approach blurs positive and normative questions and sometimes poses analytic difficulties (Kaplow 2010a). A related notion emphasized in some more recent work is that the use of unspecified weights allows one to map the Pareto frontier. One can then characterize properties of optimal income tax schedules that depend only on the Pareto principle and not on the particular SWF. This can also be done using an explicit but flexible individualistic SWF.

In considering these two sources of curvature, Mirrlees (1982) favors the view that the single, correct SWF is utilitarian (additive), so the curvature question should depend entirely on the rate of diminishing marginal utility of consumption. This view has a long history in welfare economics, beginning with Vickery's (1945) suggestion and Harsanyi's (1953) simple proof that an individual behind a veil of ignorance, facing an equal prospect of becoming any individual in the actual society, would rationally maximize expected utility. Harsanyi (1955) offered a complementary, more subtle demonstration that, if social preferences as well as individual preferences adhere to a familiar set of rationality axioms, the SWF has to be utilitarian. (He also assumed that social welfare is positively responsive to individuals' utilities and that each individual receives equal weight. That is, his analysis focused on the linearity of $W(u)$.)

Subsequent explorations show that a sort of time consistency in policy evaluation requires linearity (Hammond 1983, Myerson 1981, Ng 1981). This problem can be illustrated and extended with a simple example from Kaplow (1995). Suppose that all individuals are ex ante identical. They contemplate implementing a policy that is risky but involves a slight boost to their expected utilities. Specifically, the resulting distribution of outcomes is determinate, but it is random which individual will experience which outcome. These individuals would unanimously adopt this policy despite its resulting inequality. However, for any nonlinear SWF, there exists a small enough certainty equivalent associated with such a policy such that the social planner would reverse the policy, if feasible. But, once reversed, the individuals would unanimously favor implementing it once again. The cycling can be averted by the social planner rejecting the project in the first place. But doing so violates the Pareto principle. Hence, the Pareto principle is transgressed not only by all nonwelfarist SWFs but also by welfarist but nonlinear SWFs, once one allows for uncertainty.

8.2. *Utility*

Section 6.1 on optimal income transfers introduced a signal θ that, among other possibilities, might be associated with differences in individuals' utility functions. Motivating examples that will be variously elaborated here include physical disabilities and different family configurations. More broadly, heterogeneity in utility functions may or may not be observable; may affect the disutility of labor effort or the utility of consumption, each in qualitatively different ways; may be differentially cardinalized; and may be deemed normatively relevant or not, in a variety of ways. Heterogeneous preferences are thus a particularly heterogeneous phenomenon. Many prior treatments of preference heterogeneity, unsurprisingly, make particular (often implicit) choices on each of these dimensions and hence inevitably deliver a wide range of results. This section explores some of these issues in order to understand and reconcile prior work and to identify useful avenues for future research.

To begin, suppose that our baseline utility function, $u(c,l)$, is modified to become $u(c,\theta l)$ (Kaplow 2008a, Choné and Laroque 2010, Lockwood and Weinzierl 2015). Next, make the familiar substitution, $l = y/w$, to write $u(c,\frac{\theta}{w}y)$. One can interpret w, as before, as ability, and θ as a comparable indicator of disability. Doubling both θ and w has no effect on an individual's labor effort, before-tax income, taxes paid, consumption, or utility—that is, if T continues to depend only on y and not on θ, or if θ is unobservable in any event.

Suppose next that θ is observable and, using expression (9) in section 6.1, consider how the optimal income tax schedule should depend on θ. At any income y, a higher θ is associated with a correspondingly higher w, for the reason just given. Hence, the value of $f(w)$ will in general differ, $F(w)$ will be higher (and thus $1 - F(w)$ will be lower), and obviously the w in the denominator of the first-order condition will be higher. Thus, optimal income tax schedules will in general depend on θ, and in somewhat subtle ways. Suppose instead that θ is unobservable. In that case, which entails unobservable multidimensional heterogeneity, there will be a continuum of $\{w,\theta\}$ types at each level of income, which complicates the analysis but, for given assumptions on the joint distribution of w and θ, enables one to determine the optimum.[159]

The analysis would differ if one instead adopted a nonwelfarist SWF that differentially treated different $\{w,\theta\}$ types that exerted the same labor effort and enjoyed the same utility.[160] For example, Lockwood and Weinzierl (2015) examine the use of social welfare weights that depend explicitly on the individual's w-type but not their θ-type, reflecting the assumption that w indicates a morally relevant "ability" whereas θ indicates a morally irrelevant "taste." (That is, society wishes to offset distributions in income attributable to different abilities but not to different tastes.) They characterize cases in which a greater variance in θ favors less redistribution, reflecting that higher incomes only in part reflect ability, which should be muted in its effect on after-tax income, while they also in part reflect stronger preferences for work, which are deemed not to constitute a valid justification for redistribution. As with all choices of SWFs, value judgments are required, including that, here, accepting nonwelfarist SWF implicitly entails rejecting the Pareto principle.[161]

The foregoing analysis focuses on the second argument of the utility function and, moreover, assumed that the heterogeneity, indexed by θ, took a particular and convenient functional form. In theory, such heterogeneity regarding variations in the disutility of labor could take any form, and the significance of particular forms of heterogeneity poses an empirical question that, in most instances, is quite difficult to answer. The concrete example of certain types of disabilities—and variations in family composition, considered below—may be among

[159] Some recent work on optimal income taxation addresses challenges in making multidimensional heterogeneity tractable (Rothschild and Scheuer 2014, Jacquet and Lehmann 2021).

[160] It may not always be immediately apparent when a nonwelfarist SWF is being employed. In this setting, the core idea is that two exogenous parameters receive different social weight even when they offset and hence generate the same utility, so an individual's realized utility is not sufficient information for determining that individual's contribution to social welfare. Formally, even if the social welfare weights are used to weight individuals' utilities, when those weights depend on a trait of the state of the world other than utilities, the result is a nonwelfarist SWF. Accordingly, as discussed in section 8.1, conflicts with the Pareto principle will arise.

[161] As is already apparent from the text, one can as well interpret the "distaste" for work as a "disability," which many would believe constitutes a valid basis for compensation. Regarding both parameters as normatively symmetric could reflect a judgment that equal normative (distributive) weight should be accorded to individuals born with different$\{w,\theta\}$ combinations but having the same θ/w and hence, in a first-best world, the same opportunity sets (expressed in utility space, or in the tradeoff of the marginal utility of consumption and the marginal disutility of labor). Note further that the two-level view discussed in section 8.1 provides a welfarist explanation for the allure of this nonwelfarist view: as a matter of ideal theory, there may well be no difference (pure welfarism), but as a matter of social practice, it may be useful to inculcate norms favoring work for a variety of reasons, which implies that public expressions, exhortations, and even sometimes policies seen as instantiating that attitude may have instrumental value. A nonwelfarist SWF that combines the two levels of analysis and only adjusts welfare weights is unlikely to yield even an approximate solution to this two-level optimization problem.

the more important and more readily measurable, which requires undertaking and triangulating results from empirical investigations in a number of disciplines.

To further explore the ways in which utility functions may exhibit heterogeneity, suppose instead that individuals differ in their utility from consumption.[162] Under some formulations, the results would be similar (Kaplow 2008a). To see this, note that, while w indicates how a unit of effort (conventionally measured as hours but, in light of the preceding discussion, might also be measured in disutility) is converted into (before-tax) income, some variable θ might indicate how a unit of consumption (after-tax income) is converted into utility. An individual would be indifferent upon learning that each given choice of labor effort yielded only half as much consumption as before if each unit of consumption now generated twice as much utility.

For this case, consider a number of different senses (functional forms) in which alternative preference specifications may imply different realized utility from a given level of consumption (Kaplow 2008b). Taking one special case, suppose that the utility from consumption was additively separable from the disutility of labor and that θ scaled the overall utility of consumption. An individual with a higher θ, therefore, would have a correspondingly higher marginal utility of consumption (at all levels of consumption) and a higher total utility. Through the u_c term in the first-order condition, this would indicate a higher welfare weight, whereas through the W' term, this would favor a lower weight to the extent of the curvature of the SWF (having no effect with a utilitarian SWF). By contrast, if θ were added to effective consumption—that is, $c + \theta$ replaced c in the utility function—then a higher θ would imply a lower u_c (with diminishing marginal utility of consumption), indicating a lower welfare weight, and the W' effect would be as before because, again, the level of utility is higher. If instead, as suggested earlier, θ multiplied c, there would be opposing effects on u_c: if $\theta > 1$, u_c would be higher because each increment to c is weighted more, but it would be lower because of diminishing marginal utility (at any existing c, effective consumption would be higher so the marginal utility of effective consumption would be lower). With the functional form $\ln \theta c$, these effects would be precisely offsetting. Taking these three variations together, we can see that different forms of preference heterogeneity regarding individuals' utility from consumption can have qualitatively different implications for optimal income taxation.[163]

Finally, consider differences in utility functions associated with different family or household composition (Kaplow 1996a, 2008a). As outlined in section 6.1, we can think of the optimal regime as involving a separate income tax schedule, including separate intercepts, for each configuration, denoted again by θ.[164] Much prior work on optimal income taxation of the family takes the challenge to involve the joint treatment of two adult (potential) earners who may

[162] This possibility underlies a longstanding intuition that heterogeneity favors less income redistribution because it raises the possibility that marginal utility falls less with consumption than meets the eye because individuals with higher marginal utilities of consumption for that very reason choose to earn more. As will be seen, this intuition is not supported by all plausible formulations of the phenomenon.

[163] The discussion in the text does not distinguish cases depending on whether θ is observable. It should be clear that, if it is, the analysis of section 6.1 with separate schedules would be appropriate, with cross-type differences in tax schedules following accordingly (but involving subtleties, including that individuals earning a given y will be different types w, with all the implications of that). If θ is not observable, then those at a given income y will be of different $\{w,\theta\}$ types, with the relationship between them depending on the functional form of the heterogeneity.

[164] This approach assumes that family configurations are observable, but there are important limits due to the optionality of marriage, the difficulty of observing which dependents should be associated with which (potential) earners, and the deeper fact that there is a continuum of relationships among individuals that may form a household, which itself is neither unambiguous in principle nor always observable in fact. Many of these features could be modeled as adding dimensions of heterogeneity conditional on the categorization of households, with the usual implication that, at a given level of income y, there will be different underlying types.

have different elasticities and hence optimally be taxed differently on account of fiscal externalities.[165] By contrast, the focus in this section is on how θ also signals different utility functions, which directly imply differences in the optimal income tax schedule in the manner discussed in section 6.1.

To begin, the analysis here will take each individual family member's utility function to be an independent and relevant object (even if their utility functions exhibit interdependence). That is, the SWF is taken to be a function of each individual's utility rather than some composite utility function associated with the family as a whole. If the SWF is linear (utilitarian) and one takes a family's utility function to be the sum of its members' utility functions, there would be no difference, but when the SWF is strictly concave, differences would arise (along with additional analytical complications due to the endogeneity of each family member's utility).

If each family member has the same utility function, which may capture cases in which the family consists only of two adults, a number of issues may still arise. First, the members may share disposable income—creating some voluntary redistribution that is generally taken to be absent between unrelated individuals—but they may not share their income equally, depending on household bargaining, preferences, social norms, and other matters. To fix thinking, suppose that two adults share disposable income in some fixed proportion; perhaps the husband earns more than the wife and only shares a portion of the difference. If their total tax payments are lowered by a dollar, that increase in disposable income will, on one hand, go disproportionately to the husband, who has a lower marginal utility of consumption, but on the other hand, go in part to the wife who has a higher marginal utility of consumption. If utility is more (less) concave in consumption than $\ln c$, the total marginal utility of the dollar will be higher (lower) than under equal sharing, favoring greater (less) redistribution toward such a family. In either case, a strictly concave SWF would, ceteris paribus, place greater weight on a family that shares income unequally because of the relatively higher weight on utility gains to the less-well-off family member.

Second, suppose that there are identical utility functions and equal sharing, but now introduce economies of scale in household production. Then the analysis follows that associated with one of the above cases of heterogeneity in the utility of consumption. On one hand, a dollar to the family raises utility more because they convert that dollar into more units of effective consumption. On the other hand, because they already will have obtained a higher level of effective consumption, diminishing marginal utility makes incremental units of effective consumption less valuable. Here, the latter effect dominates when the curvature of utility is greater than that with $\ln c$. If one superimposes a strictly concave SWF, economies of scale will on that account favor reduced generosity because those benefiting from scale economies will experience higher levels of utility for a given level of disposable income.[166]

---

[165] Much of this work considers the optimal interrelationship of the income tax schedules of the two adults in the family, including the possibility that the applicable schedule may be gender based (Alesina, Ichino, and Karabarbounis 2011; Kleven, Kreiner, and Saez 2009; Frankel 2014; Gayle and Shephard 2019). See also Blau and Kahn (2007), who find that married women's labor supply elasticities fell substantially in the United States toward the end of the twentieth century, their labor supply behavior in many respects converging toward that of married men. Each prior optimal tax investigation at least implicitly takes a stance on the subjects discussed in the text that follows. Relatedly, two earners in a family unit having different elasticities may itself imply differences in their utility functions, some of which may relate to the presence of (especially young) children, whose existence (utility functions) are usually suppressed in the analysis.

[166] Another important case that bears some analytical similarity is one that incorporates family members having interdependent utilities, for example, with altruism. As explored in Kaplow (1996a, 2008a), such preferences make family members more efficient utility generators, raising the marginal utility of a dollar of disposable income and also raising utility levels.

Third, consider a special case of interest with different utility functions, specifically, a parent and a child where the child needs less consumption to obtain a given level of utility.[167] Simple cases would be where the child needs fewer calories, less expensive clothing, or is readily entertained with less expensive leisure activities. Here we have a phenomenon similar to that with economies of scale in household production. The child is taken to be a more efficient generator of utility than is an adult, which can make a marginal dollar to the household (say, shared in some fixed proportion between the parent and child) more or less valuable depending on the curvature of utility and of the SWF.

But even apart from these effects regarding the translation of disposable income into utility, the mere presence of additional family members, whether an additional adult or an additional child, implies a higher marginal utility of consumption for each individual at a given level of disposable family income, favoring more generous treatment.[168] If utility functions of adults are taken to be independent of the presence of children, their utility levels will be lower, which also favors more generous treatment to the extent of the concavity in the SWF, but the premise that children do not directly affect parents' utility is dubious.[169]

In reflecting on these phenomena, note that they do not depend qualitatively on the aggregate disposable income of a family of a given configuration. For example, if a family with a child should optimally be treated more generously than an otherwise identical one (same adults) with no children, this will be true throughout the income distribution. Hence, the notion that transfer program or income tax benefits for children need to be phased out in some sense reflects a misunderstanding of the problem. Income tax and transfer schedules should optimally be set separately for adults with no children, for those with one child, and so forth, and those schedules are linked by the common shadow value of government revenue. In a sense, those with no or fewer children may be paying for those with some or more children, just as, under separate schedules, healthy individuals may pay for benefits given to the disabled.

Furthermore, it is hardly obvious that, in absolute dollars, optimal generosity falls rather than rises with income. Compare two families with the same high incomes, where only one has children, and contemplate how much additional disposable income that family would require to have the same marginal utility of consumption per dollar as the other. Even if (without redistribution) the difference in utility levels and in marginal utilities of consumption is not that large in absolute terms at high levels of income, it may nevertheless take substantial redistribution of dollars at such income levels to equalize marginal utilities between the two families. Of course, as throughout this discussion, the analysis here abstracts from both labor

---

[167] The analysis focuses on differences in the utility from consumption associated with the presence of children, but the disutility of an adult's labor effort (utility of leisure time) may differ as well. That would not only affect the purely distributive considerations examined here but also may influence the elasticity of labor effort (although differences in the marginal utility of consumption do so as well).

[168] In empirical efforts to apply the optimal tax framework to families with children, it is important to take into account that much social support for children is in-kind, including components like free public education that are not means-tested and, relatedly, often not part of most thinking about the optimal treatment of families as a function of the numbers and ages of their children.

[169] Compare two sets of couples, each with the same earning capacity, who wish to have a child. If only the first succeeds, even though each parent's individual consumption is reduced due to some sharing with the child, by revealed preference their utility level must nevertheless be higher. That is, the utility functions of adults differ across households in some fashion, but one that plausibly implies higher, not lower utility levels for those with (wanted) children. An unappreciated implication of a strictly concave SWF is that it may accordingly imply that fewer resources should be available to families that succeed in having children.

effort and the endogeneity of all aspects of family composition.[170] As in chapter 6.1, optimal schedules for each configuration θ do not entail full compensation at each level of earnings but rather differences in disposable income that reflect all of the relevant components that influence the entirety of both income tax schedules. The analysis in this section focuses solely on the $W'u_c/\lambda$ term in the first-order condition (9) for each schedule (and also with regard to the intercepts), recognizing that both W' (if the SWF is strictly concave) and u_c each depend on θ in a number of plausibly significant ways.

9. *Conclusion*

Much has been learned in the half century since Mirrlees' (1971) pioneering effort. But much remains to be done. Starting with some of the most basic elements, the Mirrlees (1971) static framework can only be interpreted as a collapsed dynamic model that, at a minimum, views individuals (and families) over a lifetime. Among other implications, most simulations of optimal policies fail to reflect central elements in their modeling and their calibrations. Moreover, due to limitations of empirical knowledge, they also are unable to assess long-run effects of reforms, which may favor prescriptions significantly different from those offered.

Growing concern about inequality and its determinants increases the importance of continuing the recently reinvigorated research agenda addressed to multidimensional abilities and endogenous wages. Likewise, it is necessary to significantly elaborate models of founders who supply labor effort and capital (including sweat equity) that, in light of moral hazard and asymmetric information, is supplied in ways that are outside most prior analyses of optimal income taxation. Because these founders earn huge amounts of income and possess great wealth, and are disproportionately represented at the top of the income distribution, it is all the more important to pursue this line of research.

The relevance of production externalities from labor effort is increasingly understood although it is not yet clear how important are its implications for optimal income tax schedules rather than to the design of more targeted policies. By contrast, many divergent potential implications of externalities on other individuals' utilities have not been examined. Existing results reflect particular choices of both channels of influence and functional form, and empirical evidence to guide future work is limited, often in ways that may be difficult to overcome. Market power and rents have only recently been related to the optimal income taxation problem. Preliminary results suggest that there may be large mechanical implications but less significant conceptual differences in how optimal income tax analysis should be conducted.

Behavioral economics has only begun to penetrate the field of optimal income taxation. The most longstanding line of work examines individuals' systematic misunderstandings of the income tax schedule, a subject where there is significant room for novel empirical exploration using existing data from prior natural experiments. Additional research addresses myopic labor

[170] An interesting, largely unexplored aspect of the endogeneity of family composition involves the lifecycle effects noted in section 2.4 that are central in light of the fact that the static Mirrlees framework can only be interpreted as a reduced form that collapses a richer, dynamic problem. A given individual starts as a child (in a family setting that may change during childhood in various ways), continues often as a single adult for some period of time, may then at some point become part of a couple, followed (or not) by children, who in turn subsequently leave the household (although altruistic ties may remain), with retirement at some time thereafter. Addressing lifecycle behavior with respect to investments in human capital, labor supply, and savings—much less marriage, child-bearing, and divorce—involves myriad interactions. From an optimal income tax perspective, many of these concern how to think of $W'u_c$ when, in a sense, the pertinent utility function is changing over time or, put another way, important arguments of the utility function other than consumption and labor supply are changing over time.

supply and the intersection of behavioral analyses of savings and savings policies with effects on labor supply and hence on optimal income tax analysis. These subjects also seem susceptible to significant empirical study using available data to answer new questions, the answers to which can in turn illuminate what combinations of assumptions are most important to explore further.

Optimal income transfers are a socially consequential subject that has received much less attention than has optimal taxation at the very top of the income distribution. Because many transfer programs, including those like the EITC that are nominally part of income tax schedule, are categorical (for example, being primarily available to families with young children), it is necessary to extend analysis and recalibrate simulations to reflect this reality and to connect to the proper theoretical framework that involves multiple income tax schedules linked by a common shadow value of public funds. Efforts to analyze work inducements and extensive-margin responses of low-skilled workers need stronger microfoundations to craft models that better match administrative limitations and empirical evidence.

In parallel with the rise of optimal income tax analysis, there has developed a body of second-best work in public economics that takes into account interactions between other policy instruments and the income tax, particularly through effects on labor effort. Most work to date treats the analysis of each type of instrument—corrective taxes, public goods provision, estate and gift taxation—as its own subject. Yet all are amenable to a single, comprehensive treatment with regard to effects on distribution and labor supply. Moreover, the suggested approach exhibits significant modularity (even without conventional separability assumptions), enabling largely independent analysis of redistributive income taxation using the Mirrlees framework and assessment of the other instruments, each aimed at their distinctive policy targets.

Normative aspects of optimal income tax analysis are predicated on the choice of social welfare function and features of individuals' utility functions. A welfarist approach is elaborated for the former, motivated by work showing that nonwelfarist principles imply conflicts with the Pareto principle. The increasing use of social welfare weights is examined, with emphasis on their endogeneity and other features that can generate misunderstandings. The analysis of individuals' utility functions emphasizes different types of heterogeneity that can generate varying and even opposite implications. Of particular interest, distributive issues concerning taxation of the family—involving the treatment of couples and of households with children—are explored. Most existing understandings of the implications of heterogeneity, whether regarding family composition or otherwise, actually involve analysis of the effects of particular types of heterogeneity that are modeled using particular functional forms. Both empirical and normative guidance are therefore critical, although in this realm they often are not easy to come by.

This article emphasizes important theoretical channels for future research, reflecting the focus of this investigation. But many empirical avenues are exposed as well, demonstrating the well-appreciated symbiotic relationship between theoretical and empirical inquiries. With much effort and some good luck, the next half century of research on optimal income taxation will prove even more fruitful than the last.

References

Aaberge, Rolf, and Magne Mogstad. 2015. "Inequality in Current and Lifetime Income." *Social Choice and Welfare* 44 (2): 217–30.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60 (2): 323–51.

Aizer, Anna, Hilary Hoynes, and Adriana Lleras-Muney. 2022. "Children and the US Social Safety Net: Balancing Disincentives for Adults and Benefits for Children." *Journal of Economic Perspectives* 36 (2): 149–74.

Akerlof, George A. 1978. "The Economics of 'Tagging' as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning." *American Economic Review* 68 (1): 8–19.

Ales, Laurence, Musab Kurnaz, and Christopher Sleet. 2015. "Technical Change, Wage Inequality, and Taxes." *American Economic Review* 105 (10): 3061–101.

Ales, Laurence, and Christopher Sleet. 2016. "Taxing Top CEO Incomes." *American Economic Review* 106 (11): 3331–66.

Alesina, Alberto, Andrea Ichino, and Loukas Karabarbounis. 2011. "Gender-Based Taxation and the Division of Family Chores." *American Economic Journal: Economic Policy* 3 (2): 1–40.

Allcott, Hunt, Benjamin B. Lockwood, and Dmitry Taubinsky. 2019. "Regressive Sin Taxes, with an Application to the Optimal Soda Tax." *Quarterly Journal of Economics* 134 (3): 1557–626.

Allen, Franklin. 1982. "Optimal Linear Income Taxation with General Equilibrium Effects on Wages." *Journal of Public Economics* 17 (2): 135–43.

Altig, David, Alan J. Auerbach, Laurence J. Kotlikoff, Elias Ilin, and Victor Ye. 2020. "Marginal Net Taxation of Americans' Labor Supply." NBER Working Paper 27164.

Andreoni, James, Brian Erard, and Jonathan Feinstein. 1998. "Tax Compliance." *Journal of Economic Literature* 36 (2): 818–60.

Arrow, Kenneth J. 1973. "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice." *Journal of Philosophy* 70 (9): 245–63.

Atkinson, Anthony B. 1983. *The Economics of Inequality*. 2nd ed. Oxford: Clarendon Press.

Atkinson, Anthony B., and Joseph E. Stiglitz. 1976. "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics* 6 (1–2): 55–75.

Atkinson, Anthony B., and Joseph E. Stiglitz. 1980. *Lectures on Public Economics*. New York: McGraw-Hill Book Company.

Auerbach, Alan J., and James R. Hines, Jr. 2002. "Taxation and Economic Efficiency." In *Handbook of Public Economics*, Vol. 3, edited by Alan J. Auerbach and Martin Feldstein, 1347–421. Amsterdam: Elsevier.

Bastani, Spencer, Sören Blomquist, and Luca Micheletto. 2013. "The Welfare Gains of Age-Related Optimal Income Taxation." *International Economic Review* 54 (4): 1219–49.

Bastian, Jacob. 2020. "The Rise of Working Mothers and the 1975 Earned Income Tax Credit." *American Economic Journal: Economic Policy* 12 (3): 44–75.

Basu, Susanto. 2019. "Are Price-Cost Markups Rising in the United States? A Discussion of the Evidence." *Journal of Economic Perspectives* 33 (3): 3–22.

Bénabou, Roland. 2002. "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70 (2): 481–517.

Bernheim, B. Douglas. 1994. "Personal Saving, Information, and Economic Literacy: New Directions for Public Policy." In *Tax Policy for Economic Growth in the 1990s*, edited by

Charles E. Walker, Mark A. Bloomfield, and Margo Thorning, 53–78. Washington, D.C.: American Council for Capital Formation.
Bernheim, B. Douglas. 2002. "Taxation and Saving." In *Handbook of Public Economics*, Vol. 3, edited by Alan J. Auerbach and Martin Feldstein, 1173–249. Amsterdam: Elsevier.
Bernheim, B. Douglas, and Antonio Rangel. 2007. "Behavioral Public Economics: Welfare and Policy Analysis with Nonstandard Decision-Makers." In *Behavioral Economics and Its Applications*, edited by Peter Diamond and Hannu Vartiainen, 7–84. Princeton, NJ: Princeton University Press.
Bernheim, B. Douglas, and Antonio Rangel. 2009. "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics." *Quarterly Journal of Economics* 124 (1): 51–104.
Bernheim, B. Douglas, and Dmitry Taubinsky. 2018. "Behavioral Public Economics." In *Handbook of Behavioral Economics—Foundations and Applications*, Vol. 1, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 381–516. Amsterdam: North-Holland.
Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian. 2008. "The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States." In *Lessons from Pension Reform in the Americas*, edited by Stephen J. Kay and Tapen Sinha, 59–87. Oxford: Oxford University Press.
Besley, Timothy, and Stephen Coate. 1995. "The Design of Income Maintenance Programmes." *Review of Economic Studies* 62 (2): 187–221.
Bhandari, Anmol, and Ellen R. McGrattan. 2021. "Sweat Equity in U.S. Private Business." *Quarterly Journal of Economics* 136 (2): 727–81.
Bick, Alexander, Nicola Fuchs-Schündeln, David Lagakos, and Hitoshi Tsujiyama. 2019. "Why Are Average Hours Worked Lower in Richer Countries?" NBER Working Paper 26554.
Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz. 2019. "Monopoly Power and Endogenous Product Variety: Distortions and Remedies." *American Economic Journal: Macroeconomics* 11 (4): 140–74.
Blau, Francine D., and Lawrence M. Kahn. 2007. "Changes in the Labor Supply Behavior of Married Women: 1980–2000." *Journal of Labor Economics* 25 (3): 393–438.
Blundell, Richard, Alan Duncan, Julian McCrae, and Costas Meghir. 2000. "Evaluating In-Work Benefit Reform: The Working Families Tax Credit in the UK." JCPR Working Paper 160.
Blundell, Richard, and Andrew Shephard. 2012. "Employment, Hours of Work and the Optimal Taxation of Low-Income Families." *Review of Economic Studies* 79 (2): 481–510.
Blundell, Richard, and Ian Walker. 2002. "Working Families' Tax Credit: A Review of the Evidence, Issues and Prospects for Further Research." *Économie Publique—Études et Recherches* 11: 77–127.
Boadway, Robin, and Michael Keen. 1993. "Public Goods, Self-Selection and Optimal Income Taxation." *International Economic Review* 34 (3): 463–78.
Boppart, Timo, and Per Krusell. 2020. "Labor Supply in the Past, Present, and Future: A Balanced-Growth Perspective." *Journal of Political Economy* 128 (1): 118–57.
Boone, Jan, and Lans Bovenberg. 2004. "The Optimal Taxation of Unskilled Labor with Job Search and Social Assistance." *Journal of Public Economics* 88 (11): 2227–58.
Boskin, Michael J., and Eytan Sheshinski. 1978. "Optimal Redistributive Taxation When Individual Welfare Depends upon Relative Income." *Quarterly Journal of Economics* 92 (4): 589–601.
Brennan, Geoffrey. 1973. "Pareto Desirable Redistribution: The Case of Malice and Envy." *Journal of Public Economics* 2 (2): 173–83.

Brett, Craig. 1998. “Who Should Be on Workfare? The Use of Work Requirements as Part of an Optimal Tax Mix.” *Oxford Economic Papers* 50 (4): 607–22.
Brito, Dagobert L., and William H. Oakland. 1977. “Some Properties of the Optimal Income-Tax.” *International Economic Review* 18 (2): 407–23.
Browning, Edgar K. 1994. “The Non-tax Wedge.” *Journal of Public Economics* 53 (3): 419–33.
Bubb, Ryan, and Patrick L. Warren. 2020. “An Equilibrium Theory of Retirement Plan Design.” *American Economic Journal: Economic Policy* 12 (2): 22–45.
Buchanan, James M. 1975. “The Samaritan’s Dilemma.” In *Altruism, Morality, and Economic Theory*, edited by Edmund S. Phelps, 71–85. New York: Russell Sage Foundation.
Bulow, Jeremy I., and Lawrence H. Summers. 1984. “The Taxation of Risky Assets.” *Journal of Political Econom*y 92 (1): 20–39.
Chamley, Christophe, 1986. “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives.” *Econometrica* 54 (3): 607–22.
Chari, V. V., Juan Pablo Nicolini, and Pedro Teles. 2020. “Optimal Capital Taxation Revisited.” *Journal of Monetary Economics* 116: 147–65.
Chen, Lizi, and Casey Rothschild. 2015. “Screening with Endogenous Preferences.” *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association* 108: 1–44.
Chetty, Raj. 2006. “A New Method of Estimating Risk Aversion.” *American Economic Review* 96 (5): 1821–34.
Chetty, Raj. 2009. “Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implication of Evasion and Avoidance.” *American Economic Journal: Economic Policy* 1 (2), 31–52.
Chetty, Raj. 2012. “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply.” *Econometrica* 80 (3): 969–1018.
Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. “Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records.” *Quarterly Journal of Economics* 126 (2): 749–804.
Chetty, Raj, John N. Friedman, and Emmanuel Saez. 2013. “Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings.” *American Economic Review* 103 (7): 2683–721.
Chetty, Raj, Adam Looney, and Kory Kroft. 2009. “Salience and Taxation: Theory and Evidence.” *American Economic Review* 99 (4): 1145–77.
Choné, Philippe, and Guy Laroque. 2005. “Optimal Incentives for Labor Force Participation.” *Journal of Public Economics* 89 (2–3): 395–425.
Choné, Philippe, and Guy Laroque. 2010. “Negative Marginal Tax Rates and Heterogeneity.” *American Economic Review* 100 (5): 2532–47.
Christiansen, Vidar. 1981. “Evaluation of Public Projects under Optimal Taxation.” *Review of Economic Studies* 48 (3): 447–57.
Comanor, William S., and Robert H. Smiley. 1975. “Monopoly and the Distribution of Wealth.” *Quarterly Journal of Economics* 89 (2): 177–94.
Congdon, William J., Jeffrey R. Kling, and Sendhil Mullainathan. 2009. “Behavioral Economics and Tax Policy.” *National Tax Journal* 62 (3): 375–86.
Corlett, W. J., and D. C. Hague, 1953. “Complementarity and the Excess Burden of Taxation.” *Review of Economic Studies* 21 (1), 21–30.
Cowell, Frank A. 1990. *Cheating the Government: The Economics of Evasion*. Cambridge, MA: MIT Press.
Craig, Ashley. Forthcoming. “Optimal Income Taxation with Spillovers from Employer

Learning." *American Economic Journal: Economic Policy.*
Craig, Ashley, and Joel Slemrod. 2022. "Tax Knowledge and Tax Manipulation: A Unifying Model." NBER Working Paper 30151.
Cullen, Julie Berry, and Roger H. Gordon. 2007. "Taxes and Entrepreneurial Risk-Taking: Theory and Evidence for the U.S." *Journal of Public Economics*. 91 (7–8): 1479–505.
Dahan, Momi, and Michel Strawczynski. 2000. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates: Comment." *American Economic Review* 90 (3): 681–86.
Dahan, Momi, and Michel Strawczynski. 2012. "The Optimal Asymptotic Income Tax Rate." *Journal of Public Economic Theory* 14 (5): 737–55.
Dasgupta, Partha, and Peter Hammond. 1980. "Fully Progressive Taxation." *Journal of Public Economics* 13 (2): 141–54.
de Bartolome, Charles A.M.. 1995. "Which Tax Rate Do People Use: Average or Marginal?" *Journal of Public Economics* 56 (1): 79–96.
De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications." *Quarterly Journal of Economics* 135 (2): 561–644.
Diamond, Peter A. 1968. "Negative Taxes and the Poverty Problem—A Review Article." *National Tax Journal* 21 (3): 288–303.
Diamond, Peter A. 1980. "Income Taxation with Fixed Hours of Work." *Journal of Public Economics* 13 (1): 101–10.
Diamond, Peter A. 1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review* 88 (1): 83–95.
Diamond, Peter A. 2004. "Social Security." *American Economic Review* 94 (1): 1–24.
Diamond, Peter A. 2006. "Optimal Tax Treatment of Private Contributions for Public Goods with and without Warm Glow Preferences." *Journal of Public Economics* 90 (4), 897–919.
Diamond, Peter A., and James A. Mirrlees. 1971. "Optimal Taxation and Public Production II: Tax Rules." *American Economic Review* 61 (3), 261–78.
Diamond, Peter A., and Emmanuel Saez. 2011. "The Case for a Progressive Tax: From Basic Research to Policy." *Journal of Economic Perspectives* 25 (4): 165–90.
Dixit, Avinash K., and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Product Diversity." *American Economic Review* 67 (3): 297–308.
Doligalski, Pawel, Abdoulaye Ndiaye, and Nicolas Werquin. 2020. "Redistribution with Performance Pay." (Unpublished.)
Domar, Evsey D., and Richard A. Musgrave. 1944. "Proportional Income Taxation and Risk-Taking." *Quarterly Journal of Economics* 58 (3), 388–423.
Domowitz, Ian, R. Glenn Hubbard, and Bruce C. Petersen. 1988. "Market Structure and Cyclical Fluctuations in U.S. Manufacturing." *Review of Economics and Statistics* 70 (1): 55–66.
Duesenberry, James S. 1949. *Income, Saving, and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press.
Dupor, Bill, and Wen-Fang Liu. 2003. "Jealousy and Equilibrium Overconsumption." *American Economic Review* 93 (1): 423–28.
Easterlin, Richard A. 1974. "Does Economic Growth Improve the Human Lot? Some Empirical Evidence." In *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*, edited by Paul A. David and Melvin W. Reder, 89–125. New York: Academic Press.
Ebert, Udo. 1992. "A Reexamination of the Optimal Nonlinear Income Tax." *Journal of Public Economics* 49 (1): 47–73.
Eeckhout, Jan, Chunyang Fu, Wenjian Li, and Xi Weng. 2021. "Optimal Taxation and Market

Power." (Unpublished.)
Eissa, Nada, Henrik Jacobsen Kleven, and Claus Thustrup Kreiner. 2008. "Evaluation of Four Tax Reforms in the United States: Labor Supply and Welfare Effects for Single Mothers." *Journal of Public Economics* 92 (3–4): 795–816.
Eissa, Nada, and Jeffrey B. Liebman. 1996. "Labor Supply Response to the Earned Income Tax Credit." *Quarterly Journal of Economics* 111 (2): 605–37.
Epper, Thomas, Ernst Fehr, Helga Fehr-Duda, Claus Thustrup Kreiner, David Dreyer Lassen, Søren Leth-Petersen, and Gregers Nytoft Rasmussen. 2020. "Time Discounting and Wealth Inequality." *American Economic Review* 110 (4): 1177–205.
Ericson, Richard, and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies* 62 (1): 53–82.
European Union. 2004. *Guidelines on the Assessment of Horizontal Mergers Under the Council Regulation on the Control of Concentrations Between Undertakings*.
Farhi, Emmanuel, and Xavier Gabaix. 2020. "Optimal Taxation with Behavioral Agents." *American Economic Review* 110 (1): 298–336.
Farhi, Emmanuel, and Iván Werning. 2010. "Progressive Estate Taxation." *Quarterly Journal of Economics* 125 (2): 635–73.
Feldstein, Martin S. 1969. "The Effects of Taxation on Risk Taking." *Journal of Political Economy* 77 (5): 755–64.
Feldstein, Martin S. 1973. "On the Optimal Progressivity of the Income Tax." *Journal of Public Economics* 2 (4): 357–76.
Feldstein, Martin S. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics* 81 (4): 674–80.
Feenstra, Robert C., and David E. Weinstein. 2017. "Globalization, Markups, and US Welfare." *Journal of Political Economy* 125 (4): 1040–74.
Fleurbaey, Marc, and François Maniquet, 2018. "Optimal Income Taxation Theory and Principles of Fairness." *Journal of Economic Literature* 56 (3): 1029–79.
Frank, Robert H. 1999. *Luxury Fever: Money and Happiness in an Era of Excess*. Princeton: Princeton University Press.
Frankel, Alexander. 2014. "Taxation of Couples under Assortative Mating." *American Economic Journal: Economic Policy* 6 (3): 155–77.
Fullerton, Don. 1991. "Reconciling Recent Estimates of the Marginal Welfare Cost of Taxation." *American Economic Review* 81 (1): 302–08.
Garfinkel, Irwin, Laurel Sariscsany, Elizabeth Ananat, Sophie M. Collyer, Robert Paul Hartley, Buyi Wang, and Christopher Wimer. 2022. "The Benefits and Costs of a U.S. Child Allowance." NBER Working Paper 29854.
Gayle, George-Levi, and Andrew Shephard. 2019. "Optimal Taxation, Marriage, Home Production, and Family Labor Supply." *Econometrica* 87 (1): 291–326.
Gelber, Alexander M., Damon Jones, and Daniel W. Sacks. 2020. "Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test." *American Economic Journal: Applied Economics* 12 (1): 1–31.
Gentry, William M. 2016. "Capital Gains Taxation and Entrepreneurship." *Tax Law Review* 69 (3): 321–55.
Goldin, Jacob. 2015. "Optimal Tax Salience." *Journal of Public Economics* 131: 115–23.
Golosov, Mikhail, Maxim Troshkin, Aleh Tsyvinski, and Matthew C. Weinzierl. 2013. "Preference Heterogeneity and Optimal Capital Income Taxation." *Journal of Public Economics* 97: 160–75.

Golosov, Mikhail, Aleh Tsyvinski, and Iván Werning, 2007. "New Dynamic Public Finance: A User's Guide." *NBER Macroeconomics Annual 2006* 21: 317–87.
Gordon, Roger H. 1985. "Taxation of Corporate Capital Income: Tax Revenues versus Tax Distortions." *Quarterly Journal of Economics* 100 (1): 1–27.
Green, Christopher. 1967. *Negative Taxes and the Poverty Problem*. Washington, DC: Brookings Institution.
Gruber, Jon, and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics* 84 (1): 1–32.
Hall, Robert E. 2018. "New Evidence on the Markup of Prices over Marginal Costs and the Rule of Mega-Firms in the U.S. Economy." NBER Working Paper 24574.
Hall, Robert E., and Susan E. Woodward. 2010. "The Burden of the Nondiversifiable Risk of Entrepreneurship." *American Economic Review* 100 (3): 1163–94.
Hammond, Peter. 1983. "Ex-post Optimality as a Dynamically Consistent Objective for Collective Choice under Uncertainty." In *Social Choice and Welfare*, edited by Prasanta K. Pattanaik and Maurice Salles, 175–205. New York: Elsevier.
Hare, Richard M. 1973. "Rawls' Theory of Justice—I." *Philosophical Quarterly* 23: 144–55.
Hare, Richard M. 1981. *Moral Thinking: Its Levels, Method, and Point*. New York: Oxford University Press.
Harrod, Roy F. 1936. "Utilitarianism Revised." *Mind* 45: 137–56.
Harsanyi, John C. 1953. "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking." *Journal of Political Economy* 61 (5): 434–35.
Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy* 63 (4): 309–21.
Hart, Oliver. 1982. "A Model of Imperfect Competition with Keynesian Features." *Quarterly Journal of Economics* 97 (1): 109–38.
Hausman, Jerry. 1985. "Taxes and Labor Supply." In *Handbook of Public Economics*, Vol. 1, edited by Alan J. Auerbach and Martin Feldstein, 213–63. Amsterdam: North-Holland.
Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework." *Quarterly Journal of Economics* 132 (4): 1693–754.
Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2020. "Optimal Progressivity with Age-dependent Taxation." *Journal of Public Economics* 189: 104074.
Hendren, Nathaniel. 2016. "The Policy Elasticity." *Tax Policy and the Economy* 30: 51–89.
Hendren, Nathaniel. 2020. "Measuring Economic Efficiency Using Inverse-Optimum Weights." *Journal of Public Economics* 187: 104198.
Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A Unified Welfare Analysis of Government Policies." *Quarterly Journal of Economics* 135 (3): 1209–318.
Hochman, Harold M., and James D. Rodgers. 1969. "Pareto Optimal Redistribution." *American Economic Review* 59 (4): 542–57.
Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (5): 1127–50.
Hume, David. 1751 [1998]. *An Enquiry Concerning the Principles of Morals*. Edited by Tom L. Beauchamp. Oxford: Clarendon Press.
Hylland, Aanund, and Richard Zeckhauser. 1979. "Distributional Objectives Should Affect Taxes but Not Program Choice or Design." *Scandinavian Journal of Economics* 81 (2): 264–84.
Ireland, Norman J. 1981. "Optimal Income Tax in the Presence of Status Effects." *Journal of Public Economics* 81 (2): 193–212.

Jacquet, Laurence, and Etienne Lehmann. Forthcoming. "Optimal Tax Problems with Multidimensional Heterogeneity: A Mechanism Design Approach." *Social Choice and Welfare*.
Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden. 2013. "Optimal Redistributive Taxation with Both Extensive and Intensive Responses." *Journal of Economic Theory* 148 (5): 1770–805.
Johnson, Stephen, Laurence J. Kotlikoff, and William Samuelson. 2001. "Can People Compute? An Experimental Test of the Life Cycle Consumption Model." In *Essays on Saving, Bequests, Altruism, and Life-Cycle Planning*, edited by Laurence J. Kotlikoff, 335–85. Cambridge, MA: MIT Press.
Jones, Charles I. Forthcoming. "Taxing Top Incomes in a World of Ideas." *Journal of Political Economy*.
Jonsson, Magnus. 2007. "The Welfare Cost of Imperfect Competition and Distortionary Taxation." *Review of Economic Dynamics* 10 (4): 576–94.
Judd, Kenneth L. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics* 28 (1): 59–83.
Kaplow, Louis. 1989. "Horizontal Equity: Measures in Search of a Principle." *National Tax Journal* 42 (2): 139–54.
Kaplow, Louis. 1991. "Incentives and Government Relief for Risk." *Journal of Risk and Uncertainty* 4 (2): 167–75.
Kaplow, Louis. 1994. "Taxation and Risk Taking: A General Equilibrium Perspective." *National Tax Journal* 47 (4): 789–98.
Kaplow, Louis. 1995. "A Fundamental Objection to Tax Equity Norms: A Call for Utilitarianism." *National Tax Journal* 48 (4): 497–514.
Kaplow, Louis. 1996a. "Optimal Distribution and the Family." *Scandinavian Journal of Economics* 98 (1): 75–92.
Kaplow, Louis. 1996b. "The Optimal Supply of Public Goods and the Distortionary Cost of Taxation." *National Tax Journal* 49 (4): 513–33.
Kaplow, Louis. 1998a. "Accuracy, Complexity, and the Income Tax." *Journal of Law, Economics, & Organization* 14 (1): 61–83.
Kaplow, Louis. 1998b. "Tax and Non-tax Distortions." *Journal of Public Economics* 68 (2): 303–06.
Kaplow, Louis. 1998c. "Tax Policy and Gifts." *American Economic Review Papers and Proceedings* 88 (2): 283–88.
Kaplow, Louis. 2001. "A Framework for Assessing Estate and Gift Taxation." In *Rethinking Estate and Gift Taxation*, edited by William G. Gale, James R. Hines Jr., and Joel Slemrod, 164–215. Washington, D.C.: Brookings Institution Press.
Kaplow, Louis. 2004. "On the (Ir)Relevance of Distribution and Labor Supply Distortion to Government Policy." *Journal of Economic Perspectives* 18 (4): 159–75.
Kaplow, Louis. 2005. "Why Measure Inequality?" *Journal of Economic Inequality* 3 (1): 65–79.
Kaplow, Louis. 2006a. "On the Undesirability of Commodity Taxation Even When Income Taxation Is Not Optimal." *Journal of Public Economics* 90 (6–7): 1235–50.
Kaplow, Louis. 2006b. "Public Goods and the Distribution of Income." *European Economic Review* 50: 1627–60.
Kaplow, Louis. 2007a. "Optimal Income Transfers." *International Tax and Public Finance* 14 (3): 295–325.
Kaplow, Louis. 2007b. "Primary Goods, Capabilities, or . . . Well-Being?" *Philosophical Review* 116 (4): 603–32.

Kaplow, Louis. 2008a. *The Theory of Taxation and Public Economics*. Princeton, NJ: Princeton University Press.
Kaplow, Louis. 2008b. "Optimal Policy with Heterogeneous Preferences." *B.E. Journal of Economic Analysis and Policy: Advances* 8 (1: 40): 1–28.
Kaplow, Louis. 2010a. "Concavity of Utility, Concavity of Welfare, and Redistribution of Income." *International Tax and Public Finance* 17 (1): 25–42.
Kaplow, Louis. 2010b. "Taxing Leisure Complements." *Economic Inquiry* 48 (4): 1065–71.
Kaplow, Louis. 2011. "Targeted Savings and Labor Supply." *International Tax and Public Finance* 18 (5): 507–18.
Kaplow, Louis. 2012. "Optimal Control of Externalities in the Presence of an Income Tax." *International Economic Review* 53 (2): 487–509.
Kaplow, Louis. 2015a. "Government Policy and Labor Supply with Myopic or Targeted Savings Decisions." *Tax Policy and the Economy* 29: 159–93.
Kaplow, Louis. 2015b. "Myopia and the Effects of Social Security and Capital Taxation on Labor Supply." *National Tax Journal* 68 (1): 7–32.
Kaplow, Louis. 2020a. "A Unified Perspective on Efficiency, Redistribution, and Public Policy." *National Tax Journal* 73 (2): 429–72.
Kaplow, Louis. 2020b. "Competition Policy in a Simple General Equilibrium Model." NBER Working Paper 24842.
Kaplow, Louis. 2021. "Market Power and Income Taxation." *American Economic Journal: Economic Policy* 13 (4): 329–54.
Kaplow, Louis, and Steven Shavell. 2001. "Any Non-Welfarist Method of Policy Assessment Violates the Pareto Principle." *Journal of Political Economy* 109 (2): 281–86.
Kaplow, Louis, and Steven Shavell. 2002. *Fairness versus Welfare*. Cambridge, MA: Harvard University Press.
Kaplow, Louis, and Steven Shavell. 2007. "Moral Rules, the Moral Sentiments, and Behavior: Toward a Theory of an Optimal Moral System." *Journal of Political Economy* 115 (3): 494–514.
Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. 2015. "Self-Control at Work." *Journal of Political Economy* 123 (6): 1227–77.
Keane, Michael P. 2011. "Labor Supply and Taxes: A Survey." *Journal of Economic Literature* 49 (4): 961–1075.
Keane, Michael P., and Richard Rogerson. 2012. "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom." *Journal of Economic Literature* 50 (2): 464–76.
Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as Experimentation." *Journal of Economic Perspectives* 28 (3): 25–48.
Kleven, Henrik Jacobsen. 2021. "The EITC and the Extensive Margin: A Reappraisal." NBER Working Paper 26405 (revised).
Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez. 2009. "The Optimal Income Taxation of Couples." *Econometrica* 77 (2): 537–60.
Konishi, Hideo. 1995. "A Pareto-Improving Commodity Tax Reform under a Smooth Nonlinear Income Tax." *Journal of Public Economics* 56 (3): 413–46.
Kopczuk, Wojciech. 2013. "Incentive Effects of Inheritances and Optimal Estate Taxation." *American Economic Review Papers and Proceedings* 103 (3): 472–77.
Kosar, Gizem, and Robert A. Moffitt. 2017. "Trends in Cumulative Marginal Tax Rates Facing Low-Income Families, 1997–2007." *Tax Policy and the Economy* 31: 43–70.
Kremer, Michael. 2002. "Should Taxes be Independent of Age." (Unpublished.)
Kushnir, Alexey, and Robertas Zubrickas. 2021. "Optimal Income Taxation with Endogenous

Prices." (Unpublished.)
Laibson, David I. 1996. "Hyperbolic Discount Functions, Undersaving, and Savings Policy." NBER Working Paper 5635.
Laibson, David I. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics* 112 (2): 443–78.
Laroque, Guy R. 2005. "Indirect Taxation Is Superfluous under Separability and Taste Homogeneity: A Simple Proof." *Economics Letters* 87 (1): 141–44.
Layard, Richard. 1980. "Human Satisfactions and Public Policy." *Economic Journal* 90: 737–50.
Leland, Hayne E., and David H. Pyle. 1977. "Informational Asymmetries, Financial Structure, and Financial Intermediation." *Journal of Finance* 32 (2): 371–87.
Lerner, A.P. 1934. "The Concept of Monopoly and the Measurement of Monopoly Power." *Review of Economic Studies* 1 (3): 157–75.
Liebman, Jeffrey B., and Richard J. Zeckhauser. 2004. "Schmeduling." (Unpublished.)
Lindahl, Erik, 1919 [1958]. "Die Gerechtigkeit der Besteuerung: Eine Analyse der Steuerprinzipien auf der Grundlage der Grenznutzentheorie. Lund: Gleerup and H. Ohlsson." In *Classics in the Theory of Public Finance*, edited by Richard A. Musgrave and Alan Peacock, 168–76 (Chapter 4, "Positive Lösung," translated as "Just Taxation—A Positive Solution"). London: Macmillan.
Lipsey, Richard G., and Kelvin Lancaster. 1956. "The General Theory of Second Best." *Review of Economic Studies* 24 (1): 11–32.
Lockwood, Benjamin B. 2020. "Optimal Income Taxation with Present Bias." *American Economic Journal: Economic Policy* 12 (4): 298–327.
Lockwood, Benjamin B., Charles G. Nathanson, and E. Glen Weyl. 2017a. "Taxation and the Allocation of Talent." *Journal of Political Economy* 125 (5): 1635–82.
Lockwood, Benjamin B., Charles G. Nathanson, and E. Glen Weyl. 2017b. "What If Socially Useful Jobs Were Taxed Less Than Other Jobs?" *Harvard Business Review* (October 11).
Lockwood, Benjamin B., and Matthew C. Weinzierl. 2015. "De Gustibus Non Est Taxandum: Heterogeneity in Preferences and Optimal Redistribution." *Journal of Public Economics* 124: 74–80.
Mankiw, N. Gregory, and Michael D. Whinston. 1986. "Free Entry and Social Inefficiency." *Rand Journal of Economics* 17 (1): 48–58.
Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–725.
Mertens, Karel, and José Luis Montiel Olea. 2018. "Marginal Tax Rates and Income: New Time Series Evidence." *Quarterly Journal of Economics* 133 (4): 1803–84.
Meyer, Bruce D. 2010. "The Effects of the Earned Income Tax Credit and Recent Reforms." *Tax Policy and the Economy* 24: 153–80.
Meyer, Bruce D., and Dan T. Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics* 116 (3): 1063–114.
Michalopoulos, Charles, Philip K. Robins, and David Card. 2005. "When Financial Work Incentives Pay for Themselves: Evidence from a Randomized Social Experiment for Welfare Recipients." *Journal of Public Economics* 89 (1): 5–29.
Michau, Jean-Baptiste. 2014. "Optimal Redistribution: A Life-Cycle Perspective." *Journal of Public Economics* 111: 1–16.
Mill, John Stuart. 1861 [1998]. *Utilitarianism*. Edited by Roger Crisp. New York: Oxford University Press.
Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38 (2): 175–208.

Mirrlees, James A. 1976. "Optimal Tax Theory: A Synthesis." *Journal of Public Economics* 6 (4): 327–58.
Mirrlees, James A. 1982. "The Economic Uses of Utilitarianism." In *Utilitarianism and Beyond*, edited by Amartya Sen and Bernard Williams, 63–84. Cambridge, UK: Cambridge University Press.
Mirrlees, James A. 1994. "Optimal Taxation and Government Finance." In *Modern Public Finance*, edited by John M. Quigley and Eugene Smolensky, 213–31. Cambridge, MA: Harvard University Press.
Moffitt, Robert A. 2002. "Welfare Programs and Labor Supply." In *Handbook of Public Economics*, Vol. 4, edited by Alan J. Auerbach and Martin Feldstein, 2393–430. Amsterdam: Elsevier.
Moore, Dylan T., and Joel Slemrod. 2020. "Optimal Tax Systems with Endogenous Behavioral Biases." (Unpublished.)
Moser, Christian, and Pedro Olea de Souza e Silva. 2019. "Optimal Paternalistic Savings Policies." *Federal Reserve Bank of Minneapolis Institute Working Paper 17*.
Musgrave, Richard A. 1959. *The Theory of Public Finance*. New York: McGraw-Hill.
Myerson, Roger B. 1981. "Utilitarianism, Egalitarianism, and the Timing Effect in Social Choice Problems." *Econometrica* 49 (4): 883–97.
Myles, Gareth D. 1995. *Public Economics*. Cambridge, UK: Cambridge University Press.
Ng, Yew-Kwang. 1981. "Bentham or Nash? On the Acceptable Form of Social Welfare Functions." *Economic Record* 57: 238–50.
Nordhaus, William D. 2004. "Schumpeterian Profits in the American Economy: Theory and Measurement." NBER Working Paper 10433.
O'Donoghue, Ted, and Matthew Rabin. 2006. "Optimal Sin Taxes." *Journal of Public Economics* 90 (10): 1825–49.
OECD. 2017. *Inequality: A Hidden Cost of Market Power*.
Oswald, Andrew J. 1983. "Altruism, Jealousy, and the Theory of Optimal Non-Linear Taxation." *Journal of Public Economics* 20 (1): 77–87.
Pauly, Mark V. 1973. "Income Redistribution as a Local Public Good." *Journal of Public Economics* 2 (1): 35–58.
Pauly, Mark V. 1974. "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection." *Quarterly Journal of Economics* 88 (1): 44–62.
Phelps, Edmund S. 1973. "Taxation of Wage Income for Economic Justice." *Quarterly Journal of Economics* 87 (3): 331–54.
Piketty, Thomas, and Emmanuel Saez. 2013. "Optimal Labor Income Taxation." In *Handbook of Public Economics*, Vol. 5, edited by Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, 391–474. Amsterdam: Elsevier.
Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva. 2014. "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." *American Economic Journal: Economic Policy* 6 (1): 230–71.
Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *Economic Journal* 37: 47–61.
Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
Rawls, John. 1982. "Social Unity and Primary Goods." In *Utilitarianism and Beyond*, edited by Amartya Sen and Bernard Williams, 159–85. New York: Cambridge University Press.
Reck, Daniel. 2016. "Taxes and Mistakes: What's in a Sufficient Statistic?" (Unpublished.)
Rees-Jones, Alex, and Dmitry Taubinsky. 2020. "Measuring 'Schmeduling.'" *Review of Economic Studies* 87 (5): 2399–438.

Restuccia, Diego, and Guillaume Vandenbroucke. 2013. “A Century of Human Capital and Hours.” *Economic Inquiry* 51 (3): 1849–66.
Robinson, Joan. 1933. *The Economics of Imperfect Competition*. London: Macmillan.
Rognlie, Matthew. 2015. “Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity?” *Brookings Papers on Economic Activity* 1–54.
Ross, Stephen A. 1977. “The Determination of Financial Structure: The Incentive-Signaling Approach.” *Bell Journal of Economics* 8 (1): 23–40.
Roth, Jeffrey A., John T. Scholz, and Ann D. Witte, eds. 1989. *Taxpayer Compliance:* Vol. 1, *An Agenda for Research;* Vol. 2, *Social Science Perspectives*. Philadelphia: University of Pennsylvania Press.
Rothschild, Casey, and Florian Scheuer. 2013. “Redistributive Taxation in the Roy Model.” *Quarterly Journal of Economics* 128 (2): 623–68.
Rothschild, Casey, and Florian Scheuer. 2014. “A Theory of Income Taxation with Multidimensional Skill Heterogeneity.” NBER Working Paper 19822.
Rothschild, Casey, and Florian Scheuer. 2016. “Optimal Taxation with Rent-Seeking.” *Review of Economic Studies* 83 (3): 1225–62.
Rothstein, Jesse. 2010. “Is the EITC as Good as an NIT? Conditional Cash Transfers and Tax Incidence.” *American Economic Journal: Economic Policy* 2 (1): 177–208.
Roy, Andrew D. 1951. “Some Thoughts on the Distribution of Earnings.” *Oxford Economic Papers* 3 (2): 135–46.
Sachs, Dominik, Aleh Tsyvinski, and Nicolas Werquin. 2020. “Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium.” *Econometrica* 88 (2): 469–93.
Sadka, Efraim. 1976. “On Income Distribution, Incentive Effects and Optimal Income Taxation.” *Review of Economic Studies* 43 (2): 261–67.
Saez, Emmanuel. 2001. “Using Elasticities to Derive Optimal Income Tax Rates.” *Review of Economic Studies* 68 (1): 205–29.
Saez, Emmanuel. 2002a. “Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses.” *Quarterly Journal of Economics* 117 (3): 1039–73.
Saez, Emmanuel. 2002b. “The Desirability of Commodity Taxation under Non-linear Income Taxation and Heterogeneous Tastes.” *Journal of Public Economics* 83 (2): 217–30.
Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz. 2012. “The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review.” *Journal of Economic Literature* 50 (1): 3–50.
Saez, Emmanuel, and Stefanie Stantcheva. 2016. “Generalized Social Marginal Welfare Weights for Optimal Tax Theory.” *American Economic Review* 106 (1): 24–45.
Salanié, Bernard, 2003. *The Economics of Taxation*. Cambridge, MA: MIT Press.
Salanié, Bernard, 2011. *The Economics of Taxation*, 2nd ed. Cambridge, MA: MIT Press.
Samuelson, Paul A., 1954. “The Pure Theory of Public Expenditure.” *Review of Economics and Statistics* 36 (4): 387–89.
Schanzenbach, Diane Whitmore, and Michael R. Strain. 2020. “Employment Effects of the Earned Income Tax Credit: Taking the Long View.” NBER Working Paper 28041.
Scheuer, Florian. 2013. “Adverse Selection in Credit Markets and Regressive Profit Taxation.” *Journal of Economic Theory* 148 (4): 1333–60.
Scheuer, Florian. 2014. “Entrepreneurial Taxation with Endogenous Entry.” *American Economic Journal: Economic Policy* 6 (2): 126–63.
Scheuer, Florian, and Nick Netzer. 2007. “Taxation, Insurance, and Precautionary Labor.” *Journal of Public Economics* 91 (7): 1519–31.
Scheuer, Florian, and Joel Slemrod. 2020. “Taxation and the Superrich.” *Annual Review of*

Economics 12: 189–211.
Scheuer, Florian, and Iván Werning. 2017. "The Taxation of Superstars." *Quarterly Journal of Economics* 132 (1): 211–70.
Scheuer, Florian, and Iván Werning. 2018. "Mirrlees Meets Diamond-Mirrlees: Simplifying Nonlinear Income Taxation." (Unpublished.)
Schumpeter, Joseph A. 1947. "The Creative Response in Economic History." *Journal of Economic History* 7 (2): 149–59.
Seade, J. K. 1977. "On the Shape of Optimal Tax Schedules." *Journal of Public Economics* 7 (2): 203–35.
Sen, Amartya. 1970. "The Impossibility of a Paretian Liberal." *Journal of Political Economy* 78 (1): 152–57.
Sen, Amartya. 1985a. *Commodities and Capabilities*. Amsterdam: North-Holland.
Sen, Amartya. 1985b. "Well-Being, Agency and Freedom: The Dewey Lectures 1984." *Journal of Philosophy* 82 (4): 169–221.
Sen, Amartya. 1997. *On Economic Inequality*, enlarged ed. Oxford: Clarendon Press.
Shavell, Steven. 1981. "A Note on Efficiency vs. Distributional Equity in Legal Rulemaking: Should Distributional Equity Matter Given Optimal Income Taxation?" *American Economic Review Papers and Proceedings* 71 (2): 414–18.
Sher, Itai. 2021. "Generalized Social Marginal Welfare Weights Imply Inconsistent Comparisons of Tax Policies." (Unpublished.)
Sheshinski, Eytan. 1989. "Note on the Shape of the Optimal Income Tax Schedule." *Journal of Public Economics* 40 (2): 201–15.
Sidgwick, Henry. 1907 [1981]. *The Methods of Ethics*, reprint of 7th ed. Indianapolis: Hackett Publishing Co.
Simon, Herbert A. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106 (6): 467–82.
Simon, Herbert A. 1978. "Rationality as Process and as Product of Thought." *American Economic Review* 68 (2): 1–16.
Slemrod, Joel. 2019. "Tax Compliance and Enforcement." *Journal of Economic Literature* 57 (4): 904–54.
Slemrod, Joel, and Christian Gillitzer. 2014. *Tax Systems*. Cambridge, MA: MIT Press.
Slemrod, Joel, and Wojciech Kopczuk. 2002. "The Optimal Elasticity of Taxable Income." *Journal of Public Economics* 84 (1): 91–112.
Slemrod, Joel, and Shlomo Yitzhaki. 2001. "Integrating Expenditure and Tax Decisions: The Marginal Costs of Funds and the Marginal Benefit of Projects." *National Tax Journal* 54 (2): 189–202.
Slemrod, Joel, and Shlomo Yitzhaki. 2002. "Tax Avoidance, Evasion, and Administration." In *Handbook of Public Economics*, Vol. 3, edited by Alan J. Auerbach and Martin Feldstein, 1423–70. Amsterdam: Elsevier.
Slemrod, Joel, Shlomo Yitzhaki, Joram Mayshar, and Michael Lundholm. 1994. "The Optimal Two-Bracket Linear Income Tax." *Journal of Public Economics* 53 (2): 269–90.
Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick. 2019. "Capitalists in the Twenty-First Century." *Quarterly Journal of Economics* 134 (4): 1675–745.
Stantcheva, Stefanie. 2014. "Optimal Income Taxation with Adverse Selection in the Labour Market." *Review of Economic Studies* 81 (3): 1296–329.
Stantcheva, Stefanie. 2020. "Dynamic Taxation." *Annual Review of Economics* 12: 801–31.
Stevenson, Betsey, and Justin Wolfers. 2008. "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox." *Brookings Papers on Economic Activity* 1–87.

Stiglitz, Joseph E. 1976. "Simple Formulae for Optimal Income Taxation and the Measurement of Inequality." Institute for Mathematical Studies in the Social Sciences Technical Report 215.

Stiglitz, Joseph E. 1982. "Self-Selection and Pareto Efficient Taxation." *Journal of Public Economics* 17 (2): 213–40.

Stiglitz, Joseph E. 1987. "Pareto Efficient and Optimal Taxation and the New New Welfare Economics." In *Handbook of Public Economics*, Vol. 2, edited by Alan J. Auerbach and Martin Feldstein, 991–1042. Amsterdam: North-Holland.

Straub, Ludwig, and Iván Werning. 2020. "Positive Long Run Capital Taxation: Chamley-Judd Revisited." *American Economic Review* 110 (1): 86–119.

Syverson, Chad. 2019. "Macroeconomics and Market Power: Context, Implications, and Open Questions." *Journal of Economic Perspectives* 33 (3): 23–43.

Thaler, Richard H., and Shlomo Benartzi. 2004. "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy* 112 (1) (pt. 2): S164–87.

Tuomala, Matti. 2016. *Optimal Redistributive Taxation*. Oxford: Oxford University Press.

U.S. Department of Justice and Federal Trade Commission. 2010. *Horizontal Merger Guidelines*.

Veblen, Thorstein. 1899. *The Theory of the Leisure Class: An Economic Study of Institutions*. New York: Macmillan Company.

Vickrey, William. 1939. "Averaging of Income for Income-Tax Purposes." *Journal of Political Economy* 47 (3): 379–97.

Vickrey, William, 1945. "Measuring Marginal Utility by Reactions to Risk." *Econometrica* 13 (4): 319–33.

Wane, Waly. 2001. "The Optimal Income Tax When Poverty Is a Public 'Bad.'" *Journal of Public Economics* 82 (2): 271–99.

Weinzierl, Matthew C. 2011. "The Surprising Power of Age-Dependent Taxes." *Review of Economic Studies* 78 (4): 1490–518.

Weinzierl, Matthew C. 2017. "A Welfarist Role for Nonwelfarist Rules." NBER Working Paper 23587.

Werning, Iván. 2007. "Pareto Efficient Income Taxation." (Unpublished.)

World Bank Group and OECD. 2017. *A Step Ahead: Competition Policy for Shared Prosperity and Inclusive Growth*.